David B. Neale
Nicholas C. Wheeler

The Conifers: Genomes, Variation and Evolution

Springer

The Conifers: Genomes, Variation and Evolution

David B. Neale • Nicholas C. Wheeler

The Conifers: Genomes, Variation and Evolution

Springer

David B. Neale
Department of Plant Sciences
University of California, Davis
Davis, CA, USA

Nicholas C. Wheeler
Consultant
Centralia, WA, USA

ISBN 978-3-319-46806-8 ISBN 978-3-319-46807-5 (eBook)
https://doi.org/10.1007/978-3-319-46807-5

Library of Congress Control Number: 2019931904

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

The authors of this volume are sincerely pleased to dedicate our work to the many postdoctoral scholars, graduate students, research assistants, and undergraduate interns who provided the hard work, enthusiasm, and creativity to so many research projects the authors collaborated on over the years. Thanks to you all, we look forward to your ongoing contributions to our understanding of the genetics of forest trees.

Foreword

Conifers, and related gymnosperms, are among the world's oldest groups of organisms. I learned that in my first year at university while training as a forester, and even now, as a forest geneticist nearing the end of my scientific career, I still find this fascinating. Conifers evolved during the era of the dinosaurs and continue to dominate large expanses of forest around the world, even after the subsequent evolution of the angiosperms, which are exceedingly more species-diverse. What was in the DNA of conifers that made them so durable? In *The Conifers: Genomes, Variation, and Evolution*, David Neale and Nicholas Wheeler review and synthesize findings from traditional genetic studies and the most recent molecular genetic research that helps elucidate why conifers may have persisted for millions of years.

Humans have long valued conifers for their exceptional wood properties. Straighter in form, and possessing stronger yet lighter wood than most angiosperms, they have been important construction materials throughout the world for thousands of years. Today, conifers are an important forest resource in many countries. Early in the twentieth century, scientists recognized the significance of conifer genetics to the practice of forestry. Common garden studies by pioneering geneticists showed that many conifer species had considerable levels of genetic variation, making them highly suitable to "improvement" using classic breeding techniques developed on domestic plants and animals. Many traditional conifer breeding programs were subsequently initiated, practicing recurrent selection in order to improve desired traits — such as growth, resistance to disease and insects, and, more recently, adaptability to changing environments. In the 1980s, molecular genetics studies of forest trees also became a key element to understanding their basic biology. Using the technique of protein electrophoresis, these molecular approaches enhanced our knowledge of conifer genetics and became instrumental to understanding conifer mating dynamics, informing seed production practices, and developing forest tree genetic resource management programs.

As molecular biology techniques improved through the 1990s and onward, significant resources were invested in molecular genetics and genomics research due, in large part, to the progress and promises of the human genome project. While traditional forest genetics and tree breeding research waned around the world, largely due to fundamental shifts in the economics of forestry investments, forest tree molecular geneticists were keeping pace with the genomic advances occurring in most other organisms (e.g., *Arabidopsis* and humans). During the turn of this

century, study objectives changed to more scholarly questions, along with divergence in the terminology and vernacular used by forest tree breeders and forest tree molecular geneticists. To address this, Neale and Wheeler have neatly clarified the terms used in modern molecular genetics of conifers and, more importantly, have presented the most recent information of conifer genetics in a modern synthesis that integrates many aspects of their quantitative and population genetics. The book is presented in an understandable way, highlighting recent breakthroughs in gene structure and gene families, comparative genomics, phylogenetics, and landscape genomics, for example, but also identifies the interesting challenges ahead in genomics research of these marvelous organisms. Any new forest geneticist must now be versed in both classical and modern population genetics and quantitative approaches, as well as molecular genetics terminology, techniques, and bioinformatics of genomics. *The Conifers: Genomes, Variation, and Evolution* provides a much needed unification of these topics and should be required reading for new students of conifer molecular biology and genomics.

British Columbia Ministry of Forests, Lands,
Natural Resource Operations and Rural Development
Forest Improvement and Research Management Branch,
Victoria, BC, Canada

Alvin Yanchuk

Preface

The study of the genetics of forest trees began more than 100 years ago, coincident with the rediscovery of Mendel's classic works. For three quarters of the twentieth century, conifer genetics was dominated by the study of phenotypic variation in provenance trials (Chap. 8), species hybridization trials (Chap. 15), disease and insect resistance (Chap. 14), and development of tree breeding methods, all of which suggested most traits in trees were quantitatively inherited. Furthermore, many traits were adaptive and varied in response to environmental factors such as temperature and moisture.

Forest genetics research approaches, funding sources, and personnel began changing significantly in the mid-1970s as biochemical and DNA marker development allowed for population genetic and mating system investigations. These neutral markers were subsequently used to characterize the genetic basis of quantitative traits (Chaps. 9 and 11) and enhance our understanding of conservation genetics (Chap. 13) and phylogenetics (Chap. 16). Recent developments in high-throughput genome sequencing technology have resulted in a quantum leap in our knowledge of conifer genomes (Chaps. 2, 3, 4, 5, 6, 7, 12, and 17). We anticipate this knowledge will continue to increase rapidly, as it has for so many other organisms. A review of the state of knowledge in conifer forest genetics therefore seems appropriate.

In this monograph, our goal is to summarize and synthesize this body of work, specifically for conifers, in a manner that would be useful for practicing professionals in conifer genetics and genomics research but also for those from other fields of forestry and plant biology who might be curious as to what has been learned over 100+ years in this small discipline. The monograph is organized into three major parts, Genomes, Variation, and Evolution, and we have tried to synthesize and cross-reference across all chapters and sections. We have not included the very extensive literature in applied conifer breeding and tree improvement research, and instead refer readers to other important works in this area (Wright 1976; Namkoong 1979; Zobel and Talbert 1984; Mandal and Gibson 1998; Kumar and Fladung 2004; White et al. 2007; Plomion et al. 2011).While our search and summary of the relevant literature is certainly not exhaustive, we have sought to provide a comprehensive view with a modest (over 1600) number of exemplary citations.

Finally, the authors of this monograph have enjoyed and been rewarded by our years of making small contributions to this discipline of forest genetics working

individually and for much of our careers together. We have benefitted from knowing and working with a substantial number of the scientists, past and present, who have defined the forest genetics discipline. We are using this moment to give a little something back to the discipline that has given us so much. It is our perspective on the evolution of our discipline over the last 100+ years (Wheeler et al. 2015) that has shaped our final chapter (Chap. 18).

Davis, CA, USA — David B. Neale
Centralia, WA, USA — Nicholas C. Wheeler

References

Kumar, S. & Fladung, M. (2004). *Molecular genetics of breeding of forest trees*. Food Products Press, New York, London, Oxford, 436 p.

Mandal, A. K. & Gibson, G. L. (1998). *Forest genetics and tree breeding*. New Delhi: CBS Publishers and Distributors.

Namkoong, G.(1979). *Introduction to quantitative genetics in forestry*. United States Department of Agriculture Technical Bulletin No. 1588.

Plomion, C., Bousquet, J., & Kole C. (eds). (2011). *Genetics, genomics and breeding of conifers*. New Hampshire: Science Publishers.

Wheeler, N. C., Steiner, K. C., Schlarbaum, S. E., & Neale, D. B. (2015) The evolution of forest genetics and tree improvement research in the United States. *Journal of Forestry 113*(5), 500–510.

White, T. L., Adams, W. T., & Neale, D. B. (2007). *Forest genetics*. Cambridge, MA: CABI Publishing.

Wright, J. W. (1976). *Introduction to forest genetics*. New York: Academic Press, 463 pp.

Zobel, B., & Talbert, J. (1984). *Applied forest tree improvement*. New York: Wiley.

Acknowledgments

We would like to thank and acknowledge our colleagues who generously gave their time to review and edit one or more chapters. They include Joerg Bohlmann, Bruce Bongarten, Jean Bousquet, Don Dickmann, Christopher Earle, Yousry El-Kassaby, David Gernandt, Glenn Howe, Richard Sniezko, Brad St. Claire, Ross Whetten, Alvin Yanchuk, and Sam Yeaman.

We owe a huge debt of gratitude to the following people for so many things including literature searches, the drafting of tables and figures, edits of early draft versions of chapters, and the endless number of small tasks that allowed us to remain focused on the writing: Brian Allen, Greg deVeer, Emily Grau, Pat McGuire, and Valery Wyant-Wheeler.

Contents

The Conifers

1

The conifers are a diverse and ancient group of seed plants of monophyletic origin that arose more than 300 million years ago (Rothwell and Scheckler 1988). They are uniformly distinguished by their naked or exposed ovules during pollination, a trait they share with the cycads, the monotypic genus *Ginkgo*, and the gnetophytes. Collectively, these four taxa comprise the gymnosperms, which, along with the flowering plants (angiosperms), constitute the seed-bearing plants (Fig. 1.1).

The conifers are not easily circumscribed or defined. While there are many traits commonly associated with conifers, exceptions to most character states exist. Conifers are typically single-stemmed, evergreen trees with separate male (simple) and female (compound or reduced) reproductive structures (strobili or cones) either on the same (monoecy) or different plants (dioecy). Conifer wood possesses tracheids, bordered pits, and frequently, resin canals. Conifer lignin polymers are comprised almost entirely of coniferyl alcohol monolignol monomers (guaiacyl lignin), lacking syringyl elements common to most angiosperms (Boerjan et al. 2003). Leaves are simple, mostly needle- or scale-like, and often resinous (Fig. 1.2). Molecular studies indicate conifers possess a single copy of a large (~25 Kbp) inverted repeat element in the chloroplast genome (Strauss et al. 1988; Strauss and Doerksen 1990; Raubeson and Jansen 1992), while all other plants tested have two copies.

Notable exceptions to commonly held views of what defines a conifer include the shape and persistence of their leaves and the form of their female strobili. At least 15 species of conifer from five genera (*Larix, Pseudolarix, Taxodium, Metasequoia*, and *Glyptostrobus*) annually shed their leaves, and species like *Nageia wallichiana* or members of the genus *Agathis* might easily be mistaken for broad-leaved trees with their flat, wide leaves. Species of the genus *Phyllocladus* (Podocarpaceae) are essentially leafless, sporting instead modified, flattened, and photosynthetic branches that resemble celery tops. Many species have distinctly different leaf forms for juvenile and mature foliage.

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_1

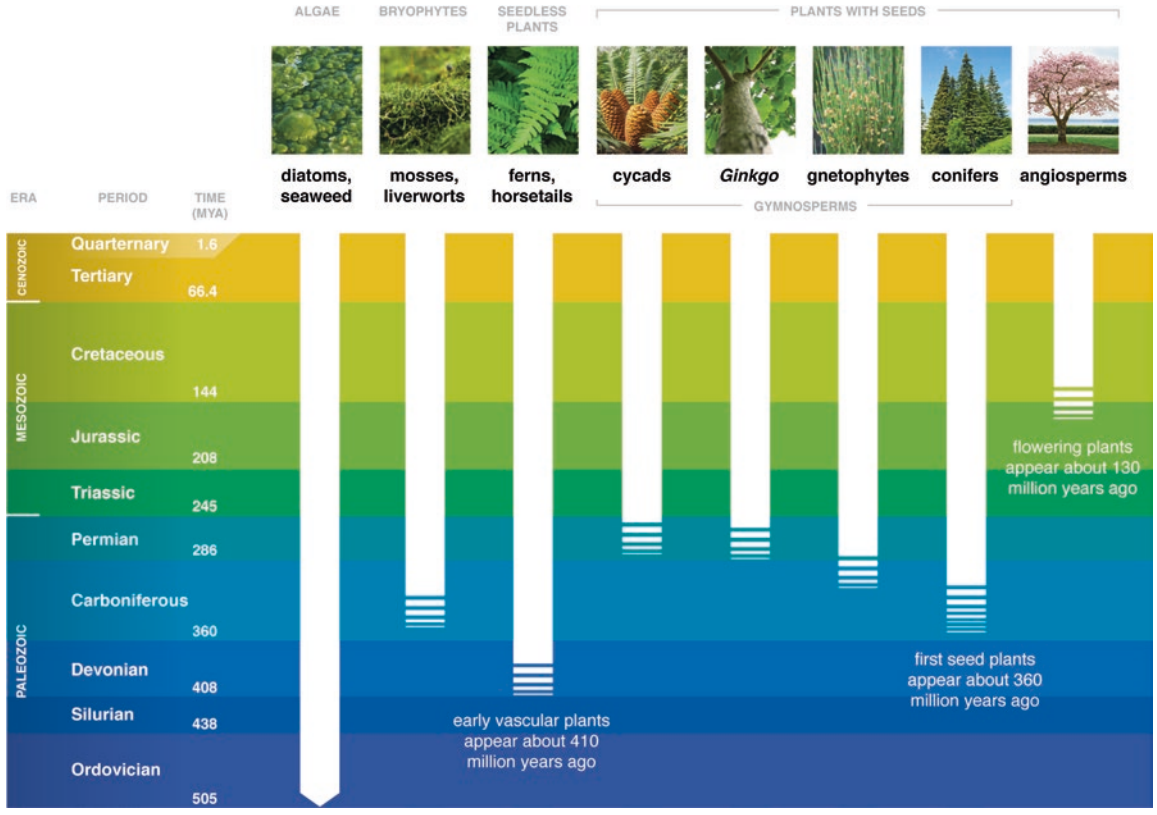


Fig. 1.1 Extant land plant taxa and their estimated evolutionary beginnings. The algae and bryophytes lack vascular elements. The earliest vascular plants, including modern ferns and their allies, are seedless, reproducing via spores. The conifers, along with the cycads, *Ginkgo biloba*, and gnetophytes, constitute the gymnosperms. Along with the angiosperms, gymnosperms represent today's seed-bearing plants

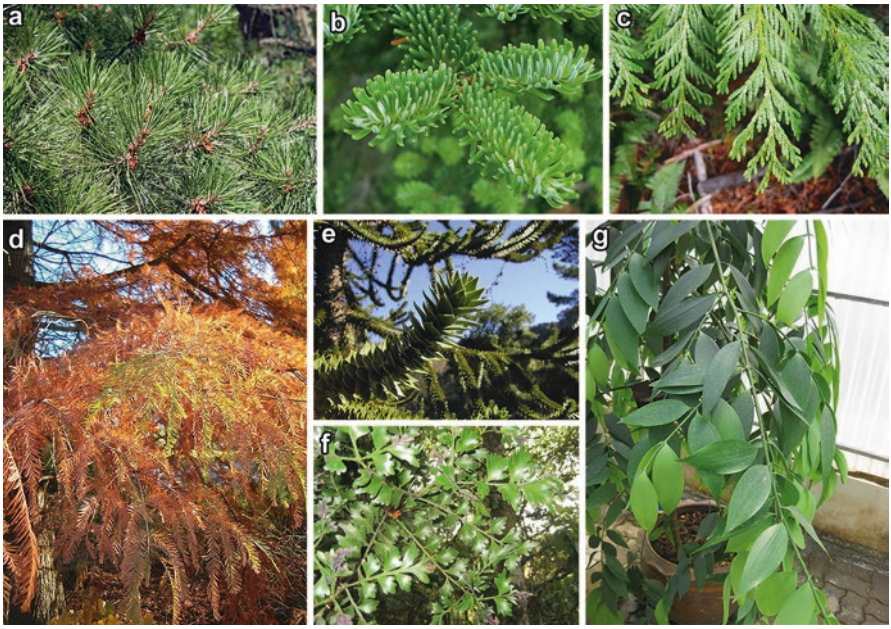


Fig. 1.2 Conifers have evolved numerous foliage types to cope with variable and often challenging environments. Residents of the Northern Hemisphere are familiar with members of the Pinaceae genera (**a**) *Pinus* and (**b**) *Abies* and the Cupressaceae genera (**c**) *Thuja* and (**d**) *Taxodium*, the latter a deciduous conifer. Less familiar to most are the Southern Hemisphere conifers such as (**e**) *Araucaria*, (**f**) *Phyllocladus*, and (**g**) *Nageia*

Not all conifers bear woody cones, as the name might imply (Fig. 1.3). The yews and related taxa in the Taxaceae and the podocarps (Podocarpaceae) generally bear a single seed atop or surrounded by a highly reduced or modified structure, often fleshy and colorful. Cones of the genus *Juniperus* (Cupressaceae) are composed of fleshy, fused scales, and are often referred to, erroneously, as berries. In short, fruiting structures of the conifers are highly variable and reflect strong selective pressures associated with modes of seed dispersal.

Tree form is another trait which varies considerably among conifers. While most conifer species grow as single-stemmed trees with strong apical dominance, multi-stemmed, shrubby species such as *Juniperus horizontalis* and *Lepidothamnus laxifolius* seldom surpass a meter in height. The most widely distributed conifer in the world, *Juniperus communis*, often grows as a low shrub, especially at higher latitudes and a subspecies of *Pinus mugo* (*P. mugo* ssp. *mugo*) is characterized by having multiple stems and short stature even though across much of its native range it is tall and single-stemmed. Many conifer species exhibit the *krummholz* growth habit near tree-line, an apparent response to adverse growing conditions.

Perhaps the greatest challenge to the traditional concept of what defines a conifer is found on the island of New Caledonia. There grows the only known heterotrophic

Fig. 1.3 The seed-bearing structures of conifers are highly diverse. The woody "pine cone" typical of the genus *Pinus* (**a**) is perhaps best recognized while the upright cones of the genus *Abies* (**b**) dry and disintegrate on the tree. The highly modified fruiting structures of *Taxus* (**c**) and *Podocarpus* (**d**) encourage seed dispersal by birds and small mammals

conifer, *Parasitaxus usta*, a small tree of 1–1.5 m, lacking chlorophyll, which survives as a parasite, growing on the roots of another conifer, *Falcatifolium taxoides*. Both species belong to the Podocarpaceae.

In summary, extant conifers are a highly diverse group of taxa representing descendants from a single common ancestor following millions of years of evolution. While most taxa share many traits in common with subsets of other taxa, there are few easily identified traits that characterize them all. The burgeoning field of genomics likely holds promise for expanding our knowledge of what constitutes a conifer (Chap. 3).

Conifer Taxonomy

Taxonomic classification of the conifers is a dynamic process, from the relatively recent discovery of new genera and species (Jones et al. 1995; Farjon et al. 2002) to ongoing revisions at virtually all taxonomic levels up to division, driven by new morphological studies, the rapid expansion of molecular studies of organelle and nuclear genomes, and the completion of monographic treatments. Disagreements frequently exist on the naming and recognition of species, genera, and even families (reviewed in Farjon 2008; Eckenwalder 2009; Farjon 2010; Gernandt et al. 2011).

Conifers have been variously recognized at the level of division or phylum (Pinophyta, Coniferophyta), class (Pinosida, Coniferae), subclass (Pinidae; Cronquist et al. 1966), and order (Coniferales), the latter considered widely accepted (Gernandt et al. 2011), though Christenhusz et al. (2011) recognized three different taxa at the level of order: the Pinales consisting solely of the family Pinaceae, the Araucariales which includes the Araucariacea and Podocarpaceae, and the Cupressales with families Sciadopityaceae, Cupressaceae, and Taxaceae. Taxonomic support for these six families, given recent morphological and molecular studies, appears strong and those six are adopted here, though recent treatments (Farjon 2001, 2008; Farjon and Filer 2013) recognize as many as eight families, including Cephalotaxaceae and Phyllocladaceae with the previous six (Fig. 1.4).

Classification of species and genera has fluctuated to an even greater degree over the last 60 years. Dallimore et al. (1967) accepted 53 genera, while more recent treatments recognized as many as 71 (Farjon 2001, 2008, 2010; Eckenwalder 2009; Gernandt et al. 2011; Christenhusz et al. 2011; Table 1.1). This increase in genera has been coincident with a general decline in the number of recognized species, though this number too remains elusive to define, with a range of 546 to 670, as noted in recent treatments (Table 1.1). The highly variable estimates likely result from how accepted taxa are treated at the species and subspecies levels.

Taxonomic diversity of families ranges widely (Farjon 2008; Farjon and Filer 2013). The Pinaceae has ~231 species in 11 genera, the largest of which is *Pinus* with ~113 recognized species. The other large families are the Podocarpaceae

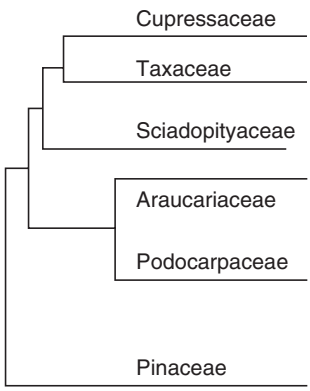


Fig. 1.4 Cladogram of the six conifer families most frequently recognized today. (Adapted from Gernandt et al. 2011). Some authorities recognize two additional families: the Cephalotaxaceae, here subsumed in the Taxaceae, and the Phyllocladaceae, here subsumed in the Podocarpaceae

Table 1.1 Reported number of conifer taxa for family, genus, and species as listed in four recent authoritative treatments

	Number of taxa			
Taxon	A	B	C	D
Family	8	6	6	6
Genus	69–70	71	67	69
Species	615–630	670	546	–

A = Farjon (2008), Farjon and Filer (2013)
B = Gernandt et al. (2011)
C = Eckenwalder (2009)
D = Christenhusz et al. (2011)

(18 genera, 175–190 species, including the phylloclads) and Cupressaceae (30 genera, 135 species). Of the 30 genera in the Cupressaceae, 17 are monotypic (Farjon 2008). The Araucariaceae (3 genera, 37–41 species), Taxaceae including the formerly recognized Cephalotaxaceae (6 genera, 36 species), and Sciadopityaceae (1 genus, 1 species) are considerably smaller. Enumeration and descriptions of all conifer species and their distributions are provided in excellent references by Eckenwalder (2009) and Farjon and Filer (2013).

Though diverse and seemingly species-rich, the conifers in general may be in decline and have been since the Mesozoic, an era some paleobotanists have referred to as the Age of the Conifers. As many as 20 conifer families, 12 of which are now extinct, have been recognized and it has been estimated the Jurassic flora may have had as many as 20,000 conifer species (Debazac 1964 cited in Farjon 2008). Clearly, a great deal of diversity has been lost in the last 65 million years, and the existing inventory of conifers includes many species struggling to survive. Farjon (2008, pp. 184–185) lists 39 species considered to be relicts that are currently known to exist in a single locality, often confined to a few square kilometers. The International Union for Conservation of Nature Red List of Threatened Species (IUCN 2017), which tracks the status of organisms worldwide, listed 196 of the 605 conifer species reviewed to be vulnerable, endangered, or critically endangered in 2015. In some cases, such as with *Abies beshanzuensis*, the number of known living trees can be counted on one hand (Yang et al. 2013). Some genera and families are significantly more threatened than others (Table 1.2), no doubt a reflection of their evolutionary past and their inability to adapt to new and changing conditions, as well as range fragmentation, human development, and mismanagement. A more thorough look at the evolutionary history of conifers and their relationships to one another is explored in Chap. 16.

Geographic Distribution and Biogeography

Conifers are widely distributed across much of the world's landmasses, occurring on all continents excepting Antarctica (Fig. 1.5, Farjon and Filer 2013). They form extensive, circumboreal forests across North America and Eurasia, the largest

Table 1.2 The number of species listed as vulnerable, endangered, or critically endangered for select genera as noted in the IUCN Red List (April 2015)

	Status					
Genus	Vulnerable	Endangered	Critically endangered	Total threatened	Number of species in genus	Percent of species threatened
Abies	4	7	4	15	46	32.6
Agathis	4	5	–	9	16	56.3
Amentotaxus	3	2	–	5	6	83.3
Araucaria	5	6	1	12	17	70.6
Cupressus	4	3	–	7	15	46.7
Dacrydium	2	3	1	6	21	28.6
Juniperus	4	7	2	13	53	24.5
Picea	7	6	2	15	36	41.7
Pinus	7	9	2	18	113	15.9
Podocarpus	11	19	5	35	97	37.1
Taxus	1	3	1	5	8	62.5
Widdringtonia	–	–	2	2	4	50.0

Fig. 1.5 The worldwide distribution of conifers. (From Farjon 2008, p. 166)

recognized floristic region in the world (Takhtajan 1986), and they are common to predominant elements in floras throughout North and Central America, large areas of South America, Malesia (the biogeographical region straddling the Indomalayan and Australasian ecozones), Australia, New Zealand, China, and Eastern Africa. Conversely, they are rare or absent in deserts, steppes, the Arctic tundra, alpine regions, and the great lowland tropical rainforests of the Amazon and Congo Basins, predominantly due to ecological preferences and tolerances. The absence of conifers in large areas of otherwise habitable land on the Indian subcontinent, Southern Africa, and southern South America is discussed in some detail by Farjon (2008) and may be variously attributed to historical periods of climate change, the breakup of the supercontinent Gondwana, and vicariance events like mountain building, flood basalts, and continental drift. Conifers are commonly dominant in stressful or extreme environments where water and temperature regimes are limiting to plant growth (Richardson and Rundel 1998) but less successful in ameliorating environments that favor the rapid growth of angiosperms (Coomes et al. 2005).

The study of the distribution of organisms across large geographic regions is known as biogeography, the roots of which date to the studies of Alfred Russell Wallace on the Malay Archipelago (Wallace 1876). Over the last several decades, biogeographical studies seeking to explain the distribution patterns of organismal groupings have been significantly influenced by the near-universal acceptance of plate tectonics or continental drift (Moss and Wilson 1998). For the ancient conifer lineages, which have been evolving for over 300 million years, continental drift helps explain much of the current worldwide distribution patterns of extant taxa.

The consensus theory today is that the Earths' landmasses have moved about on continental plates for at least two billion years (Zhao et al. 2004), periodically forming supercontinents, where all plates are roughly in contact with each other. The most recent of these supercontinents, known as Pangea, likely formed about

300 million years ago (mya), about the time early conifers appear in the fossil record. Pangea began breaking up roughly 200 mya into two minor supercontinents called Gondwana and Laurasia which roughly equate to todays' Southern and Northern Hemisphere landmasses, respectively (Fig. 1.6). The breakup of the land masses had a profound effect on the subsequent distribution of conifer taxa. Extant conifer diversity is largely (90%) contained in major clades (families or subfamilies) that are confined to either the Northern or the Southern Hemisphere (Leslie et al. 2012, Fig. 1.6). The Podocarpaceae and Araucariaceae, early fossils of which occurred on Gondwanan landmasses prior to the breakup of Pangea, remain largely restricted to the Southern Hemisphere, while the Pinaceae, Sciadopityaceae, and Taxaceae are predominantly of Northern Hemisphere origin (Critchfield and Little Jr 1966; Eckenwalder 2009; Farjon and Filer 2013). The large and diverse Cupressaceae are distributed worldwide but are divided into subfamilies that are mostly hemisphere specific: Cupressoideae in the north, Callitroideae in the south (Mao et al. 2012).

Similarly, generic and species diversity in the conifers appears to have been significantly influenced by continental- and hemispheric-scale phenomena like the distribution of oceans, mountain-building, climate change, and continental drift (Leslie et al. 2012). Conifer taxa are clearly not distributed randomly around the world. Farjon (2008, Ch. 24) notes that all extant families, 83% of genera (gen) and over half of all species (sp) occur in 14 centers of diversity around the Pacific Ocean, many of which are islands like Japan (18 gen, 45 sp), New Caledonia (14 gen, 44 sp), and Taiwan (17 gen, 26 sp). In the eastern Pacific, the Pacific Northwest (13 gen, 29 sp), California (14 gen, 48 to 50 sp) and southern Mexico/Guatemala (7 gen, 39 sp) host forests of remarkable conifer diversity. The Klamath Mountains of northwest California alone is home to 35 species (Griffin and Critchfield 1976; Kauffmann 2012), nearly half of which can be found within remarkably diverse forest plots. Around the Pacific, the distribution of rare and endemic species is often coincident with these centers of diversity. Approximately 70% of all conifer species occur in the Northern Hemisphere.

Life History

The conifers are large, long-lived, woody perennial plants that often grow in extensive panmictic populations covering vast portions of the boreal and temperate regions of the world and in mixed stands in tropical and subtropical forests. They are the dominant life forms in a diverse array of ecosystems, demonstrating the capacity to adapt to highly variable climatic and edaphic conditions. Their success and persistence, over 300 million years of continental shifting, climate changes, mountain building and volcanism, must rest in large part with the spectrum of life history traits they exhibit and, by inference, the store of genetic diversity they maintain.

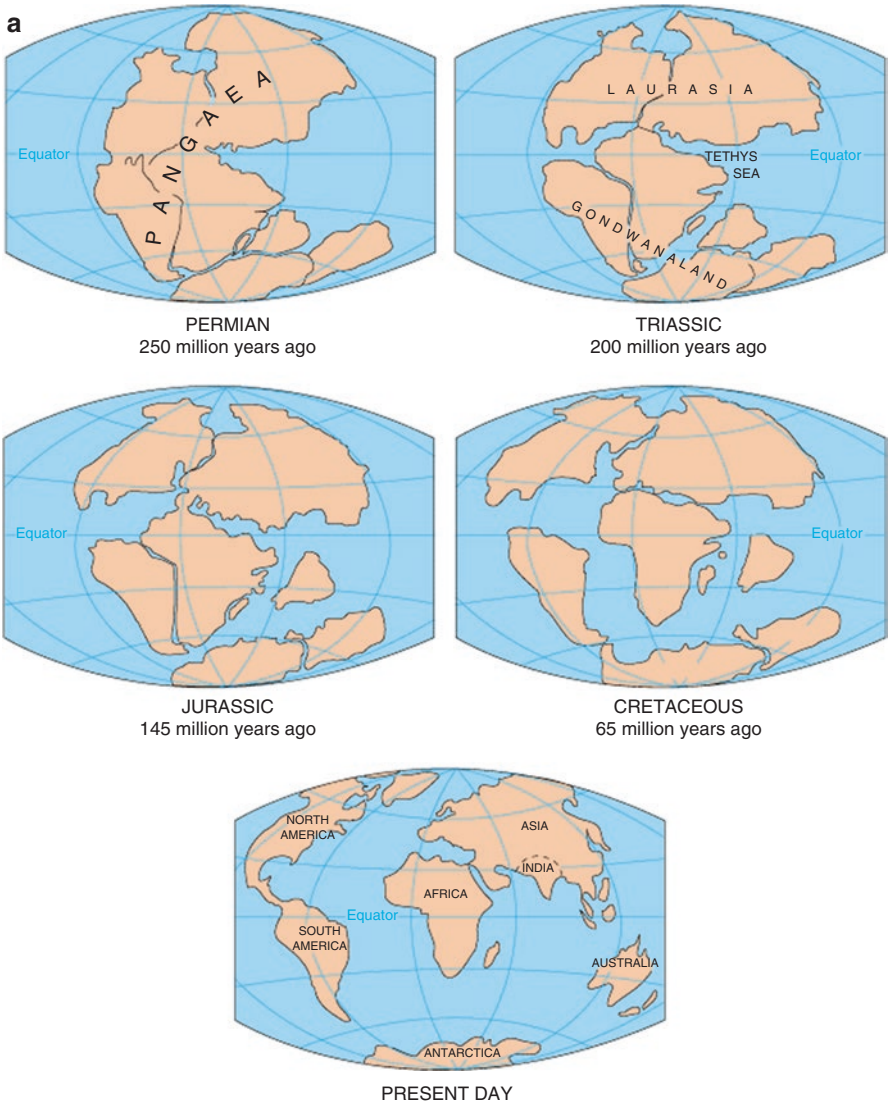
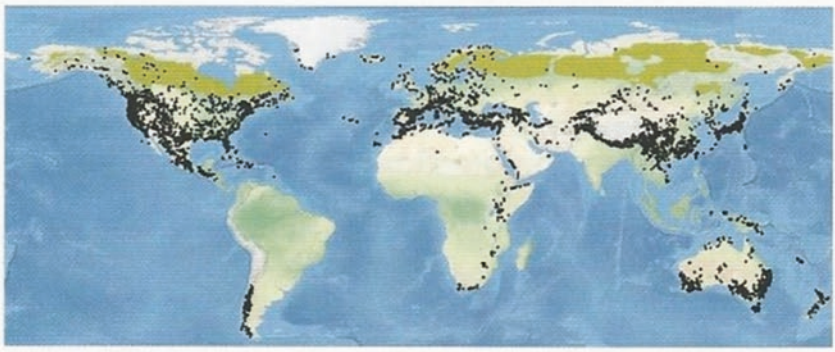
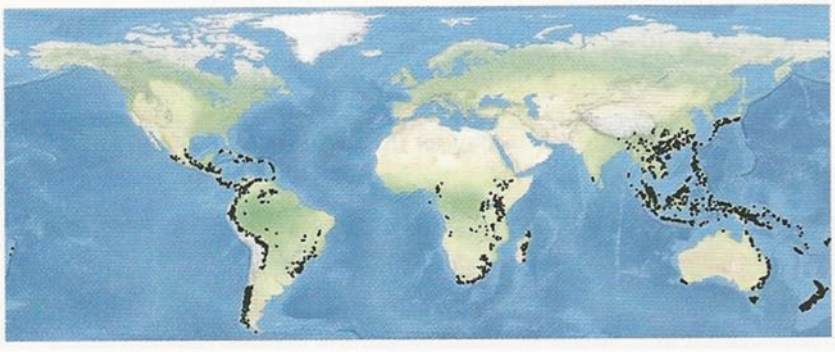


Fig. 1.6 (**a**) Hypothetical pattern of continental shifting following the breakup of the supercontinent Pangea (from http://www.rocksinmyheadtoo.com/Pangea.htm), and global distribution of (**b**) the Cupressaceae (map GTC-5, Farjon and Filer 2013, p. 14), (**c**) the Pinaceae (map GTC-6, Farjon and Filer 2013, p. 14), and (**d**) the Podocarpaceae (map GTC-10, Farjon and Filer 2013, p. 15)



Map GTC-5. The global distribution of Cupressaceae.

Map GTC-6. The global distribution of Pinaceae.



Map GTC-10. The global distribution of Podocarpaceae.

Fig. 1.6 (continued)

Fig. 1.7 Conifers are the dominant plants on expansive areas of (**a**) boreal and (**b**) high-elevation temperate ecosystems in the Northern Hemisphere

Ecological Tolerance

Conifers have adapted to and successfully competed with angiosperms in most (17 of 21) major vegetation types in which vascular plants occur (Farjon 2008, p. 36–37). They tolerate extremes of temperature and moisture, from arctic tundra to hot deserts, and from less than 250 mm/year to well over 5000 mm/year of precipitation. Conifers are dominant across the vast boreal forests (Fig. 1.7) of the North American and Eurasian continents where they thrive under short growing seasons, extreme cold, heavy winter snow loads and, frequently, what otherwise might be considered desert conditions (precipitation less than 375 mm/year). At these high latitudes, low evapotranspiration demands do not inhibit the development of closed stands. The modest number of species that survive in these conditions possess many physiological and morphological adaptions that contribute to their success. All have relatively

tall, narrow crowns that both shed snow in the winter and capture low-angle sunlight at the end of the growing season. Evergreen crowns (*Abies, Pinus*, and *Picea*) require less energy to produce and maintain than annually replacing the entire leaf surface and permit photosynthetic activity as soon or as late as conditions permit. *Larix* species avoid cold damage to foliage by dropping needles annually. All species have deep dormancy capacity and the biochemistry to tolerate extreme cold events that might otherwise damage or kill primary and secondary meristematic tissues (buds, cambium). Conifers that inhabit the higher elevations of the world's major mountain ranges share many of the same adaptations. Adaptation to winter cold in temperate and boreal trees involves an array of complex genetic, physiological, and developmental processes for which most conifers in these environments exhibit remarkable diversity (Howe et al. 2003).

Elsewhere, in the Northern Hemisphere, conifers often grow in areas with high evaporative stress (Gernandt et al. 2011) and may be considered as xerophytes (Mirov 1967). In large part, members of the Pinaceae and Cupressaceae possess an array of traits that confer drought tolerance, like thick cuticles and epicuticular wax layers on needles, the ability to shut down stomata under unfavorable conditions, and mycorrhizal associations that can significantly enhance water uptake (Molina and Trappe 1984; Smith and Brewer 1994). By contrast, many of the Southern Hemisphere Podocarpaceae and Araucariaceae have evolved under more mild, wetter conditions (Leslie et al. 2012) and are generally less cold and drought hardy. Exceptions to this are common however. Though Podocarps have a pan-tropical distribution across continents, they are largely montane dwelling in Africa and South America where droughty conditions exist (Addie and Lawes 2011).

Though conifers do not typically tolerate saturated or flooded soils, here too exceptions exist such as New Zealand's Kahikatea swamp forests (with *Dacrycarpus dacrydioides*), the cypress swamps of the Southeastern United States (with *Taxodium disticum*), and the temperate rainforests of the Pacific Northwest (US), Canada, and Southeast Alaska. Conifers in temperate rainforests in both hemispheres adapt by establishing extensive, but shallow, spreading root systems close to the surface. Few if any conifers tolerate saline conditions (Farjon 2008).

While most conifers grow exceptionally well on good soils, such as the deep, uplifted marine sediments of the coastal northwestern United States, they are more typically found on nutrient-poor, often shallow soils in upland and mountainous regions where angiosperms compete less well (Fig. 1.7). Their success on poor soils has been attributed, in large part, to the near-universal association and coevolution (Brundrett 2008) between conifers and mycorrhizae, a symbiotic plant–fungus relationship that enhances nutrient and moisture uptake, provides plant hormones to facilitate root growth, and functions to reduce some soil toxins and protect against other deleterious fungi (Malloch et al. 1980; Molina and Trappe 1984). Mycorrhizae are classified by how they physically associate with plant roots. Ectomycorrhizae, which form extensive mycelia mats or mantles among the tree's fine root tips (extracellular), are found in association with members of the Pinaceae, while vascular–arbuscular (VA) mycorrhizae, which form intercellular associations, are associated with all other conifer families (summarized by Brundrett 2008; see also Malloch et al. 1980; Molina and Trappe 1984).

Soils play a significant role in the distribution of conifers worldwide and can influence both species composition and successional status of conifer forests. While this often is manifest in a tree's tolerance to drought or moisture, it may also result from adaptation to unusual soil chemistries. To illustrate, three examples are drawn from the diverse forests of western North America. The short-lived, pioneering *Pinus contorta* ssp. *contorta*, common from coastal muskeg to upland sites, may become an edaphic climax species on deep, excessively drained sand and gravel soils along the Pacific Coast of western North America, where other temperate rain forest species simply cannot compete or survive. At mid-elevations, in the Klamath Mountains of northern California and southern Oregon, ultramafic soils (serpentine soils) support unique plant assemblages including some conifers, like *Pinus jeffreyi*, that tolerate the heavy metal concentrations in such soils, often without competition from other conifers (Sawyer 2006; Kauffmann 2012). And lastly, the Great Basin bristlecone pine (*Pinus longaeva*), known for specimens of remarkable age, grows almost exclusively on limestone soils at elevations more than 2800 meters in the western United States, virtually to the exclusion of other species. *Pinus longaeva* has combined adaptive tolerances to cold, drought, intense solar radiation, and soil chemistry challenges, and in so doing, largely avoids physical challenges from disease and insect pests that cannot survive in these rarified conditions.

Light and fire represent the remaining abiotic factors that contribute to the success or failure of conifer establishment and survival. Most conifers are early successional or pioneering species, occupying newly disturbed sites with full-sunlight conditions and exposed mineral soil. Such species are shade intolerant. Disturbance is key to conifer succession (Peattie 1953; Daubenmire 1968). Primary succession, on newly created sites, results following land shaping events like glacial retreat, floods, lava flows, and avalanches. More commonly, secondary succession of conifers follows existing stand disturbance caused by factors such as fire, wind-throw, timber harvest, or pestilence. At least in the Northern Hemisphere, fire has played, by far, the greatest single role in shaping conifer stand dynamics and evolution, and has been doing so for a very long time (Muir and Lotan 1985; Keeley and Zedler 1998; Pausas and Keeley 2009; Keeley 2012). Conifers have evolved an array of fire-adapted traits, the origins of which date at least to the Cretaceous period (65–145 mya, He et al. 2012a). Such traits include bark thickness, pyriscence (a special case of serotiny), branch shedding, the presence of a grass stage, and re-sprouting capacity (Keeley and Zedler 1998; Keeley et al. 2011; He et al. 2012a). Fire-adapted traits are most pronounced in the Pinaceae, particularly in the genus *Pinus*, and to a lesser extent the Cupressaceae. Pyriscence refers to the habit of retaining seed in cones that are sealed by a resinous coating and are triggered to open when heated by fire. The trait, common in the genus *Pinus*, subsections Contortae, Oocarpae, and Sylvestris, can result in large crown-borne seed banks and the release of millions of seeds per hectare following fire (Wheeler and Critchfield 1985; Stevenson 1991). Extreme examples of regeneration success following fire in *Pinus contorta* stands in northern British Columbia, Canada exceed 20,000 stems per hectare (Wheeler, personal observation, 1975; Fig. 1.8).

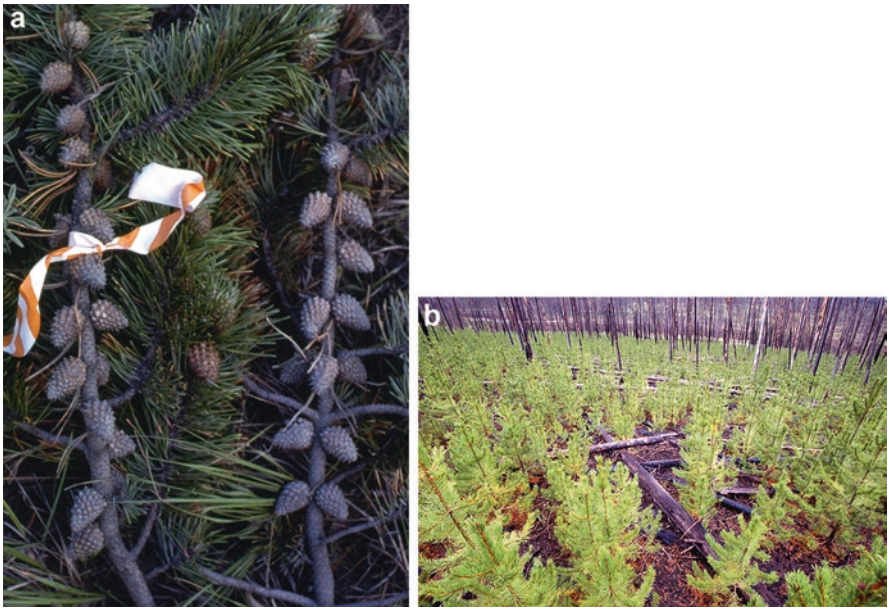


Fig. 1.8 (**a**) Serotinous cones may accumulate for decades in the crowns of *Pinus contorta* ssp. *latifolia* producing enormous crown-borne seed banks. The heat of low to moderate intensity fire causes the cones to open and release seed, which often leads to (**b**) overstocked stands. (NPS Photo by Jim Peaco)

Shade intolerance is often cited as a significant factor in the inability of conifers to compete with rapidly growing angiosperms under favorable growing conditions (Farjon 2008). Members of the Pinaceae and Cupressaceae are probably least shade tolerant, though genera such as *Abies*, *Tsuga*, and *Thuja* produce some remarkably tolerant species, capable of germinating and surviving under near complete canopy closure for decades before assuming normal growth following stand release. The range of shade tolerance among species often results in the replacement of a stand's pioneering species with one or more seral stages leading to a climax or self-perpetuating stand with different species mixes (Daubenmire 1968).

Many species in the Podocarpaceae and Araucariaceae exhibit varying levels of shade tolerance that allow them to establish and thrive in very modest size disturbances or even under an established canopy (Lamont et al. 1991). This has been particularly well documented in review papers on conifer silviculture in New Zealand (Stewart 2002) and Africa (Addie and Lawes 2011), the latter of which suggests that shade tolerance is a key factor in the maintenance of podocarps in Afro-montane mixed forests. Consequently, conifers in the Southern Hemisphere, especially in the Asia-Pacific and African regions, most commonly occur in small stands or as single trees (Verkaik et al. 2007), and generally in mixed species groupings, in stark contrast to the vast conifer stands noted in the Northern Hemisphere.

Conifer Mating System, Life Cycle, and Reproduction

Conifers are almost exclusively wind-pollinated and predominantly outcrossing, traits that encourage widespread gene flow and maintenance of genetic diversity. Outcrossing is facilitated by a spectrum of factors (reviewed in White et al. 2007). As noted previously, conifers have separate or unisexual reproductive structures (strobili) either on the same (monoecy) or different plants (dioecy). In monoecious species, pollen and ovulate strobili are often spatially separated, pollen produced on lower branches and seed on upper branches. Timing of pollen shed and seed cone receptivity are often offset, with pollen shed typically occurring first (by one to several days). Perhaps more importantly, selfing and lower levels of consanguineous mating typically lead to relatively severe inbreeding depression manifested as aborted seed or non-thrifty seedlings that die early or grow very slowly (Sorensen and Miles 1974; Fowler and Park 1983; Wilcox 1983). Self-infertility in conifers, particularly in the well-studied Pinaceae, appears to be a function of high genetic load as defined by the estimated number of lethal equivalent alleles a tree carries (Sorensen 1969). While conifers lack self-incompatibility mechanisms common to angiosperms, they do possess polyembryony, wherein ovules may have multiple (1–10) archegonia, each with a viable egg cell. Multiple fertilizations may occur, but multiple embryos per seed are rare, suggesting competition among embryos based on genetic load is severe (Owens and Blake 1985). This mechanism likely controls inbreeding and contributes to differential male reproductive success (Nakamura and Wheeler 1992a, b). The average number of lethal equivalents in well-studied species is typically quite high (>10), but significant variation can occur among trees within species (i.e., 3–27.6 in *Pseudotsuga menziesii*), insuring that some selfing (5–20%) is still likely to occur (Sorensen 1969; Piesch and Stettler 1971).

Conifers exhibit a diplohaplonic life cycle (reviewed in Williams 2009) in which both diploid sporophytic and haploid gametophytic phases occur concurrently on the same plant (Fig. 1.9). Male and female spores are borne in separate structures (microsporangium and megasporangium, respectively) occurring on current year's growth, near the ends of branches. Reproductive buds typically differentiate in late winter to early spring. Following meiosis and a brief series of mitotic divisions, microspores (pollen) are released and megaspores become receptive, housed within developing cones or upon receptacles, as noted previously. Pollination occurs in early to late spring, but fertilization is delayed for a few to 12 or more months, as is the typical case in the genus *Pinus* (Fig. 1.9). Following fertilization early in the year after cone initiation, the second-year cone matures and sheds seed in the fall. For most conifers, the process is completed within 6–9 months but, in some cases, may take as much as 3 years (Owens and Blake 1985; Williams 2009).

A notable feature of the conifer life cycle is the haploid nature of the megagametophyte. A product of meiosis, this female-derived tissue is highly valued by geneticists who have used it extensively for genetic mapping of segregating markers and traits (Chap. 11) and as the basis for developing reference genome sequences (Chap. 3).

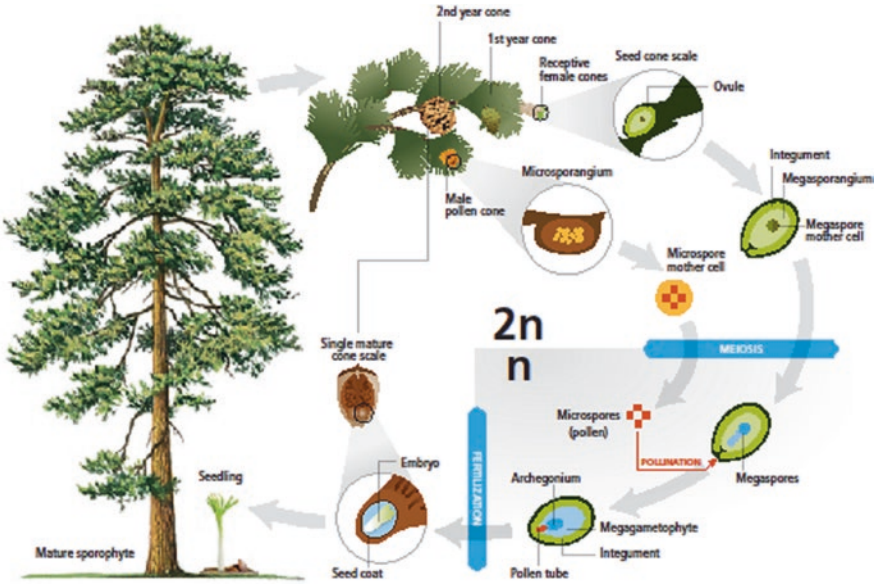


Fig. 1.9 The reproductive cycle of a typical conifer. (From Neale et al. 2014)

Most conifers have an extended period of juvenility (10–25 years), during which they are reproductively incompetent. Williams (2009) differentiates between reproductive onset and reproductive competence. In the former condition, plants may produce strobili given strong external stimuli, such as girdling (Wheeler et al. 1985; Wheeler and Bramlett 1990) or treatment with growth regulators (Wheeler et al. 1980). In extreme cases, members of the Cupressaceae have been stimulated to flower by age one (Pharis and Morf 1967). Once reproductively competent, healthy trees will continue to produce pollen and seed for hundreds—even thousands—of years. Conifers simply do not fit neatly in the r/k reproductive strategy model developed by MacArthur and Wilson (1967). They possess most of the ecological characteristics of *k* strategists (long-lived, large, stable environments), but have a prodigious reproductive capacity typical of r strategists. Individual trees can produce tens of thousands of seed and millions of pollen grains, annually. Williams (2009) estimates a single hectare of *Pinus taeda* can release 1×10^{11} pollen grains daily during anthesis. Those living in or near conifer forests are frequently reminded of the capacity of trees to produce literally tons of pollen over short periods, covering windshields and yard furniture with a gritty, yellow layer that flows almost like liquid and accumulates to measurable depths (Fig. 1.10).

Not only is conifer pollen production often copious, it may disperse over great distances, even under modest wind conditions. Pollen dispersal distances up to several hundred kilometers from the nearest source stands have been reported for members of the Pinaceae (reviews by Lanner 1966; Koski 1970; Burczyk et al. 2004a, b) but viability of such well-traveled pollen remains suspect. Physical measures of pollen dispersal distances suggest most pollen falls within a few hundred

Fig. 1.10 A heavy band of *Picea sitchensis* pollen left on coastal rocks following high tide in a Southeast Alaska protected inlet. (Photo courtesy of Richard Billings)

meters of the source tree (reviewed in Di-Giovanni and Kevan 1991). Effective pollen dispersal, as measured by genetic markers, suggests most seed are sired by trees located nearby (Erickson and Adams 1989), but very high rates of contamination (1–89%) from pollen sources located well away (0.5 to many km) from the seed tree are common for Pinaceae species (reviewed in Wheeler and Jech 1992; Adams and Burczyk 2000).

Wind and gravity are predominant modes of seed dispersal for most members of the Pinaceae, Cupressaceae, and Araucariaceae while seeds surrounded by fleshy arils and epimatia, common to the Taxaceae and Podocarpaceae, are typically distributed by birds and small mammals, as are the large, wingless seeds of some pine species. Point source studies (reviewed in White et al. 2007) reveal wind dispersed seed typically falls within 50 meters of a parent tree but such studies must significantly underestimate the tail of the leptokurtic curve that describes seed dispersal. In North America, following the last glacial retreat, species of pine, spruce, and larch successfully migrated up to 4000 km in under 12,000 years, an average annual distance of 330 m. Exotic pine species planted in New Zealand have established at distances up to 18 km from the nearest seed source (Bannister 1965). And long-distance dispersal (22 km) of *Pinus edulis* seeds by Clark's nutcracker (*Nucifraga columbiana*) has been noted in western North America (Vander Wall and Balda 1977) though most bird-dispersed seed is likely cached much closer to the seed-bearing parent trees.

Asexual reproduction of conifers in the wild is relatively less common than in angiosperms but when it occurs may be of importance. Following fire or harvest, *Cunninghamia lanceolata* produces stump sprouts (coppice) that grow rapidly and will reoccupy a site quickly (Li and Ritchie 1999a, b). The sprouts, in turn, are easily propagated by direct planting of cuttings, or of rooted cuttings. Vegetative propagules have been used for reforestation and afforestation in China for over 800 years and may have contributed to over 80% of all planted *Cunninghamia lanceolata* in the recent past (Li and Ritchie 1999a, b). *Sequoia sempervirens* produces sprouts

from basal lignotubers, and re-sprouting from epicormic buds on boles and branches of flooded or burned trees occurs in some species (i.e., *Pseudotsuga menziesii*, *Pinus rigida*). Layering of branches on the ground occurs in several members of the Cupressaceae, in *Picea*, and in *Phyllocladus aspleniifolius* (Gernandt et al. 2011). The entire population of the recently discovered relict *Wollemia nobilis* in Australia is reported to be a single clone. Coppicing occurs on virtually all adult stems. The fragile existence of this species may well be a function of its ability to reproduce asexually. Commercial production of conifer vegetative propagules has enjoyed considerable attention in recent decades. Ritchie (1991) reported that more than 65 million rooted cuttings of 11 conifer species were being produced annually around the world, easily half of which were of *Cryptomeria japonica* in Japan, which has been vegetatively propagated for over 500 years (Toda 1974). To this may now be added the millions of steckling (rooted cuttings) of *Cunninghamia* in China noted by Li and Ritchie (1991b). A recent review of natural vegetative propagation in gymnosperms (Lamant 2012) has been enhanced and presented online at the Gymnosperm Database (Earle 1997–2019).

The Largest, Tallest, and Oldest Organisms on the Planet

This brief look at the world of conifers would be incomplete without noting some remarkable facts about these ancient seed-bearing plants. Previous discussions have focused on conifers in general, their ecological preferences, distribution and biogeography, and taxonomic or phylogenetic placement. The following discussion is focused on individual conifer trees and their status as the largest, tallest (arguably), and oldest living organisms on the planet, all of which currently exist within the State of California in the United States. All trees noted here are single-stemmed and not part of a long-lived clone.

The largest trees in the world, as measured by stem volume or by a points system (http://www.americanforests.org/our-programs/bigtree/), belong to *Sequoiadendron giganteum*, the giant sequoias, and the largest of these, the General Sherman, has a stem volume of 1489 m^3 (55,040 ft^3), stands 83.8 meters (275 ft) tall, has a diameter of 7.7 m (25 ft), and is in excess of 2500 years old (Van Pelt 2001, Fig. 1.11a). *Sequoiadendron giganteum* grows on the western slopes of the Sierra Nevada Mountains where they are restricted to some 75 groves covering less than 15,000 ha, although fossil evidence suggests they once were considerably more widespread in the western United States (Axelrod 1959, 1964).

The related coastal redwoods (*Sequoia sempervirens,* Fig. 1.11b), growing approximately 700 km to the north and west of the *S. giganteum* stands, routinely exceed 360 feet in height, the tallest of which is currently measured at 378 feet (http://www.conifers.org/cu/Sequoia.php). *Sequoia sempervirens* also have exceptional life spans, many exceeding 2000 years.

Worldwide, examples of ancient trees, with ages typically verified by ring counts, number 28 or more species with specimens that exceed 1000 years of age (Table 1.3, Pojar and MacKinnon 2004; Farjon 2005, 2008; Earle 2011; Brown 2013;

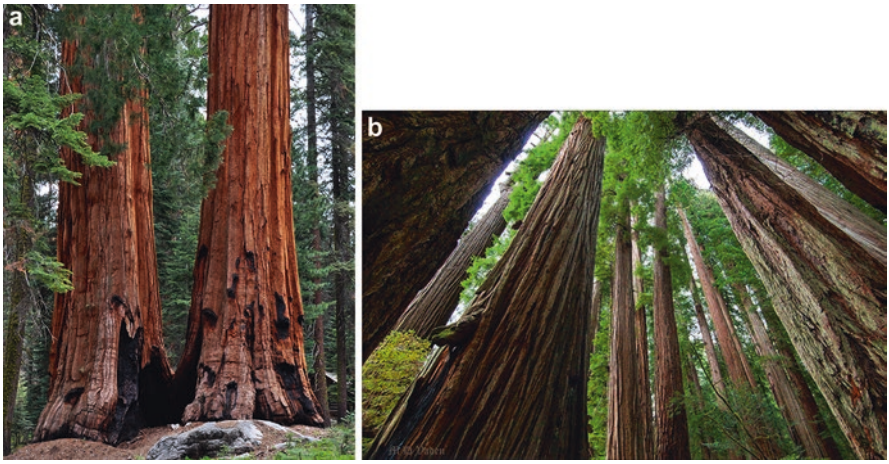


Fig. 1.11 (**a**) A pair of *Sequoiadendron giganteum* growing in Sequoia National Park. (Photo credit to Dr. Bruce Bongarten) and (**b**) a thrifty stand of *Sequoia sempervirens*. (Photo credit to M. D. Vaden)

Table 1.3 Life spans of conifer species known to exceed more than 1000 years

>1000 years
Araucaria araucana, *Agathis australis*, *Chamaecyparis obtusa*, *Cryptomeria japonica*, *Dacrydium cupressinum*, *Juniperus scopulorum*, *Larix lyalli*, *Picea sitchensis*, *Pinus albicaulis*, *Pinus edulis*, *Pinus flexilis*, *Pseudotsuga menziesii*, *Taxodium distichum*, *Taxodium mucronatum*, *Taxus baccata*, *Thuja occidentalis*, *Thuja plicata*, *Tsuga heterophylla*, *Tsuga mertensiana*
>2000 years
Juniperus occidentalis, *Lagarostrobus franklinii*, *Pinus aristata*, *Pinus balfouriana*, *Sequoia sempervirens*, *Cupressus nootkatensis*
>3000 years
Fitzroya cupressoides, *Sequoiadendron giganteum*
>4000 years
Pinus longaeva

Wikipedia contributors 2017). By far, the oldest trees belong to *Pinus longeava* (the Great Basin bristlecone pines), the record for which was determined recently to be 5060 years old (Brown 2013).

Genetic Diversity

We have referred often to the considerable diversity that exists among and between the genera and family level taxa of conifers, but relatively little about genetic variation within species or populations. As will be detailed in future chapters of Sections One and Two, the conifers in general retain large stores of variation, at virtually

all levels considered, from single nucleotide polymorphisms (SNPs), non-coding repetitive elements, and gene structure, to measures of phenotypic, proteomic, and metabolomic variation. That is not to say that some species are not genetically depauperate, a condition that certainly puts them at risk of near-term extinction, as alluded to in an earlier reference to relict species in this chapter. It is the objective of this volume to illustrate and summarize, to the extent possible in the space provided, the extent or lack thereof, of genetic variation in conifer genomes and whole plants, and discuss its evolutionary significance for this ecologically and economically important group of plants.

Summary

The conifers are a diverse and phylogenetically ancient group of seed plants of monophyletic origin that arose more than 300 million years ago. Once the predominant land plants on Earth, the conifers have been in decline since the Mesozoic era, coincident with the rise of flowering plants (angiosperms). Today some 670 conifer species in six families are recognized, though the numbers vary by authority. The conifers are not easily circumscribed or defined. While there are many traits commonly associated with conifers, exceptions to most character states exist. They are typically single-stemmed, evergreen trees with separate male (simple) and female (compound or reduced) reproductive structures (strobili or cones) either on the same (monoecy) or different plants (dioecy). Along with the distantly related cycads, *Ginkgo biloba*, and gnetophytes, the conifers comprise the gymnosperms, all of which bear naked seeds. The conifers are typically large, long-lived, woody perennial plants that often grow in extensive panmictic populations covering vast portions of the boreal and temperate regions of the world and in mixed stands in tropical and subtropical forests. They are the dominant life forms in a diverse array of ecosystems, demonstrating the capacity to adapt to highly variable climatic and edaphic conditions. Conifers are almost exclusively wind-pollinated and predominantly outcrossing, traits that encourage widespread gene flow and maintenance of genetic diversity. While most conifer species maintain high levels of genetic diversity within and among robust populations, dozens of species are today listed as threatened or endangered, in part a function of shrinking populations and reduced genetic variation. Conifers are notable for their size and age. The tallest (*Sequoia sempervirens*), most massive (*Sequoiadendron giganteum*), and oldest (*Pinus longaeva*) individual organisms on Earth are all conifers growing within a few hundred miles of one another in the state of California, United States of America.

Part I

Genomes

Genomes: Classical Era

2

The Beginnings of Genome Research in Conifers

The study of conifer genomes began in 1933 with the publication by Sax and Sax (1933) of chromosome numbers found in several conifer species. From 1933 until the beginning of the application of recombinant DNA technologies in the 1970s, the study of conifer genomes was done using classical cytogenetic techniques. From the 1970s until the late 1990s, a suite of developing DNA technologies was applied to the study of conifer genomes. We have labeled the years between 1933 and the late 1990s as the *classical era* in the study of conifer genomes. Beginning in the late 1990s, high-throughput DNA sequencing began to be applied. We have labeled this period as the *modern era* which will be covered in Chap. 3. In conifers, like all higher plants, there are three genomes; the nuclear genome (nDNA), the chloroplast genome (cpDNA), and the mitochondrial genome (mtDNA). In this chapter we will discuss what was learned about chromosome number, ploidy, karyotypes, genome size, and basic genome content in all three genomes during the *classical era.*

Chromosome Number and Polyploidy

The first reported study on the number of chromosomes in conifers was by Sax and Sax (1933) using acetocarmine smears of seed megagametophyte cells. In a sample of 22 conifers, they found either 11 or 12 pairs of chromosomes for all species except *Pseudolarix amabilis*, for which they found 22 pairs. Over the next 30 years, dozens of papers were published on the chromosome counts in many species of conifers. The body of literature summarized and reviewed by Khoshoo (1961) included chromosome numbers in 264 gymnosperm species, most of which were conifers (Table 2.1). Haploid (1N) chromosome numbers in the Southern Hemisphere families Podocarpaceae and Araucariaceae range from 10 to 13 with the exception of species in *Podocarpus* section *Stachycarpus* subsection *Euprumnopitys* having 18 or 19. The 1N base number is 12 ($x = 12$) in the Pinaceae

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_2

Table 2.1 Chromosome number (2N) and genome size (1N) in several conifer species grouped by family

Genus	Species	Chromosome number	Genome size (Mbp)
Araucariaceae			
Agathis	*alba*	26	
	australis	26	
	dammara		13,300.80
	robusta	26	
Araucaria	*angustifolia*	26	
	araucana		22,249.50
	bidwillii	26	
	columnaris	26	
	cunninghamii	26	
	heterophylla	26	
Wollemia	*nobilis*		13,887.60
Cupressaceae			
Actinostrobus	*pyramidalis*	22	10,415.70
Athrotaxis	*cupressoides*	22	9877.80
	laxifolia	22	
	selaginoides	22	9828.90
Austrocedrus	*chilensis*		10,660.20
Callitris	*canescens*	22	
	columellaris	22	
	endlicheri	22	
	preissii	22	
	rhomboidea	22	8802.00
	verrucosa	22	
Calocedrus	*decurrens*		14,914.50
	formosana		16,626.00
Chamaecyparis	*formosensis*		8361.90
	lawsoniana	22	10,562.40
	obtusa	22	9095.40
	pisifera	22	9030.20
	thyoides		10,171.20
Cryptomeria	*japonica*	22	10,073.40
Cunninghamia	*konishii*	22	12,469.50
	lanceolata	22	19,364.40
Cupressus	*arizonica*	22	11,051.40
	cashmeriana	22	
	duprezziana	22	
	funebris	22	
	guadalupensis	22	
	leylandii	22	
	lusitanica	22	10,513.50
	macnabiana	22	
	macrocarpa	22	

(continued)

Table 2.1 (continued)

Genus	Species	Chromosome number	Genome size (Mbp)
	nootkatensis		11,149.20
	sempervirens	22	11,149.20
	torulosa	22	
Diselma	*archeri*		8850.90
Fitzroya	*cupressoides*		17,115
Fokienia	*hodginsii*		10,855.80
Glyptostrobus	*pensilis*		8850.90
Juniperus	*bermudiana*	22	
	cedrus		10,953.60
	chinensis		24,205.50
	communis	22	10,843.60
	foetidissima		12,909.60
	formosana	22	
	horizontalis	22	11,719.70
	monosperma	22	
	oxycedrus		11,442.60
	phoenicea	22	
	pingii		21,760.50
	procera	22	
	rigida	22	
	sabina	22, 24	21,907.20
	scopulorum		11,491.50
	squamata	44	11,882.70
	virginiana	22	11,687.10
Libocedrus	*bidwillii*	22	
	plumosa	22	
Metasequoia	*glyptostroboides*	22	9437.70
Microbiota	*decussata*		8948.70
Neocallitropsis	*pancheri*		12,469.50
Pilgerodendron	*uviferum*		13,740.90
Platycladus	*orientalis*		9951.20
Sequoia	*sempervirens*	66	28,215.30
Sequoiadendron	*giganteum*	22	10,171.20
Taiwania	*cryptomerioides*	22	11,418
Taxodium	*distichum*	22	8997.60
	huegelii	22	8850.90
Tetraclinis	*articulata*	22	12,567.30
Thuja	*koraiensis*		11,687.10
	occidentalis	22	11,344.80
	plicata	22	11,051.40
	standishii	22	12,127.20
	sutchuenensis		11,833.80
Thujopsis	*dolabrata*	22	11,638.20
Widdringtonia	*nodiflora*	22	10,171.20

(continued)

Table 2.1 (continued)

Genus	Species	Chromosome number	Genome size (Mbp)
	schwarzii		9584.40
	wallichii		10,513.50
Pinaceae			
Abies	*balsamea*	24	16,039.20
	cephalonica	24	17,701.80
	concolor	24	
	homolepis		19,413.30
	koreana		17,848.50
	magnifica	24	
	nordmanniana	24	17212.8
	pindrow	24	
	pinsapo		18,630.90
	procera		16,821.60
	sibirica	24	
	veitchii	24	
Cathaya	*argyrophylla*		24,205.50
Cedrus	*atlantica*	24	
	deodara	24	16,234.80
	libani	24	15,501.30
Keteleeria	*evelyniana*	24	23,667.60
Larix	*decidua*	24, 48	12,714.00
	eurolepis	24	
	gmelinii	24, 36	12,567.30
	griffithii		13,105.20
	kaempferi	24	12,909.60
	occidentalis	24	
	polonica	24	
	sibirica	24	
Picea	*abies*	24	19,902.30
	alcoquiana	24	
	asperata	24	
	engelmannii	24	
	glauca	24	
	jezoensis	24	
	koyamae	24	
	likiangensis	24	
	mariana	24	
	maximowiczii	24	
	omorika	24	18,533.10
	orientalis	24	
	pungens	24	20,586.90
	rubens	24	
	sitchensis	24	
	smithiana	24	

(continued)

Table 2.1 (continued)

Genus	Species	Chromosome number	Genome size (Mbp)
Pinus	*albicaulis*		31,589.40
	aristata		27,824.10
	armandii	24	31,882.80
	ayacahuite	24	
	balfouriana		23,569.80
	banksiana	24	22,249.50
	bungeana	24	32,176.20
	canariensis	24	31,442.70
	caribaea	24	22,151.70
	cembra	24	30,098.00
	cembroides	24	
	clausa	24	
	contorta	24	19,022.10
	culminicola		28,557.60
	densiflora	24, 48	24,498.90
	echinata	24	
	edulis	24	28,753.20
	elliottii	24, 36	
	flexilis	24	30,513.60
	gerardiana	24	35,061.30
	halepensis	24	25,819.20
	heldreichii		29,584.50
	jeffreyi	24	
	kesiya	24	
	koraiensis	24	30,318.00
	lambertiana	24	
	longaeva		25,672.50
	luchuensis	24	
	massoniana	24	25,134.60
	merkusii	24	30,220.20
	monophylla		30,953.70
	montezumae	24	
	monticola	24	29,975.70
	mugo	24	22,119.10
	nigra	24	24,803.60
	palustris	24	
	parviflora	24	29,731.20
	patula	24	
	peuce	24	
	pinaster	24	28,264.20
	pinea	24	27,970.80
	ponderosa	24	21,809.40
	pumila		29,046.60
	pungens	12	

(continued)

Table 2.1 (continued)

Genus	Species	Chromosome number	Genome size (Mbp)
	radiata	24, 48	23,716.50
	resinosa	24	
	rigida	24	
	roxburghii	24	30,171.30
	sabiniana	24	
	sibirica		28,899.90
	strobus	24	28,557.60
	sylvestris	24	23,667.60
	tabuliformis	24	
	taeda	24	21,662.70
	thunbergii	24	24,841.20
	virginiana	24	19,902.30
	wallichiana	24	28,997.70
Pseudolarix	*amabilis*	44	25,525.80
Pseudotsuga	*menziesii*	26	17,212.80
Tsuga	*canadensis*	24	18,533.10
	caroliniana	24	20,489.10
	chinensis		20,391.30
	diversifolia	24	
	heterophylla		16,821.60
	jeffreyi		16,674.90
	mertensiana		17,750.70
	sieboldii		18,386.40
Podocarpaceae			
Acmopyle	*pancheri*		8557.50
	sahniana		6748.20
Afrocarpus	*falcatus*		5672.40
	gracilior	24	
	mannii		4987.80
Dacrycarpus	*imbricatus*		5232.30
Dacrydium	*balansae*		6894.90
	elatum		6405.90
	gracile		6503.70
	nausoriense		6846.00
	nidulum		6797.10
Falcatifolium	*taxoides*		10,953.60
Halocarpus	*bidwillii*		8215.20
Lagarostrobos	*franklinii*		4938.90
Lepidothamnus	*fonkii*		4645.50
Manoao	*colensoi*		13,545.30
Microcachrys	*tetragona*		4058.70
Nageia	*nagi*		5476.80
Pherosphaera	*fitzgeraldii*		4205.40
	hookeriana		4107.60

(continued)

Table 2.1 (continued)

Genus	Species	Chromosome number	Genome size (Mbp)
Phyllocladus	*aspleniifolius*		7775.10
	trichomanoides		7188.30
Podocarpus	*cunninghamii*		8264.10
	gnidioides		8459.70
	lawrencei		7775.10
	macrophyllus		9486.60
	nivalis		8215.20
Prumnopitys	*ladei*		5476.80
Retrophyllum	*rospigliosii*		5770.20
Saxegothaea	*conspicua*		4987.80
Sundacarpus	*amarus*		6650.40
Sciadopityaceae			
Sciadopitys	*verticillata*	20	18,949.00
Taxaceae			
Amentotaxus	*argotaenia*	22	
	yunnanensis		29,535.60
Cephalotaxus	*fortunei*	24	
	harringtonii	24	25,672.50
Pseudotaxus	*chienii*		16,919.40
Taxus	*baccata*		11,137.00
	brevifolia		11,198.10
	canadensis		11,344.80
	cuspidata	24	11,198.10
	hunnewelliana	24	
	media	24	
	sumatrana		10,904.70
	wallichiana		11,589.30
Torreya	*californica*		21,564.90
	nucifera	22	21,809.40
	taxifolia		21,124.80

Chromosome numbers from Khoshoo (1961) and genome sizes from Zonneveld (2012)

(with the exception of *Pseudotsuga menziesii* with $x = 13$), $x = 11$ in the Cupressaceae and the Taxodiaceae (now included in the Cupressaceae), and both $x = 11$ and $x = 12$ in the Taxaceae. Based on the phylogenetic relationships among taxa of the conifers (see Chap. 16), it would appear that the ancestral number of chromosomes is $x = 11$.

Polyploidy is ubiquitous in the angiosperms and is regarded as one of the primary mechanisms of evolution and speciation in that group (Soltis et al. 2010). In the gymnosperms, however, and especially in the conifers, polyploidy is extremely rare (Khoshoo 1959; Ahuja 2005; Yang et al. 2012). There are many examples of naturally rare or chemically induced polyploids (Ahuja 2005); however, there are only three known naturally occurring and widespread polyploids in conifers and all are from the Cupressaceae: *Fitzroya cupressoides* (Alerce) (Hair 1968),

Juniperus chinensis "Pfitzeriana" (Sax and Sax 1933), and *Sequoia sempervirens* (coast redwood) (Hirayoshi and Nakamura 1943; Stebbins 1948). Alerce and Pfitzeriana are tetraploid ($4x = 44$) and coast redwood is hexaploid ($6x = 66$).

The origin of the tetraploid Pfitzeriana clearly seems to be from hybridization between the diploids *J. chinenesis* and *J. sabina* (De Luc et al. 1999). However, the origins of the tetraploid Alerce and hexaploid coast redwood are not fully resolved. Ahuja (2005) suggested that Alerce is an autotetraploid, whereas in a more recent study, Yang et al. (2012) suggested that it may be an allotetraploid from a hybrid between *F. cupressoides* and a species of *Diselma*. In the case of coast redwood, based on meiotic chromosome preparations, Stebbins (1948) first proposed an allopolyploid origin whereas Ahuja and Neale (2002) proposed that it may have originated from an autoallopolyploid or segmental polyploidy event. In a much more recent study based on transcriptome sequencing in coast redwood and the related *Sequoiadendron giganteum* (giant sequoia) and *Metasequoia glyptostroboides* (dawn redwood), Scott et al. (2016) concluded that coast redwood was of autopolyploid origin. Ultimately, full genome sequencing may resolve the origin of this sole example of hexaploidy in conifers.

Genome Size

It has been known since the pioneering work of Miksche (1967) that conifer genomes are very large. The first comprehensive survey of conifer genome sizes was done by Ohri and Khoshoo (1986). Later, Murray (1998) reported genome sizes for 117 gymnosperm species. The early literature on the estimation of DNA content, or C-value, was reported in picogram units, whereas in the more recent literature it is reported in units of base pairs. For simplicity, we will use base pair (bp) units and convert picograms (pg) to base pairs with the approximate conversion of 1 pg = 980,000,000 bp = 980 megabase pairs (Mb). The most recent and comprehensive survey of conifer genome sizes is that of Zonneveld (2012) where estimates of 172 species from 64 of the 67 genera of conifers are provided. In that paper, genome-size estimates are given in 2C pg amounts; we have converted these to 1C Mb estimates (Table 2.1).

Estimates of genome size range from 4067 Mb (*Microcachrys tetragona*) to 35,084 Mb (*Pinus gerardiana*)—nearly a tenfold difference. One species not measured in this study was *Pinus lambertiana* (sugar pine), recently sequenced by Stevens et al. (2016) with an estimated genome size of 31,000 Mb. By contrast, many of the angiosperm tree species have genome sizes in the range of 500 to 1000 Mb. In general, species of the Pinaceae have larger genomes (~20,000 Mb+) than species in other families of conifers.

One topic that interested researchers in the early years was intraspecific variation in DNA content and any adaptive differences found within species that might be associated with differences in DNA content. Note that at this stage in the development of conifer genomics, there was no idea with regard to how many

protein-coding genes there were within a species or that differed between species, so the intraspecific differences observed would be due to either coding DNA, noncoding DNA, or both. Now that several conifer genomes have been completely sequenced (Chap. 3), we can see that these differences are almost entirely in the amount of noncoding DNA.

In an early study on intraspecific variation in DNA content, Miksche (1968) showed differences in genome size among populations of *Picea glauca*, *P. sitchensis*, and *Pinus banksiana* and that there was an increase in size from south to north in latitude. Likewise, El-Lakany and Sziklai (1971) and Dhir and Miksche (1974) found similar results in *Pseudotsuga menziesii* and *Pinus resinosa*, respectively. In other studies, however, either no differences were found among populations (Dhillon et al. 1978) or no correlations with latitude were found (Joyner et al. 2001). In one study, a relationship with longitude was found (Bogunic et al. 2006). In one of the more comprehensive studies on this topic, Wakamiya et al. (1993) found correlations with genome size and growth and seed characteristics as well as with climatic factors in several species of *Pinus*. Interest in this topic has declined in recent years, likely due to inconclusive or contradictory results. What have emerged in its place are studies of landscape genomics (Chap. 12) where differences in allele frequencies across geographical or environmental gradients or both have been observed in protein-coding loci that may underlie complex adaptive traits.

It is clear that genome size varies considerably among conifer species which leads to the question, how and why? Are there evolutionary or adaptive explanations to these differences or are these differences purely random and not causative in some way? Ahuja and Neale (2005) attempted to address this question and gave four hypotheses for the large genomes of conifers: (1) whole-genome duplication, polyploidy; (2) duplication of specific sets of genes; (3) repetitive DNA content; and (4) large intron sizes in coding genes. As we have seen in the large survey of Zonneveld (2012) and other reports, there is a clear lack of evidence for polyploidy in conifers with just a couple of exceptions. However, in a recent paper by Li et al. (2015), the authors argue that three ancient whole genome duplications occurred in gymnosperms and that polyploidy has in fact been important in the evolution of conifers. This assertion is based entirely on data from transcriptome sequencing (Chap. 3); full genome sequencing will provide much more insightful data and ultimately resolve this debate.

The other three hypotheses were difficult to accurately assess in 2005 without data derived from large gene-sequencing or whole-genome-sequencing projects. As we will see in Chap. 3, conifers have more or less the same number of genes as all other plants. There are clearly cases where gene families have amplified in classes of genes important to the function of woody perennials but these differences do not nearly account for the large genome sizes. Likewise, it has been shown that gene intron lengths in conifers can often be quite large (Chap. 5), but again this difference cannot account for the large genome sizes. It has become clear that the large sizes of conifer genomes can be attributed to the amplification of some types of repetitive DNA. This topic will be covered in detail in Chap. 4.

Karyotype Analysis

The study of conifer chromosomes by karyotype analysis began more than 100 years ago; the first published studies were by Ferguson (1904) and Lewis (1908). The basic approach to karyotype analysis is to make chromosome preparations from some cell type, often dividing root-tip cells, and then stain the chromosome preparations with a dye and visualize the chromosomes with a microscope. The number and morphology are depicted in photographs or idiograms (Fig. 2.1). The questions most often addressed in the early conifer karyotype studies were: (1) Could chromosomes within species be distinguished from one another based on size and morphology? (2) Could differences in chromosomes distinguish among species and do these differences generally follow accepted phylogenies? (3) Are there differences in chromosomes that can be attributed to the population of origin or due to influences of the environment? As would be expected, contrasting results to these three questions were obtained. An important question not addressed in these early karyotype studies was whether homologous chromosomes among species could be identified based on karyotype. This question was not really addressed until fluorescent in situ hybridization (FISH) techniques became available (see below).

In a series of papers by authors such as Saylor (1961, 1964, 1972), Pedrick (1967, 1968, 1970), Schlarbaum and Tsuchiya (1975a, 1975b, 1976), and more (see White et al. 2007, Chapter 2), it was well established that conifer chromosomes within species were often very similar in size and not easily distinguished but when stained, differences were revealed and simple idiograms could be drawn to represent the morphological differences among chromosomes.

Results from studies comparing karyotypes among related species have produced inconsistent results. In some taxa, clear differences in karyotype were observed and paralleled phylogenetic relationships (Schlarbaum et al. 1983; Schlarbaum and Tsuchiya 1984a; Nkongolo and Klimaszewska 1995; Dagher-Kharrat et al. 2001) and in other studies few differences among species were found (Nkongolo 1996; Mehes-Smith et al. 2011).

Similar to studies that observed differences in DNA content due to population origin, a few studies have observed differences in chromosome morphology due to population of origin (Guttenberger et al. 1996; Sedelnikova and Muratova 2002). Finally, there were a few studies that observed chromosome aberrations due to cultural conditions (Owens 1967; Korshikov et al. 2012).

FISH provided an important technical advance in cytogenetics and karyotype analysis in conifers (Islam-Faridi and Nelson 2011). The first paper reporting use of this technology in conifers was by Brown et al. (1993) in *Picea glauca*. Subsequently, several reports using FISH in pine (Doudrick et al. 1995; Lubaretz et al. 1996; Hizume et al. 2002; Liu et al. 2003a; Shibata et al. 2005; Cai et al. 2006; Islam-Fardi et al. 2007; Deng et al. 2008; Mehes-Smith et al. 2011), spruce (Brown and Carlson 1997; Vischi et al. 2003; Shibata and Hizume 2008), and Douglas-fir (Amarasinghe and Carlson 1998) were published. The general approach is to hybridize a fluorescent-labeled DNA probe to chromosome preparations and then visualize and

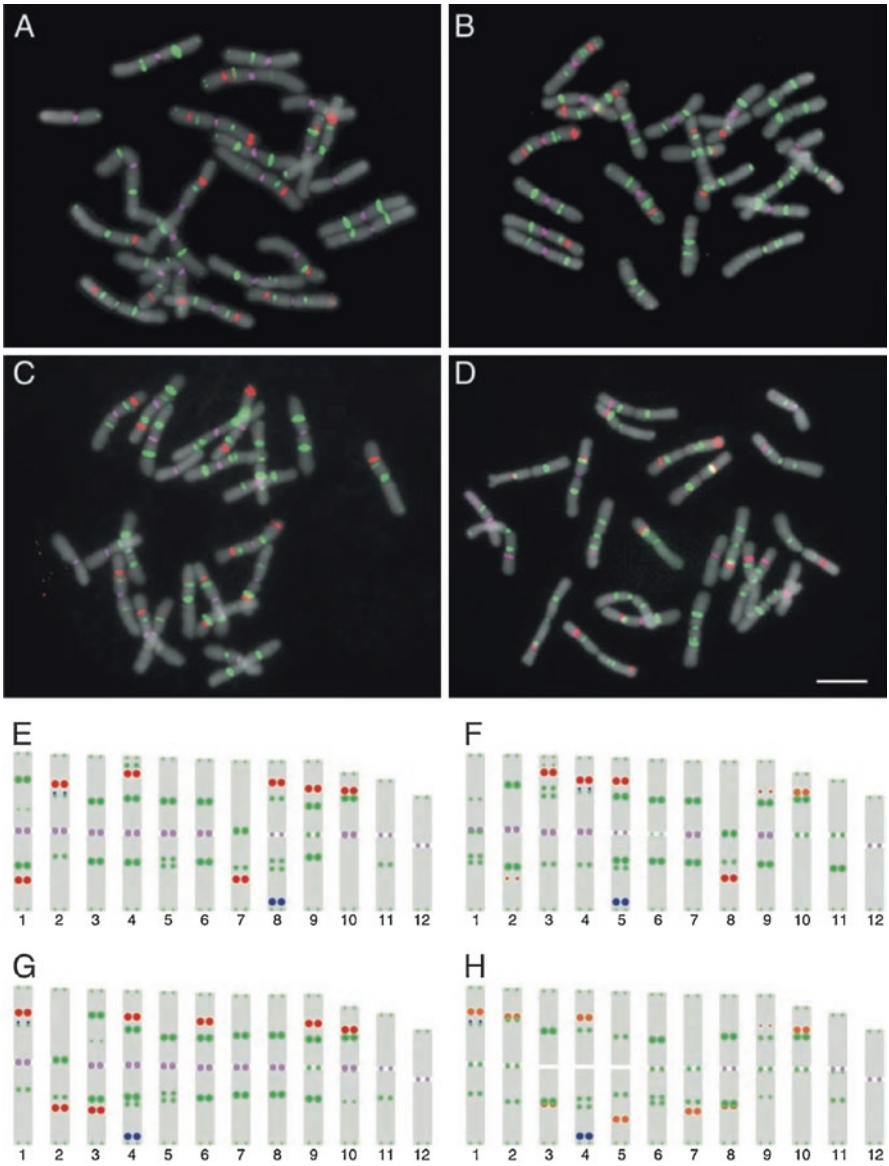


Fig. 2.1 FISH images of somatic chromosomes (**a**–**d**), and idiograms of FISH karyotypes (**e**–**h**), of *Pinus densiflora* (**a**, **e**), *P. thunbergii* (**b**, **f**), *P. sylvestris* (**c**, **g**), and *P. nigra* (**d**, **h**). (**a**–**d**) The *red* signal corresponds to the 45S and 5S rDNA probes, *magenta* represents the *PCSR* probe, and *green* corresponds to the telomere sequence probe. (From Hizume et al. (2002))

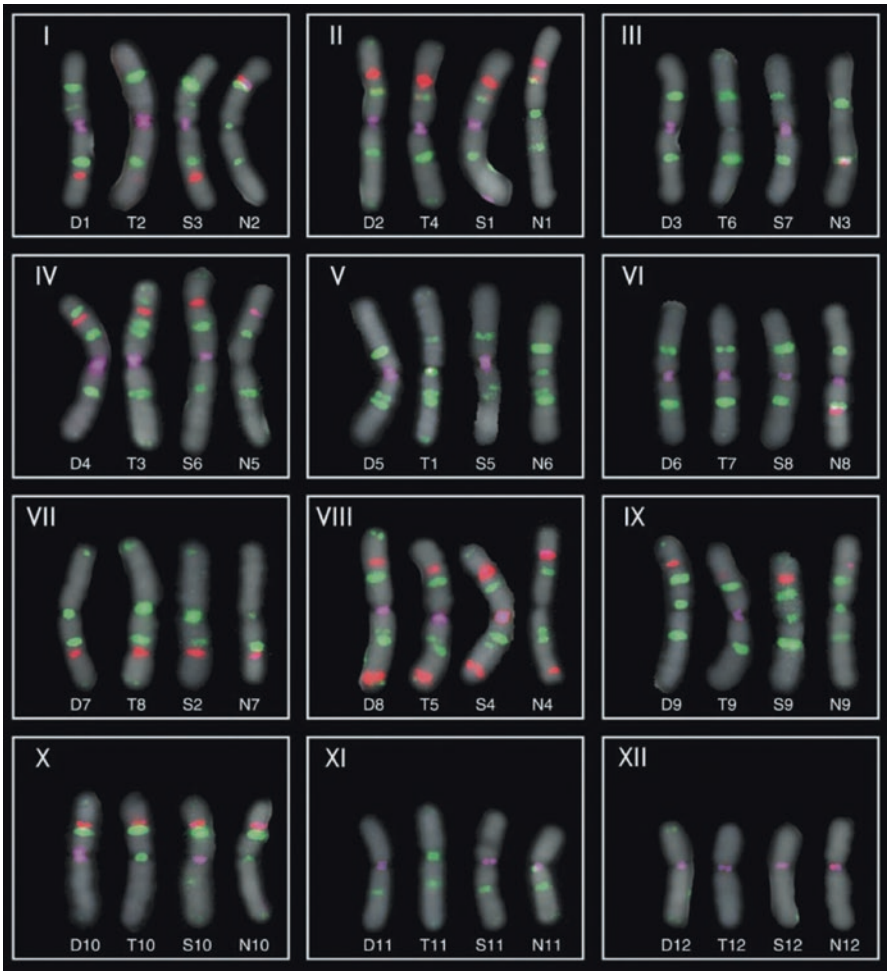


Fig. 2.2 Comparison of FISH chromosomes probed with 45S rDNA and 5S rDNA (*red*), *PCSR* (*magenta*), and telomere sequence (*green*) probes in four *Pinus* species. The chromosome groups are numbered *I–XII*. The *letters* (*D*, *T*, *S*, *N*) appended to each chromosome number identify *P. densiflora*, *P. thunbergii*, *P. sylvestris*, and *P. nigra*, respectively. (From Hizume et al. (2002))

photograph with fluorescent microscopy. DNA probes used are from a highly repeated DNA sequence such as ribosomal DNA (rDNA) or telomeric repeat sequences. An example of FISH in four pine species is shown in Fig. 2.2. FISH technology can clearly distinguish chromosomes within species and differences among species. It has now become possible to link genetic linkage maps based on segregation analysis (Chap. 11) to individual chromosomes. An advance of fundamental importance provided by FISH is the assignment of homologous chromosomes among species which is necessary for a variety of comparative genomic and evolutionary analyses (Chap. 17).

There has been recent advancement in techniques for chromosome banding. Using chromomycin A_3 and 4′,6-diamidino-2-phenylindole (DAPI) banding in a number of conifers, Hizume (2015a, b), Hizume and Kan (2015), and Hizume et al. (2016) have improved the resolution and differentiation of conifer chromosomes significantly.

Genome Content

The classical era of genome studies in conifers provided a great deal of new information about chromosome number, chromosome morphology, and genome size through cytogenetic and karyotyping techniques. What remained almost entirely unknown was the composition of these chromosomes and genomes. The acquisition of this knowledge would await the application of recombinant DNA technologies. However, a couple of pioneering papers were published in the early 1980s in advance of recombinant-DNA-based research. The technique of DNA reassociation kinetics was first used by Rake et al. (1980) in two pine and two spruce species to estimate the relative fractions of the genome by their repetitive content. This study established that a very significant portion of the genome was of a highly repeated nature, but what these sequences were could not be determined. The authors speculated that the large size and repetitive DNA content of the pine genomes, relative to the spruces, may be due to higher ploidy. One of the pines studied was *Pinus lambertiana*; it would not be until 31 years later that its genome was sequenced (Stevens et al. 2016) and the exact composition of all this repetitive DNA would become known (Chaps. 3 and 4): no evidence for polyploidy was found in this study.

In an elegant reassociation kinetics study by Kriebel (1985), it was determined that the DNA fractions did not fall into discrete classes, as shown in the seminal work in sea urchins by Britten and Kohne (1968), but were distributed in a more continuous manner (Table 2.2). Kriebel (1985) further conjectured that much of the genome was probably made up of ancient and diverged repeated sequences and that only 0.1% of the genome might code for protein-coding genes. Thirty years later following the sequencing of the first conifer genomes (Chap. 3 and 4) we can see that Kriebel was exactly correct.

Table 2.2 Sequence components of the *Pinus strobus* genome

Component	Fraction of fragments	Average no. of copies	Complexity in nucleotide pairs
Highly repetitive	0.11	18,800	2.8×10^4
Middle repetitive	0.4	948	2.1×10^6
Low repetitive	0.18	77	1.2×10^7
Single copy	0.24	1	1.3×10^9

Modified from Kriebel et al. (1985)

In a short review paper by Kinlaw and Neale (1997), it was proposed that conifer genomes may have much larger gene family sizes than in angiosperms. This assertion was based on Southern hybridizations of genomic conifer DNA with cDNA probes. Again, 30 years later following the sequencing of conifer genomes, it can be seen that this hypothesis is only partly correct and much of what was observed in the early years was certainly pseudogene sequences and not large functional multigene families.

A more precise and detailed understanding of the content of conifer genomes would not come until 2009–2010 when the first papers were published on sequencing a few large pieces of DNA from conifers (Hamberger et al. 2009; Kovach et al. 2010). The large pieces of DNA were bacterial artificial chromosomes (BACs) which by the year 2000 had become the DNA cloning approach for sequencing large and complex eukaryotic genomes. In the first study, Hamberger et al. (2009) constructed a BAC library of 1.1M clones and then screened this library for clones that would contain genes of interest, in this case a terpenoid synthase gene and a cytochrome P450 gene. The two BAC clones containing these genes were then completely sequenced; the two clones being 172 kbp and 94 kbp long. It can be seen that the protein-coding genes sought by the researchers occupy only a very small part of the BAC clone and most of the clone is occupied by highly repeated sequence elements (Fig. 2.3). Furthermore, it can be seen that in the case of these two BAC clones only one protein-coding gene was found in a piece of DNA of approximately 100 kbp or more. An average genome-wide estimate of gene density could not accurately be estimated from a sample of two, but early indications were that gene density is not high and that novel strategies to sequence the entire genome would be needed.

In the second study, Kovach et al. (2010) sequenced 10 BAC clones from *Pinus taeda*, but in addition did a small amount of whole-genome shotgun sequencing

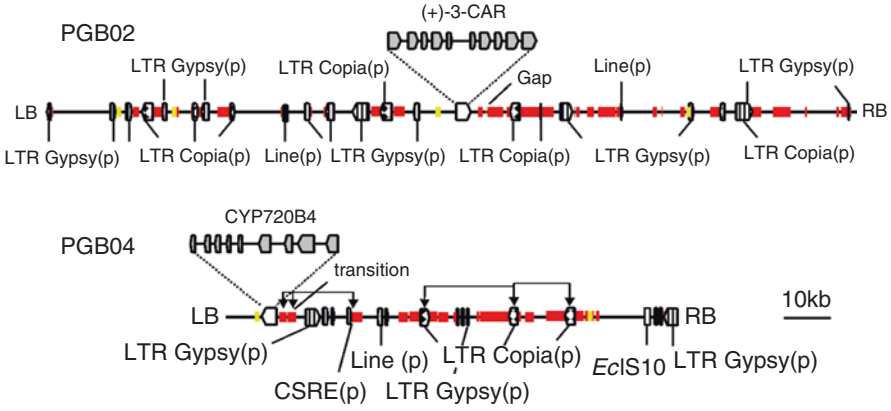


Fig. 2.3 Structure of *Picea glauca* genomic DNA of BAC clones PGB02 and PGB04. The position of the target genes 3CAR and CYP720B4 is indicated. Red and yellow bars represent repeated segments and segments with similarity to DNA transposons, respectively. (From Hamberger et al. (2009))

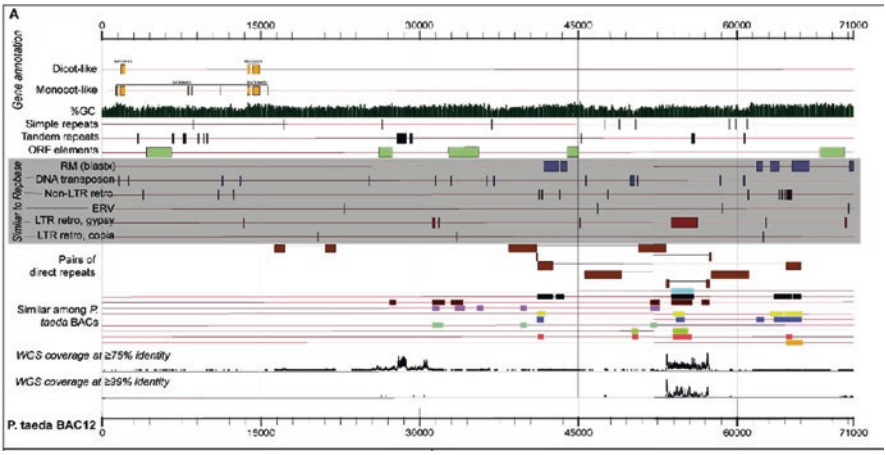


Fig. 2.4 *Pinus taeda* BAC12 (clone Pt314B2) illustrates several new trends found in the pine genome. The length of BAC12 is shown along the horizontal axis. Shown above the axis are tracks of annotated genes (dicot and monocot parameters), similarity hits to Repbase [RM (blastx); DNA transposons; Non-LTR retroelements; ERV (endogenous retroviruses); LTR retroelements, copia; LTR retroelements, gypsy], and other elements identified in this study (simple repeats, tandem repeats, ORF elements, pairs of direct repeats, and regions of similarity among BACs). The bottom two tracks indicate WGS coverage at ≥75% identity and at ≥99% identity. (From Kovach et al. (2010))

using next-generation sequencing technology. This study confirms the results of Hamberger et al. (2009) that the gene density in conifers is very low and that nearly all of the genome is comprised of various types of highly repetitive elements (Fig. 2.4). However, this study deepened the understanding of conifer genome content by showing that many of these repetitive elements were unique to conifers, and furthermore, most of these repetitive elements are very old and have diverged from one another over evolutionary time. This last assertion is important because it suggested that full-genome sequence assembly in conifers might not be as difficult as in some angiosperm species with smaller genomes, but with a much higher proportion of more recently duplicated repetitive DNA. In a final study of this type, Liu et al. (2011a) made very similar conclusions for the *Taxodium distichum* var. *distichum* genome, again based on partial sequencing of BACs.

Organelle Genomes: Chloroplast and Mitochondria

The organelle genomes were studied rather intensely in the classical genomics era, in large part because technologies of the time allowed access to these genomes, whereas access to the much larger nuclear genome was still quite limited. The primary reason for easier "access" was simply due to the fact that these genomes are very small in size and are found in very high copy number within cells, so that large amount of DNA could be isolated and molecular methods used for study. Topics of

study during this period included (1) genome size and structure, (2) gene content, (3) inheritance, and (4) variation. These topics were summarized in a review by Hipkins et al. (1994) and are briefly described here.

Size and Structure

The chloroplast genome in plants, derived from a cyanobacterium-like endosymbiont, is a small circular molecule found in multiple copies in each chloroplast; there are multiple chloroplasts in each photosynthetic cell. In angiosperms, this circular molecule is generally 140–160 kb. It was known that the conifer chloroplast genome was smaller, about 120 kb, but it was not known until the work of Strauss et al. (1988) that this was due to the lack of one of the two large inverted repeat regions that are found in angiosperm chloroplasts. Later, once a conifer chloroplast genome was sequenced for the first time (Wakasugi et al. 1994), it was seen in *Pinus thunbergii* that a small portion of the second repeat was present (Tsudzuki et al. 1992), thus it could be surmised that for some reason there was a loss of the second repeat in conifers following the divergence of angiosperms and gymnosperms. The mitochondrial genome was rarely studied in conifers in the classical era and all that was really known was that there were many copies per mitochondrion and that copies might differ in size.

Gene Content

Organelle DNA studies done during the classical era were conducted primarily using restriction mapping techniques and Southern hybridization using DNA probes cloned from angiosperm cpDNA. These combined approaches showed that conifer cpDNA had basically the same small gene content (tRNA, rRNA, and protein-coding genes) as angiosperms but this approach would not reveal genes that would be unique to conifer cpDNA. This would have to wait until the first conifer cpDNA genome was sequenced (Wakasugi et al. 1994, see Chapter 3) and likewise for conifer mtDNA genomes (Chap. 3).

Inheritance

The study of the inheritance of cpDNA and mtDNA was an active area of research in the classical era. In angiosperms it was known that both organelle genomes were almost always uniparentally and maternally inherited. In a study by Ohba et al. (1971), it was shown that a chlorophyll mutant in *Cryptomeria japonica* was transmitted by the male parent and that the cpDNA must then be paternally inherited. The first molecular confirmation of paternal inheritance of cpDNA in a conifer was shown in *Pseudotsuga menziesii* (Neale et al. 1986) using restriction fragment

length polymorphism (RFLP) markers. Paternal inheritance of cpDNA in conifers was then confirmed in a large number of studies; *Pinus* (Wagner et al. 1987, 1989, 1992; Neale and Sederoff 1989; Dong et al. 1992; Boscherini et al. 1994), *Picea* (Szmidt et al. 1988a; Stine et al. 1989; Stine and Keathley 1990; Sutton et al. 1991), *Larix* (Szmidt et al. 1987), *Sequoia* (Neale et al. 1989), and *Calocedrus* (Neale et al. 1991).

So if the cpDNA was paternally inherited in conifers, would the mtDNA also be inherited paternally or would it still be strictly maternally inherited as it is in angiosperms and mammals? The first reports from members of the Pinaceae all showed maternal inheritance of mtDNA; *Pinus* (Neale and Sederoff 1989; Wagner et al. 1991a), *Picea* (Sutton et al. 1991), *Pseudotsuga* (Marshall and Neale 1992), and *Larix* (Deverno et al. 1993). However, the very surprising result of paternal inheritance of mtDNA was shown for two members of the Cupressaceae; *Sequoia* (Neale et al. 1989) and *Calocedrus* (Neale et al. 1991). Unfortunately, there has been very little follow-up to these early studies of inheritance of organelles in conifers done with ultrastructural approaches to understand the cellular mechanisms responsible for these novel modes of inheritance.

Variation

The general dogma that had arisen from studies in angiosperms during this period was that genetic variation could be found among species that was informative for evolutionary and phylogenetic studies, but intraspecific variation was very low or nonexistent. Conifer geneticists discovered enough intraspecific variation in a small number of full-sib crosses in order to establish inheritance of cpDNA and mtDNA, but population-level studies would need to be done to establish better estimates of population diversity of organelle genomes. These types of studies will be discussed in greater detail in Chap. 9. The first and largest study of this type was that of Wagner et al. (1987) in which 902 trees of *Pinus contorta*, *P. banksiana*, and their hybrids in a zone of introgression were typed for cpDNA RFLP variation (Chap. 9). This study established that intraspecific cpDNA variation in conifers could be found and used in population genetic studies. Subsequently, many other studies were done with smaller population samples to establish the existence of intraspecific variation in a number of conifer species (Szmidt et al. 1988a; White 1990; Ali et al. 1991; Hong et al. 1993; Ponoy et al. 1994; Nelson et al. 1994; Hipkins et al. 1995; Tsumura et al. 2000). Likewise, a small number of studies also established the intraspecific mtDNA variation could also be found in conifers (Strauss et al. 1993; Dong and Wagner 1993). A fundamentally important result of these findings is that researchers working in population genetics of species of the Pinaceae now had an experimental system where the maternal lineage (mtDNA, maternal gamete), paternal lineage (cpDNA, paternal gamete), and biparental lineage (nDNA, zygote) could now all be followed simultaneously from the same set of population samples (Chap. 9).

Summary

The *classical era* of genome research in conifers is best characterized as the study of chromosomes using cytogenetic techniques. The chromosome number for a large number of species was determined to vary little ($x = 11$ to $x = 13$), with just a couple of exceptions of polyploid species. However, what does vary more significantly is total DNA content, varying from ~4000 Mbp to ~35,000 Mbp; nearly an order of magnitude difference. Furthermore, genome sizes in conifers can be 20–40X larger than those of most angiosperm species. Large genomes prohibited full genome sequencing in conifers for a decade or more until next-generation sequencing technologies became available. Karyotyping technologies steadily evolved during this period, notably the application of fluorescent in situ hybridization (FISH) such that morphological differences among chromosomes and among species could readily be seen. As genome research has moved into the *modern era* (Chap. 3), few researchers trained in classical cytogenetic techniques remain, an unfortunate circumstance given this research approach still has much to contribute.

Gene and Genome Sequencing in Conifers: Modern Era

3

A Short History of DNA Sequencing in Conifers

We have defined the period up to the late 1990s as the *classical era* of the study of conifer genomes (Chap. 2) and everything after that as the *modern era*. The distinction between these two eras is based largely on the availability of DNA sequences. DNA sequencing of conifer DNA in fact began much earlier. The first report of sequencing of conifer DNA, to our knowledge, was that of Kenny et al. (1988). In this study, Kenny et al. (1988) cloned a small piece of *Pinus contorta* genomic DNA (gDNA) and sequenced the DNA manually using the chain termination method of Sanger (Sanger et al. 1977). They then compared the DNA sequence and the translated amino acid sequence to other published actin gene sequences. In the decade that followed, there were dozens of similar reports where short pieces of DNA (either from gDNA or complementary DNA (cDNA)) were sequenced and compared to sequence entries in growing databases of DNA sequences. This very early period of DNA sequencing will be covered briefly as it pertains to an understanding of gene structure in conifers (Chap. 5). In this chapter, we will begin in the late 1990s with high-throughput expressed sequence tag (EST) sequencing, the primary technology used to study conifer genomes for the ensuing 15 years or more. Then we will cover gene sequencing using a next-generation sequencing (NGS) technology, called RNA-seq, that began in 2010. Finally, we will summarize the work on full genome sequencing in conifers that began in 2013.

Expressed Sequence Tag Sequencing

Adams and Kelley (1991) first presented the idea for a rapid method to generate DNA sequences from protein-coding genes called *expressed sequence tags* (ESTs). The idea was to first make cDNA from some tissue (in their case, it was human brain tissue) and then partially sequence a large random sample of these cDNAs. The partial random EST nucleotide sequences, translated amino acid sequences, or both

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_3

were then compared to sequences in databases, and their identity and function might be inferred. This approach was proposed as an alternative to full genome sequencing of eukaryotic organisms which was not technically or economically feasible at the time. A literature search in 2017, using the key words "expressed sequence tags," yielded 726,000 papers as a testimony to the impact of this technology.

The predecessor to EST sequencing in conifers was the sequencing of a small number of anonymous cDNAs that were used as restriction fragment length polymorphism (RFLP, Chap. 11) probes in *Pinus taeda* (Kinlaw et al. 1997) and *Pseudotsuga menziesii* (Jermstad et al. 1998). The goal of these very early studies was to simply assign a gene identity to a genetic marker residing on a genetic map.

The first large EST sequencing report in a conifer was for *Pinus taeda* (Allona et al. 1998) (Table 3.1). In this study, RNA was isolated and cDNA constructed from xylem tissue of compression wood of 6-year-old trees. Sequences were obtained from 1097 clones and then compared to databases to infer the identity and function of these expressed genes. Not surprisingly, genes involved in cell-wall formation and lignin and carbohydrate synthesis and metabolism were found. Nevertheless, this paper marked a major breakthrough, characterizing, for the first time, the type of genes being expressed in a specific tissue and developmental state. Over the next 10 years, many similar EST sequencing papers were published (Table 3.1), almost exclusively in just a few pine and spruce species, undoubtedly due to the significant cost of doing this research.

In *Pinus taeda,* there were three large EST projects in the United States. Kirst et al. (2003) continued and expanded on the work of Allona et al. (1998) generating 59,797 EST sequences from xylem from a suite of developmental states. A very interesting result from this study was that when a rigorous test of sequence homology between *Pinus taeda* and *Arabidopsis thaliana* was applied, it was found that 90% of the sequences were in common, despite the great evolutionary divergence of these species (Fig. 3.1). This early result suggested that there may be great conservation in gene content and function between gymnosperms and angiosperms.

Table 3.1 Results reported from a few large EST sequencing projects in conifers

Species	Tissue	EST (no.)	Contig (no.)	Singleton (no.)	References
Cryptomeria japonica	Male strobili	36,011	7,686	15,972	Futamura et al. (2008)
Picea glauca	Many	49,101	9,354	7,224	Pavy et al. (2005)
Picea mariana	Needle	4,594	497	2,234	Mann et al. (2013)
Picea sitchensis, P. glauca	Many	147,146	19,941	26,804	Ralph et al. (2008)
Pinus radiata	Xylem	6,389	952	2,352	Li et al. (2009a)
Pinus taeda	Xylem	1,097	107	736	Allona et al. (1998)
Pinus taeda	Xylem	59,797	8,070	12,307	Kirst et al. (2003)
Pinus taeda	Embryos	68,721	5,274	6,880	Cairney et al. (2006)
Pinus taeda	Water stressed roots	12,918	6,765	n/a	Lorenz et al. (2006)

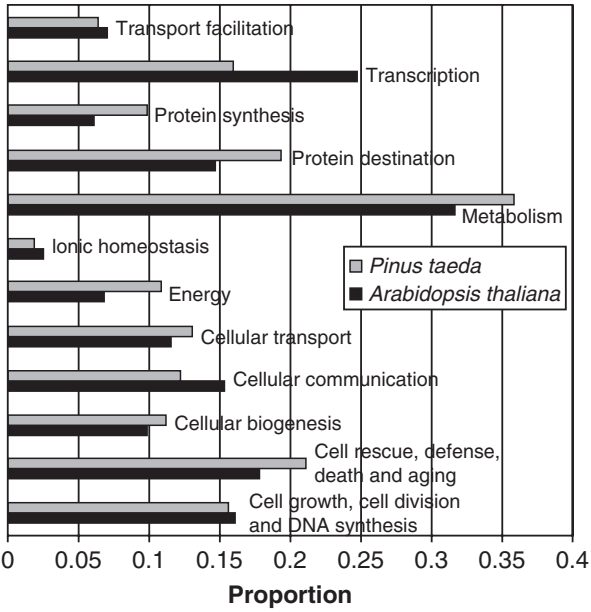


Fig. 3.1 *Pinus taeda* xylogenesis UniGene set, classified by cellular functional categories, compared with *Arabidopsis thaliana*. The proportion of *A. thaliana* genes in each functional category is relative to the 12,922 total predicted genes that were assigned by the *Arabidopsis* Genome Initiative. (From Kirst et al. (2003). Copyright (2003) National Academy of Sciences, USA)

However, as research has progressed, it can be seen that this may be true in many cases, but in other cases, there appear to be novel genes and functions in conifers (Chap. 17). Pavy et al. (2005) added value to this resource with a more advanced bioinformatics analysis of the Kirst et al. (2003) ESTs. In a second EST project, Cairney et al. (2006) reported on 68,721 ESTs from zygotic and somatic embryos. When compared to the existing database of *Pinus taeda* ESTs, they found that 28% were unique to the embryo ESTs, but when compared to a small set (108) of angiosperm embryogenesis-related genes, 77% were found in common. So, like the xylem EST studies, the general conclusion could be made that conifers do not have a large number of unique genes, but the profound differences in the whole plant phenotypes must be related to the expression of the genes and the interactions among these genes. In the third *Pinus taeda* EST project of that era, Lorenz et al. (2006) generated 12,918 ESTs from root tissues of drought-stressed rooted cuttings and found suites of genes previously implicated in drought tolerance in other plants.

Researchers in Canada were also productive contributors to conifer EST resources with their work in spruce. Pavy et al. (2005) completed a comprehensive study of 49,101 ESTs generated from 16 different tissue types and experimental treatments in *Picea glauca*. The complete collection of ESTs was then classified as belonging to different protein families (Fig. 3.2). Even though this presentation combines expressed genes over different tissues and treatments, it illustrates the abundance of mRNAs from different types of genes. We will see in Chap. 6 how ESTs were used to make cDNA microarrays to study gene expression more precisely. In an even more comprehensive study in several spruces (*Picea sitchensis*, *P. glauca*, and the *P. glauca-P. engelmannii* complex), Ralph et al. (2008) produced

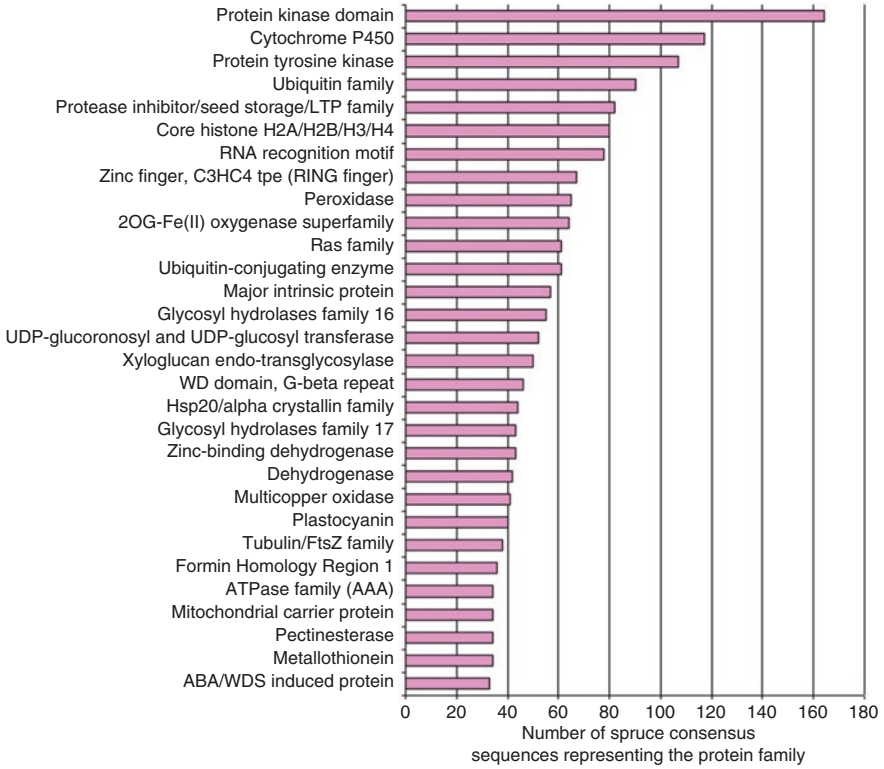


Fig. 3.2 Occurrence of the 30 most abundant protein families in the *Picea glauca* EST project. (From Pavy et al. (2005))

206,875 ESTs from 20 different cDNA libraries. A unique and extremely valuable contribution of this project was the development of 6464 sequences of the entire length of the mRNA, called *full-length cDNAs* (FLcDNAs). Standard EST projects produced only partial sequences which are sufficient for making a prediction of the identity and function of the gene but do not provide the complete coding sequence of the gene that is necessary for more detailed studies of gene structure and function. Subsequently the Canadian group that produced the original *P. glauca* EST resource (Pavy et al. 2005) produced a much more complete resource including FLcDNAs (Rigault et al. 2011).

Smaller EST resources were also developed in other conifers. These included 6389 ESTs from earlywood and latewood tissues in *Pinus radiata* (Li et al. 2009a). This study revealed ESTs unique to the lower quality juvenile wood. This information could be used in breeding or genetic engineering for trees with a higher proportion of mature latewood. In *Cryptomeria japonica*, Futamura et al. (2008) produced 19,437 ESTs from pollen strobili. Allergenic pollen is a large public health issue in Japan, and this information might be used to breed or engineer male sterile or less allergenic trees.

Specific bioinformatic and database resources were also developed to support the application and curation of EST resources. Resources such as ConiferEST (Liang et al. 2007), EuroPineDB (Fernández-Pozo et al. 2011), and Conifer DBMagic (Lorenz et al. 2012) were built, but unfortunately some of these resources are no longer available. The long-standing forest tree genome database TreeGenes (https://treegenesdb.org) curates the most complete collection of conifer ESTs.

Gene Discovery Using Next-Generation Sequencing

The Sanger sequencing method was used to generate ESTs and conduct gene discovery in conifers through the late 2000s. By the end of that decade, several companies introduced NGS technologies that dramatically increased throughput and decreased cost. The technique called RNA-seq was introduced which is very similar to the approach used in EST sequencing where mRNA is isolated first, converted to cDNA, and then sequenced. The fundamental difference between EST sequencing and RNA-seq is the depth to which the cDNA is sequenced, being much higher in the latter enabling accurate quantitation of the messenger RNA (mRNA) population and thus a measure of gene expression. This topic will be covered in Chap. 6, but, for now, we will just discuss RNA-seq as a method for gene discovery.

The first conifer gene discovery report using the RNA-seq approach was with the Roche 454 sequencing platform in *Pinus contorta* (Parchman et al. 2010) (Table 3.2). Here, they reported 586,732 sequence reads resulting in 17,000 unique genes (Fig. 3.3). This study established the power of this approach to generate state-of-the-art genomic information in a species where there were very few existing resources. Several more conifer gene discovery projects were reported over the next few years using either the Roche 454 or the Illumina sequencing platforms (Table 3.2). Two obvious trends can be taken from Table 3.2: (1) The number of sequence reads increased dramatically as researchers moved from the Roche 454 platform to the Illumina platform and (2) the number of unique genes discovered does not vary profoundly given the large differences in sequencing depth and tissue types used among studies. Looking at these data, one could easily conclude that conifers, in general and in total, might have about 30,000–40,000 expressed genes. This number is only slightly greater than the average number of genes found in angiosperms, most of which have genome sizes 10–100 times smaller than those of conifers.

The goals of these early RNA-seq studies varied. One goal was to develop a complete catalog of all expressed genes in a species, such as the reports for *Pinus pinaster* (Canales et al. 2014) and *P. lambertiana* (González-Ibeas et al. 2016). In others, it was to discover genes related to a specific function and study their expression, such as taxol synthesis in *Taxus cuspidata* (Wu et al. 2011b), drought stress in *Pseudotsuga menziesii* (Müller et al. 2012), monoterpene synthesis in *Pinus contorta* and *P. banksiana* (Hall et al. 2013), lignin and cellulose biosynthesis in *Cunninghamia lanceolata* (Huang et al. 2012a), and disease resistance in *Pinus monticola* (Liu et al. 2013a). A final goal was to discover genetic variation

Table 3.2 Results obtained from several RNA-seq studies in conifers

Genus	Species	Sequencing platform	Sequence read (no.)	Putative transcript (no.)	References
Cunninghamia	*lanceolata*	Illumina	40 M	83,248	Huang et al. (2012a)
Picea	*abies*	Illumina	70 M	38,419	Chen et al. (2012a)
Pinus	*contorta*	454	586,732	17,000	Parchman et al. (2010)
		454/Illumina	1.3 M	30,000	Hall et al. (2013)
	banksiana	454/Illumina	1.4 M	30,000	Hall et al. (2013)
	tabuliformis	454	911,302	46,584	Niu et al. (2013)
	pinaster	454/Illumina	~300 M	26,020	Canales et al. (2014)
	monticola	Illumina	380 M	39,439	Liu et al. (2013a)
	lambertiana	Illumina/PacBio	2.5 B/1.6 B	33,113	González-Ibeas et al. (2016)
Pseudotsuga	*menziesii*	454	3.6 M	170,859	Müller et al. (2012)
		454	2.76 M	25,002	Howe et al. (2013)
Taxus	*cuspidata*	454	81,148	20,357	Wu et al. (2011a)

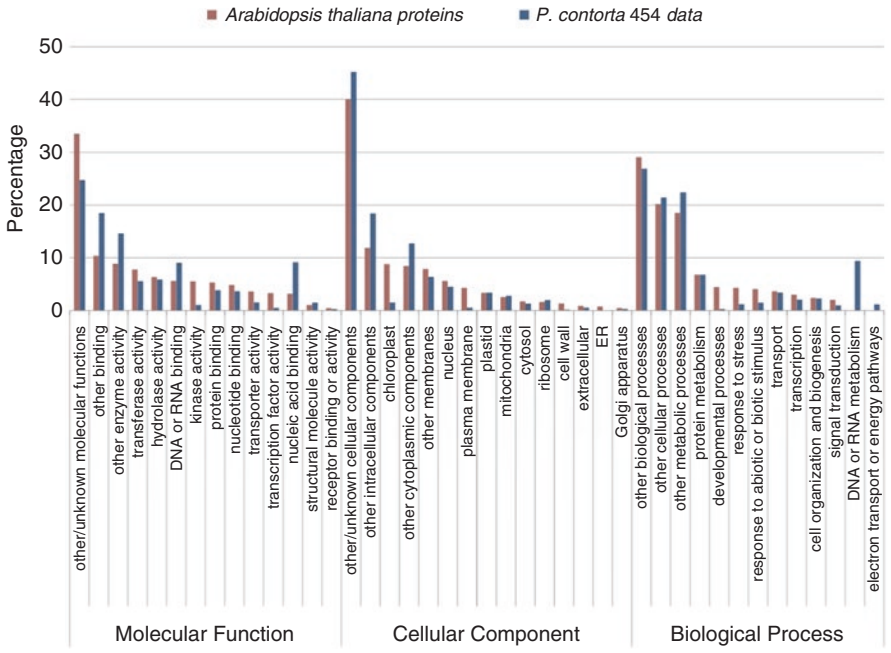


Fig. 3.3 Gene ontology assignments for *Pinus contorta* and *Arabidopsis thaliana*. Proportion of annotated contigs and singletons from *P. contorta* 454 ESTs and annotated *A. thaliana* proteins that matched various gene ontology (GO) categories. (From Parchman et al. (2010))

(single-nucleotide polymorphisms, SNPs) and use this in marker studies, such as in *Pinus contorta* (Parchman et al. 2010) and *Pseudotsuga menziesii* (Howe et al. 2013) or in evolutionary studies, such as in *Picea abies* (Chen et al. 2012a) and *Pinus tabuliformis* (Niu et al. 2013).

Clearly, NGS and RNA-seq studies opened the door to development of important genomic resources in non-model species, including conifers. Table 3.2 lists only 10 of the 600+ conifer species, but we should expect this number will grow exponentially over the next several years.

Conifer Reference Genome Sequences

The discipline of the study of genetics of eukaryotic organisms changed profoundly at the turn of the twentieth century with the sequencing of the human genome. Now, instead of research being focused on one or a small number of genes in individual studies, entire genomes were the focal point of studies. This turning point coincides with the usage of the word *genomics* versus *genetics.* The change from *genetics* to *genomics* could not only be seen in individual research studies but also in things such as the naming of scientific journals, biotech companies, and university departments; a testimony to how profound this change was.

The first tree genome to be sequenced, and only the third plant genome to be sequenced, was *Populus trichocarpa* (Tuskan et al. 2006). This poplar species has a relatively small genome (485 Mb), and accompanied with its potential for a feedstock species for cellulosic ethanol production, it was strategic and technically possible to sequence using largely Sanger sequencing technology. Following soon thereafter was the sequencing of a few other tree species, all of which were angiosperms, with genome sizes similar to, or only slightly larger, than poplar. Most were horticultural species of great economic value (see Neale et al. 2017a for a review of all tree genomes sequenced to date). The genome sequencing of any conifer, the smallest being nearly 10 Gb (Chap. 2), was simply not economically possible; further, it was thought to be technically too difficult to assemble the sequence because of the high repetitive DNA content (however, see discussion of BAC sequencing in Chap. 2). With the arrival, around 2010, of NGS for genome sequencing of eukaryotes, the possibility of sequencing a 20 Gb conifer genome suddenly became feasible.

Nearly simultaneously around 2010, three conifer genome sequencing projects were funded. The first was for *Picea abies* in Sweden funded by the Wallenberg Foundation, the second was for *Picea glauca* in Canada by Genome Canada and regional funding agencies, and the third was in the United States for *Pinus taeda*, *P. lambertiana*, and *Pseudotsuga menziesii* funded by the US Department of Agriculture. These projects were all funded in excess of ten million USD each. Because each of these conifer genomes would by far be the largest genomes of any kind ever sequenced, these projects were mandated to develop new sequencing strategies or, at the very least, modify existing strategies that would produce a successful outcome at a reasonable price. A genome sequencing project for any organism generally has three major components—*sequence, assemble*, and *annotate.* In the next section, we will briefly compare and contrast the approaches used by these projects for each of these three tasks.

Sequencing, Assembly, and Annotation Strategies

A detailed discussion of genome sequencing and assembly strategies as they have been applied to sequencing eukaryotic genomes is beyond the scope of this volume; however, we will attempt to briefly describe the approaches used to sequence conifer genomes to date. Two general strategies have been used to sequence eukaryotic genomes, commonly referred to as (1) BAC by BAC or hierarchical shotgun sequencing and (2) Whole Genome Shotgun (WGS). These approaches and their differences were well chronicled during the competition for sequencing of the human genome where the publicly funded project used the former and the private effort of Celera Corporation used the latter. All conifer genomes have been sequenced using WGS; the BAC-by-BAC approach would have been simply cost prohibitive.

One of the very first considerations in determining a genome sequencing strategy is the source of DNA that will be sequenced. Sequence assembly becomes much more difficult and the results more error prone if there is any sequence polymorphism

between two sequence reads from the same homologous region. This can occur if the DNA used is from either more than one individual or one individual that is heterozygous (allelic differences). The variation in the former case is easily overcome by using DNA from only a single individual, and nearly all genome sequencing projects are done this way. However, heterozygosity of a single individual is more difficult to overcome. For many self-pollinating agricultural crops, a single highly inbred individual can be used which has greatly reduced heterozygosity. Highly inbred, and certainly self-pollinated, plant materials in conifers are difficult to produce due to the significant amount of inbreeding depression that would result (Chap. 1). Furthermore, such materials were not available for any of the three original conifer genome sequencing projects. But there is a tissue even better than the inbred diploid tissue, the haploid conifer seed megagametophyte! Depending on the size of the seed, the amount of DNA that could be isolated from a single megagametophyte is not abundant (1–5 μg) but enough for several NGS technologies. Megagametophyte DNA was used in the US conifer genome projects and the Swedish spruce project. Diploid needle tissue DNA was used in the Canadian spruce project.

The Swedish, Canadian, and US conifer genome sequencing projects all used a form of WGS sequencing using NGS sequencing platforms. We refer the reader to the original publications (Table 3.3) on these genome sequences for details and will only describe the methods briefly here. As an example, an overall depiction of the *Pinus taeda* sequencing strategy is shown in Fig. 3.4. All projects shared a similar approach for where to begin. WGS is applied to short cloned DNA fragments (300–500 bp) which were sequenced redundantly to a fairly great depth (up to 60X). These sequences were assembled into contigs (a few hundred to a few thousand bp each). Next, much longer cloned pieces of DNA (1–40 kbp) were sequenced, either completely or partially, and these sequences were used to merge the much smaller contigs into longer contiguous sequences called scaffolds. The algorithms and software to assemble sequences differed among these projects, but all used some form of the two basic assembly algorithms—Overlap Layout Consensus (OLC) and De Bruijn graph. The genome assemblies for these first conifer genomes were very fractured. In fact, what was reported was the complete, or near complete, sequencing of the genome, but the genome sequence was in several million pieces, essentially none of which were "ordered." For this reason, the first conifer genome sequences, generally reported as version 1.0 (V1.0), were called *draft* sequences, as it would take several more iterations of sequencing and assembly to produce a contiguous length of sequence from one end of a chromosome to the opposite end and to assign these sequences to individual chromosomes identified by karyotypes (Chap. 2) and genetic maps (Chap. 11).

The final step after sequencing and assembly is annotation of the genome sequence. Initially, one of several annotation software applications is used in an automated process to identify the coding from the noncoding portions of the genome. As we saw in Chap. 2, only a very small portion of the genome codes for genes, the rest is made up of highly repetitive sequences of some kind (see sections below and Chaps. 4 and 5). Then by comparison to databases of repetitive sequences and gene sequences, the identity and probable function of the genome sequences

Table 3.3 Summary statistics of sequenced conifer genomes

Species	ID	Genome size (Gb)	Total assembled (Gb)	N50 contig (kbp)	N50 scaffold (kbp)	Number scaffolds (M)	Assembler	References
Picea abies	Z4006	19.6	12	0.6	0.72	10.2	CLC	Nystedt et al. (2013)
Picea glauca x (*P. sitchensis* x *P. engelmannii*)	PG29	15.8	20.8	5.4	22.9	7.1	ABySS	Birol et al. (2013)
Picea glauca	WS77111		22.4		20		ABySS	Warren et al. (2015)
Pinus taeda	20–1010	21.6	22.1	25.4	107	1.5	MaSuRCA	Zimin et al. (2017)
Pinus lambertiana	5038	31.6	33.9	4.25	246.6	58.4	MaSuRCA + SOAP	Stevens et al. (2016)
Pseudotsuga menziesii	412–2	16.1	16.6	44.1	340.7	9.2	MaSuRCA + SOAP	Neale et al. (2017b)

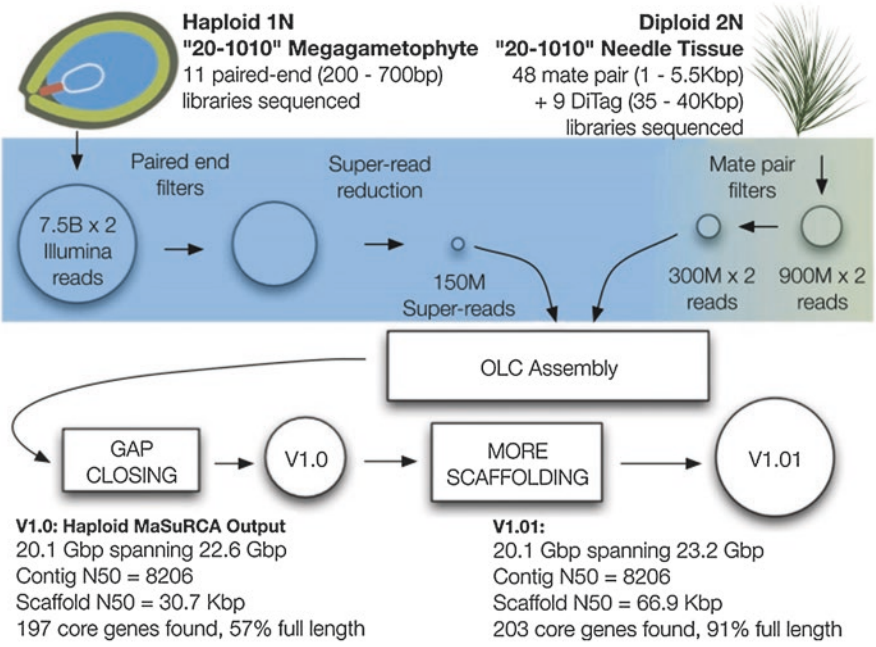


Fig. 3.4 Sequencing and assembly schematic for the *Pinus taeda* genome. (From Neale et al. (2014))

can be inferred. The automated annotation process is just a first approximation of sequence identity and function, subsequent detailed and manual annotation follows by experts in different types of repetitive DNA or coding genes to more robustly establish identity and function.

Summary Statistics of Published Conifer Genome Sequences

As of 2017, genome sequences of six conifer species have been published (Table 3.3). Note that we have indicated in Table 3.3 that the spruce sequence reported in Birol et al. (2013) as *Picea glauca* represents a white spruce hybrid, which is the typical form in which this species occurs in Western Canada.

Standard summary statistics that generally reflect the quality and completeness of these genome sequences are shown in Table 3.3. All of these genomes are extremely large—nearly 20 Gb and more, *Pinus lambertiana* being by far the largest conifer genome and genome of any kind to be sequenced. A comparison of the estimated genome size and the amount of sequence assembled shows that all were essentially completely sequenced apart from *Picea abies* where only about 60% of the genome was actually sequenced. The N50 contig and N50 scaffold sizes are standard measures of sequence contiguity. It can be seen how contiguity has steadily improved with each new published sequence with the scaffold N50 of *Picea abies*

being just 0.72 kbp to that of *Pseudotsuga menziesii* being 340.4 kbp, a nearly 500-fold improvement. Nevertheless, all of these genome sequences are very highly fractured, containing millions of scaffolds. As of 2017, none of these genome sequences have been anchored to chromosomes or genetic maps. Clearly, much work lies ahead toward improving the quality of these valuable resources. Nevertheless, an enormous amount has been learned about the content of conifer genomes which we will cover briefly in the next two sections and in greater detail in Chaps. 4 and 5.

Discovery of the Noncoding DNA Content of Conifer Genomes

It was well established in the earlier work in reassociation kinetics that conifer genomes must be largely made up of some type of repetitive DNA (Rake et al. 1980; Kriebel 1985; Chapter 2). But without the ability to sequence DNA, the nature of this repetitive DNA was unknown. In Chap. 4, we trace the history of sequencing conifer repetitive DNA, from the early work of sequencing single small pieces of DNA to the most recent sequencing of entire genomes, toward gradually understanding the nature of the repetitive DNA content in conifers. In brief, the sequencing of the first six conifer genomes has revealed that the largest portion of the repetitive DNA is made up of long terminal repeat (LTR) retrotransposons (Fig. 3.5), that these elements increased in number in ancestral genomes many millions of years ago, and that they have persisted in conifer genomes, possibly due to the lack of mechanisms in conifers to purge these elements from the genome.

Discovery of the Number and Types of Coding Genes in Conifers

Gene discovery approaches based on RNA as the starting material (either EST sequencing or RNA-seq) provide a first approximation of the exact number of protein-coding genes in a genome, but for several reasons, these may be under- or overestimates. The reason they might be underestimated is because the gene must be expressed in the tissue sample from which the RNA was isolated or else it will not be detected. EST or RNA-seq projects that exhaustively sample different tissues under many different developmental states or inductive conditions (biotic or abiotic stress treatments) might come close to detecting every single coding gene, but it is well known from many eukaryotic systems that some genes are expressed at very low levels, at very limited time points, or both, so these might easily escape detection. What might just as easily happen is that the true number of genes is overestimated. This is basically a function of the bioinformatic analyses to determine whether expressed gene copies are from the same or different gene. The common reason two expressed gene copies derived from the same gene might be assigned to different genes are allelic differences between copies and alternative splicing (see Chap. 5). Therefore, full genome sequencing and assembly ultimately provides the best estimate of the exact number of coding genes.

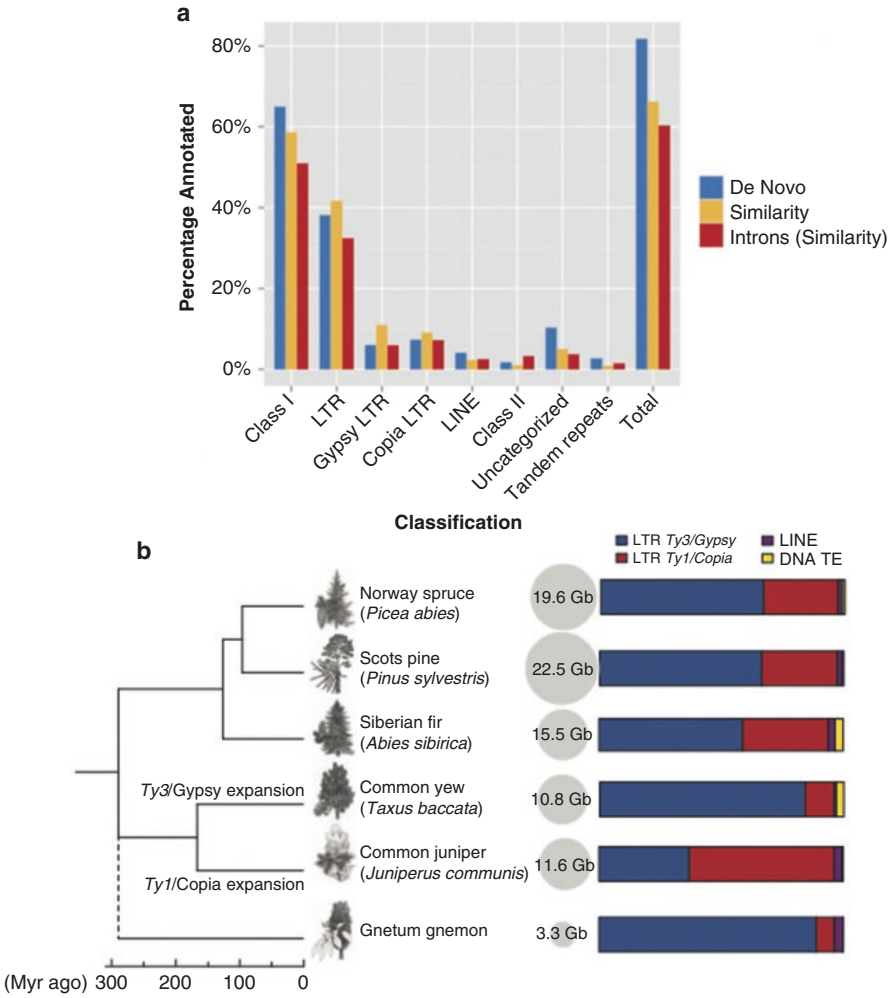


Fig. 3.5 Interspersed and tandem repetitive content of the *Pinus taeda* genome. (**a**) Overview of repetitive content in the *P. taeda* genome for similarity (blue) and de novo (yellow) approaches. (**b**) Conifer genomes contain expansions of a diverse set of LTR-RTs. Distribution of different classes of transposable elements from six gymnosperm species. (From Neale et al. (2014))

The estimated gene number derived from full genome sequencing projects in conifers varies enormously (Table 3.4). The estimates shown in Table 3.4 range from a minimum number based on high-quality gene models to an upper limit just based on the number of unique genes derived from automated annotation. As with all genome sequencing projects, the true number will become known in time, but it is expected to be in the range of 30,000–50,000 genes, which is very similar or slightly higher than that for angiosperms. Furthermore, the number of genes among conifer species is likely to be very similar—the difference in genome sizes among species being due almost entirely to the repetitive DNA content (Chap. 5).

Table 3.4 Comparison of gene number and gene structure in conifers

Species	Gene number	Average length (bp)	Average intron length (bp)
Picea abies	26,359–58,587	931	1020
Picea glauca	16,386–105,000	1421	603
Pinus lambertiana	13,936–71,117	1330	8039
Pinus taeda	9,024–50,172	1562	12,875
Pseudotsuga menziesii	27,257–54,830	1169	2301

It can also be seen that the average size of genes does not vary much among species, the estimate of 931 bp from *Picea abies* likely being an underestimate due to the highly fractured assembly (Table 3.3). However, quite a large difference in average intron size is observed. These estimates are likely affected by differences in the quality of the assemblies, and better estimates will result over time. Nevertheless, the nature of and activity differences in retrotransposons and the size and content of introns will be an active research question in the years ahead.

Aside from the exact number of genes found in conifers versus other plants, another question is whether similar types of genes are found and if there are large numbers of genes unique to conifers. Results from *Pinus taeda* (Fig. 3.6) and the other sequenced conifers suggest that the same classes of genes found in angiosperms are found in conifers, although differences in relative abundances are found. There are also groups of genes unique to conifers. Ultimately, it will become clear what are the fundamental genetic differences between angiosperms and gymnosperms (conifers) and what is the relative gain or loss of genes between these plant groups.

Chloroplast and Mitochondrial Genome Sequencing

The first conifer chloroplast genome to be completely sequenced was for *Pinus thunbergii* (Wakasugi et al. 1994). The chloroplast genome was only 119,707 bp in size, much smaller than angiosperm cpDNA genomes, in large part due to the missing copy of the large inverted repeat (Chap. 2). However, it was shown that a small segment of the second inverted repeat was present, thus providing support for the hypothesis that the inverted repeat copy was lost in conifers following divergence of the angiosperms. The genome sequence was annotated and showed that it contained 4 rRNA genes, 32 tRNA genes, 61 protein-coding genes, and 11 open reading frames (ORFs) (Fig. 3.7). This included nearly all genes found in angiosperm cpDNA except, interestingly, the loss of 11 *ndh* genes. The authors concluded that

Fig. 3.6 (continued) (*Arabidopsis thaliana, Glycine max, Populus trichocarpa, Ricinus communis, Theobroma cacao*, and *Vitis vinifera*), early land plants (*Selaginella moellendorffii* and *Physcomitrella patens*), and a basal angiosperm (*Amborella trichopoda*). (**b**) Gene ontology molecular function term assignments by family for all species (red), conifers (green), and *Pinus taeda* exclusively (blue). (From Neale et al. (2014))

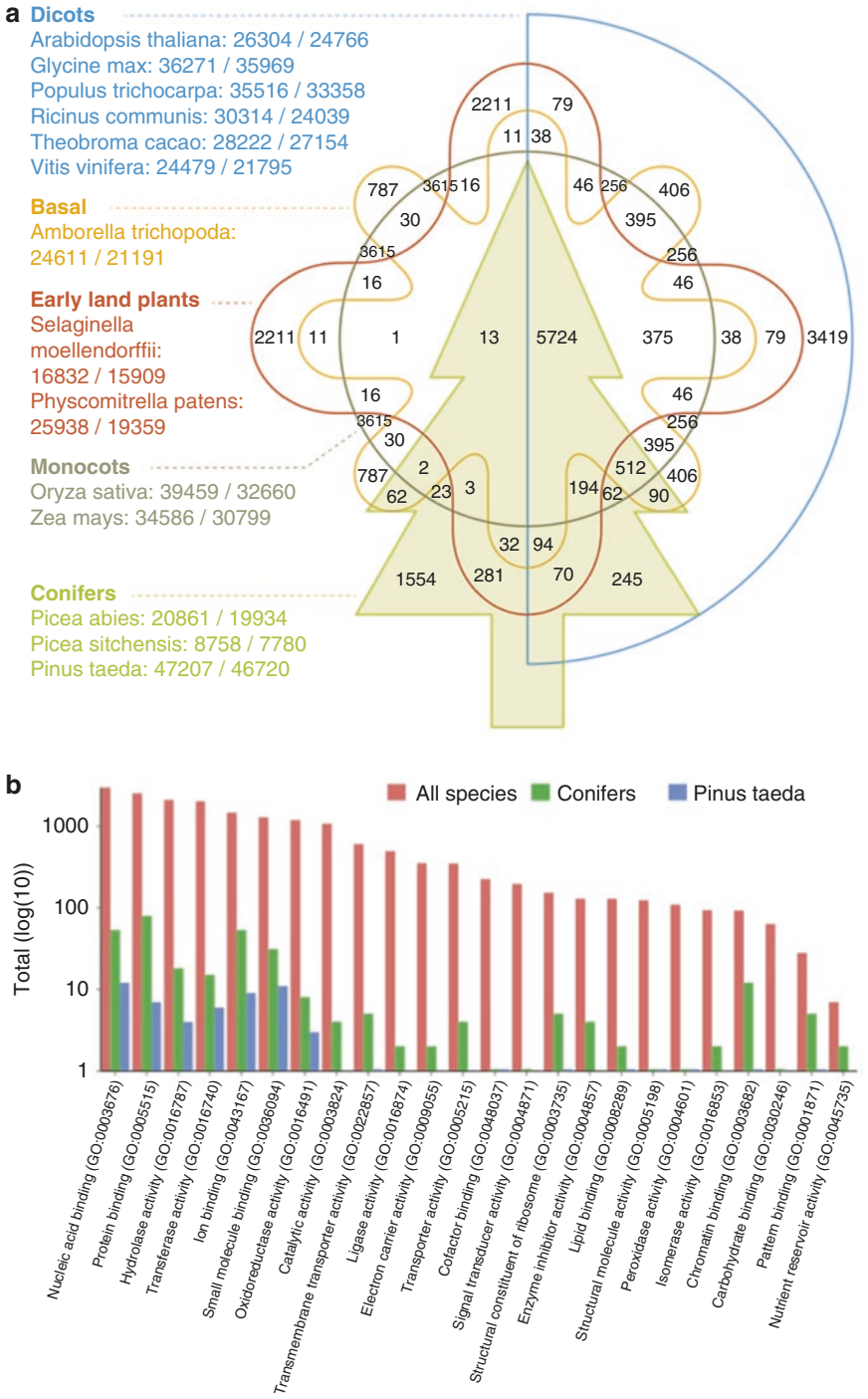


Fig. 3.6 Unique gene families and gene ontology term assignments in conifers. (**a**) Identification of orthologous groups of genes for 14 species split into five categories: conifers (*Picea abies, P. sitchensis*, and *Pinus taeda*), monocots (*Oryza sativa* and *Zea mays*), dicots

these genes may have been transferred to the nDNA genome or were not functionally necessary. To date, the *P. thunbergii* nDNA genome has not been sequenced to determine which hypothesis might be correct.

The second conifer chloroplast genome to be sequenced was that for *Cryptomeria japonica*, a member of the Cupressaceae (Hirao et al. 2008). It was found to be slightly larger, 131,810 bp, than the *Pinus thunbergii* cpDNA genome, and all the *ndh* genes were present. This observation then inspired additional research to better understand conifer chloroplast genome evolution and determine if differences tracked established evolutionary relationships.

In the subsequent 20 years following the initial sequencing of conifer chloroplast genomes, dozens of cp genomes have been sequenced from all families of conifers (Cronn et al. 2008; Wu et al. 2011a; Yi et al. 2013; Guo et al. 2014; Hsu et al. 2014; Vieira et al. 2014; Wu and Chaw 2014; Zhang et al. 2014a; Jackman et al. 2016; Wu and Chaw 2016). All of these chloroplast genomes were sequenced with NGS technologies. A brief summary of all this work is that all conifer chloroplast genomes are similar, all missing one copy of the inverted repeat, but that small differences in the presence/absence of some genes and rearrangements of the genomes follow phylogenetic lines.

Sequencing of mitochondrial genomes in conifers has lagged far behind. For the three conifer species where the mitochondrial genome has been sequenced, it was essentially a by-product of the sequencing of the nuclear genome: *Picea abies* (Nystedt et al. 2013), *Picea glauca* (Jackman et al. 2016), and *Pinus taeda* (Neale et al. 2014). The *Picea abies* mitochondrial genome was nearly four times longer than that of the *Pinus taeda* assembly (4 Mb versus 1.2 Mb). The *Picea glauca* mitochondrial genome (5.6 Mb) was the only one annotated yielding 8 rRNA genes, 29 tRNA genes, 106 protein-encoding genes, and, very interestingly, 6265 ORFs (Jackman et al. 2016). None of these ORFs could be annotated by comparisons to the closest sequenced and annotated gymnosperm mitochondrial genome, raising the question as to whether conifer mitochondrial genomes may have many novel genes. This will certainly be an active research topic as more and more conifer genomes (nuclear, chloroplast, and mitochondrial) are sequenced.

Summary

The *modern era* of genome research in conifers, as would be true for the study of all forms of life, was enabled by first the development and then the application of high-throughput capacity to sequence DNA. The first approach for sequencing large amounts of DNA was through the expressed portion of the genome, which was called expressed sequence tag sequencing. In this approach, mRNA was first isolated from tree tissues, converted to cDNA, and then cloned. The cDNAs could then be sequenced in a high-throughput manner and their identity inferred by comparisons to DNA sequence databases. This approach was, however, quite costly in the Sanger sequencing era and was only applied to a small number of conifers of great economic value. Nevertheless, this approach provided a view of the number of

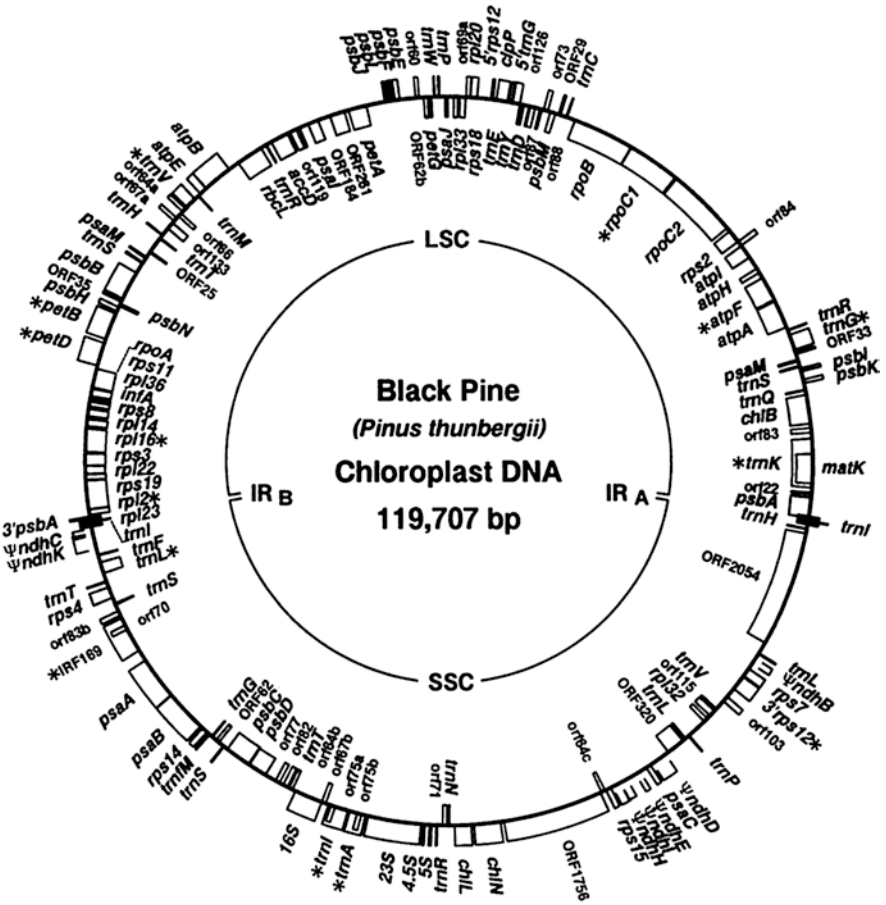


Fig. 3.7 Gene map of the *Pinus thunbergii* chloroplast genome. Genes shown on the inside of the circle are transcribed clockwise, and genes on the outside are transcribed counterclockwise. (From Wakasugi et al. (1994). Copyright (1994) National Academy of Sciences, USA)

expressed genes that might be in a conifer and what proportion might be similar or unique compared to those discovered in other plants. Generally, the number and similarity proved not to be profoundly different from that found in the model plant *Arabidopsis thaliana* or the model tree *Populus trichocarpa.* The next generation of expressed gene sequencing, called RNA-seq, enabled this approach to be applied to many more species due to its much lower cost.

The other important development in the modern era was the ability to sequence entire conifer genomes. Full eukaryotic genome sequencing, beginning with the human genome, began in the 1990s, but due to the very large size of conifer genomes, the complete sequencing of a conifer genome did not occur until 2013. This was finally made possible by the reduced cost of NGS and more advanced bioinformatics methods to assemble genome sequences. As of 2018, only five

conifer genome sequences have been reported (*Picea abies*, *Picea glauca*, *Pinus taeda*, *Pinus lambertiana*, and *Pseudotsuga menziesii*) due to the initial technical and financial challenges; however, efficient strategies and protocols have now been developed such that tens, if not hundreds, of conifer genomes will be sequenced in the near term. As has been true for all plant and animal species with a reference genome sequence, this resource is transformative and has accelerated conifer research in many ways.

Noncoding and Repetitive DNA

4

Introduction

It is well established that conifer genomes are comprised largely of noncoding and repetitive DNA (Chaps. 2 and 3). However, the much smaller fraction of the genome comprised of protein-coding genes has been the focus of most research (see Chaps. 5, 6, and 7) because these genes ultimately lead to phenotypes and the genetic differences among species and individuals within species. Nevertheless, some attention has been paid to characterizing the nature of noncoding and repetitive DNA and what functional significance, if any, this DNA might have. In this chapter, we will introduce the types of noncoding and repetitive DNA that have been discovered in conifers and what is known about how and when in evolutionary time this DNA was amplified and became so abundant. We will also point to how variability in this DNA has been used in population genetic studies (Chap. 9) and evolutionary and phylogenetic studies (Chaps. 15 and 16). There are four general classes of noncoding and repetitive DNA that will be discussed: (1) ribosomal DNA, (2) tandemly duplicated DNA, (3) transposons and retrotransposons, and (4) pseudogenes.

Ribosomal DNA

The genes that code for ribosomal RNA (referred to in the literature as either rDNA or rRNA genes) exist in many thousands of copies in plants. There are separate genes for the two basic subunits: (1) the 18S-5.8S-26S rRNA gene and (2) the 5S rRNA gene. The rRNA genes make up a small proportion (<5%) of the total repetitive DNA content in conifer genomes. We saw in Chap. 2 that the location of these genes on chromosomes can be identified using fluorescence in situ hybridization (FISH) technique. The regions where the 18S-5.8S-26S genes are found on chromosomes are called nucleolus organizing regions (NORs) which are found on several,

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_4

but not all, chromosomes depending on the species. The 5S rRNA gene has only been found at one location (Karvonen et al. 1993; Brown and Carlson 1997; Amarasinghe and Carlson 1998).

An example of the structure of the 18S-5.8S-26S rRNA gene in *Pinus sylvestris* is shown in Fig. 4.1. The regions in between the 18S, 5.8S, and 26S regions are called internal transcribed spacers (ITS). The regions in between each gene are the intergenic spacer (IGS). There were two major topics of interest in conifer rRNA genes that emerged over a fairly brief period between 1988 and 1998: (1) copy number variation in 18S-5.8S-26S rRNA genes and (2) restriction site variation in the ITS and IGS.

The first published study on rRNA gene copy number was done in *Pseudotsuga menziesii* (Strauss and Tsai 1988). They reported relative copy number among 54 trees across latitudinal, longitudinal, and elevational gradients. A significant amount of variation was found among trees, and increasing copy number was correlated with increasing latitude, longitude, and elevation. These results were consistent with earlier reports of genome size and these geographic variables (Chap. 2). In a study in *Picea rubens* and *P. mariana*, similar correlations with copy number and these geographic variables were also found (Bobola et al. 1992). In another study in *Pinus rigida*, very large differences (up to 12-fold) were found among populations in the Pine Barrens of New Jersey, USA, and individuals in more stressful environments had higher copy number (Govindaraju and Cullis 1992). These early studies suggested that rRNA gene copy number in conifers may be adaptive. Interest in this topic thereafter declined, and no new studies were published after the early 1990s.

The other topic of interest was searching for genetic variation in rRNA genes using restriction enzymes, particularly in the ITS and IGS regions. Studies in other plant systems established that genetic variation could be found, and it may be adaptive. Studies in *Pinus sylvestris* (Karvonen et al. 1993, Karvonen and Savolainen 1993) reported that restriction site variation could be found among trees in the ITS and IGS regions, although no link to adaptive patterns of variation was established. Later, studies such as one in the *Pinus* subsection *Cembroides* showed that variation in the ITS region was informative in phylogenetic studies (Gernandt et al. 2001).

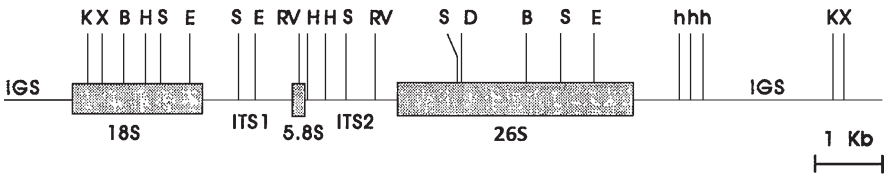


Fig. 4.1 Schematic restriction map of the *Pinus sylvestris* rDNA repeat showing the site of a 0.4-kb deletion in the ITS1 region. (From Karvonen et al. 1993)

Tandem Repeats: Satellite, Minisatellite, and Microsatellite DNA

Tandemly repeated DNA sequences, first discovered in human and other animal and plant genomes, are also found in conifers. These classes of tandemly repeated DNA are distinguished by the size of the repeat unit. Satellite DNAs are the largest (>100 bp). Minisatellites are repetitive sequences of 10–60 bp that can vary in the repeat number at individual loci. This type of variation is called *variable number tandem repeats* (VNTRs) and has been used as a genetic marker to fingerprint individuals in humans and other plants and animals. In two early studies, both using the bacteriophage M13 DNA probe, variation in minisatellite sequences was found in *Pinus torreyana* (Rogstad et al. 1988) and in *Picea abies* (Kvarnheden and Engström 1992), but there is no record that minisatellite markers were ever used in standard genetic applications in conifers.

Microsatellites are short repeat sequences of 1–5 bp and are ubiquitous in plants and animals. Unlike minisatellites, microsatellites have been used extensively in all kinds of applications in conifers. Microsatellite sequences provide an excellent type of genetic marker because the di-, tri-, and tetra-nucleotide repeat sequences can be found in multiple copy numbers at individual loci, thus providing a rich source of allelic variation for genetic studies. Microsatellite genetic markers in plants are often called *simple sequence repeats* (SSRs). Because SSRs are generally found in noncoding DNA regions, they are applied as neutral genetic markers (Chap. 9).

The first report of developing SSRs in conifers was in *Pinus radiata* (Smith and Devey 1994). They reported that the $(CA)_n$ repeat was found, on average, once in every 750 bp. Note, in this chapter, we will only discuss the type of microsatellite sequences and their abundance in genomes; in Chap. 9, we will discuss the application of SSR markers. Since 1994, there have been many reports of the successful development of SSR genetic markers in conifers, a few of which report the frequency of different repeat types. For example, Echt et al. (1996) reported an average density of 1/6250 bp of four different di- and tri-nucleotide repeats in *Pinus strobus* and *P. taeda*. Other studies reported much higher densities, for example, 1/194–1/409 bp in *Picea abies* (Pfeiffer et al. 1997). It would not be until large pieces of DNA or entire genomes were sequenced in conifers that it became clear just how abundant microsatellites are in conifers and what proportion of the total repetitive DNA is microsatellite DNA.

The most comprehensive study to estimate the abundance of tandemly repeated DNA in conifer genomes, prior to complete genome sequencing, was in *Pinus taeda* (Wegrzyn et al. 2013). In this study, the presence of satellite, minisatellite, and microsatellite sequences was searched for in a sample of 103 sequenced BAC clones and 90,954 sequenced fosmid scaffolds which represented approximately 1% of the *P. taeda* genome. In total, these repeated DNA sequences make up only 2.6% of the region sequenced: 0.96% satellite, 1.56% minisatellite, and 0.09% microsatellite (Table 4.1). In a similar study in *Taxus mairei* (Hao et al. 2011), the distribution of different microsatellite repeat-size classes is shown in three conifers relative to what has been observed in angiosperms (Fig. 4.2). Interestingly, the distributions in *Picea*

Table 4.1 Summary of tandem repeats from BAC and fosmid sequences in *Pinus taeda*

Type of repeat	Total loci	Copy number	Variant (no.)	Total length in bp (% of sequence sets)
Microsatellites				
Dinucleotide	2967	64,740	10	126,254 (0.046)
Trinucleotide	645	9,657.70	39	28,440 (0.010)
Tetranucleotide	282	3,899.30	46	15,316 (0.006)
Pentanucleotide	172	3,167.20	75	15,560 (0.006)
Hexanucleotide	402	3,427.80	172	20,303 (0.07)
Heptanucleotide	499	4174	153	28,765 (0.010)
Octanucleotide	176	920.1	135	7184 (0.003)
Total	**5143**	**89,986.10**	**630**	**241,822 (0.09)**
Minisatellites				
9–30 bp	31,363	84,572.90	26,428	1,631,083 (0.588)
31–50 bp	11,800	30,641.90	10,989	1,164,786 (0.420)
51–70 bp	5316	13,775.30	5192	805,210 (0.290)
71–100 bp	3518	8642.90	3473	722,282 (0.260)
Total	**51,997**	**137,633**	**46,082**	**4,323,361 (1.559)**
Satellites				
101–200 bp	5183	12,589.30	5110	1,710,062 (0.617)
201–300 bp	857	2141.20	854	524,329 (0.189)
301–400 bp	280	691.6	280	236,623 (0.085)
>400 bp	162	405.9	162	179,726 (0.065)
Total	**6482**	**15,828**	**6406**	**2,650,740 (0.956)**
Grand total	**63,622**	**243,447.10**	**53,118**	**7,215,923 (2.602)**

From Wegrzyn et al. (2013)

glauca and *Pinus taeda*, members of the Pinaceae, look very similar to each other, but quite different from that in *Taxus mairei*, a member of the Taxaceae. These early estimates of the proportion of conifer genomes comprised of tandemly repeated sequences based on sequencing of a small number of BAC clones, fosmid clones, or both have largely been validated by full genome sequencing. Wegrzyn et al. (2014) reported that the estimated proportion of tandem repeats in *Pinus taeda*, *Picea glauca*, and *Picea abies* were 2.86, 2.71, and 2.40%, respectively. Thus, as important as these sequences have become as genetic markers, they account for only a very small portion of the total genome space (Fig. 4.3). This raises the question of how useful these markers would be in complex trait dissection research and marker breeding (see Chap. 11).

Transposons and Retrotransposons

The first time it was suggested, based on experimental data, that the large size of conifer genomes might be due to transposons and retrotransposons was in a meeting abstract in 1990 (Kossack et al. 1990). It was shown that a *Ty3-gypsy*-like retrotransposon was in very high copy number in *Pinus radiata*. This report would not be

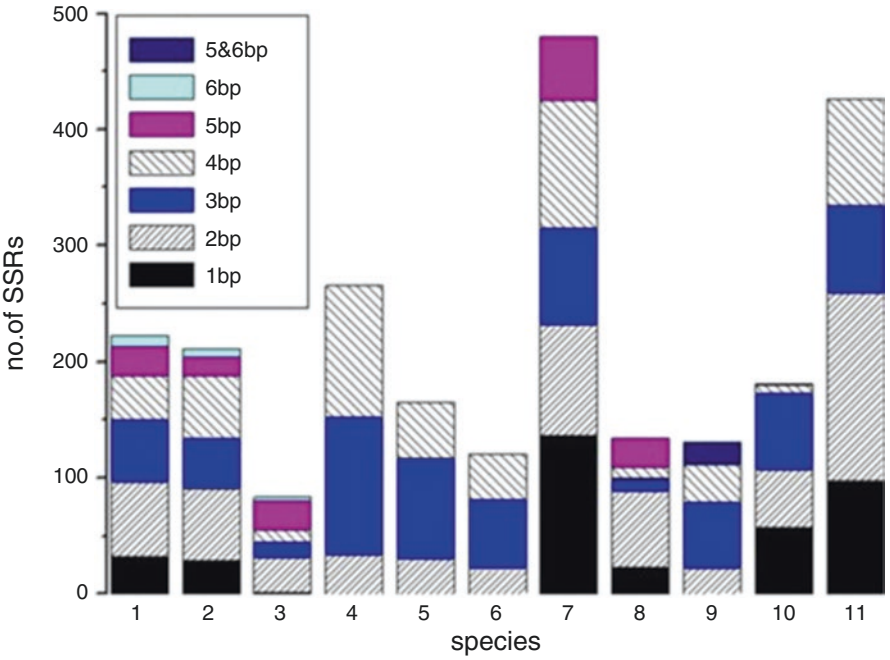


Fig. 4.2 Distribution and frequency of SSRs detected in the nuclear genome of various plant species. The SSR numbers of all species were those within 1 Mb of genomic sequences. 1, *Picea glauca*; 2, *Pinus taeda*; 3, *Taxus mairei*; 4, *Brachypodium distachyon*; 5, wheat; 6, maize; 7, *Vitis vinifera*; 8, apple; 9, carrot; 10, *Brassica*; 11, papaya. (From Hao et al. 2011)

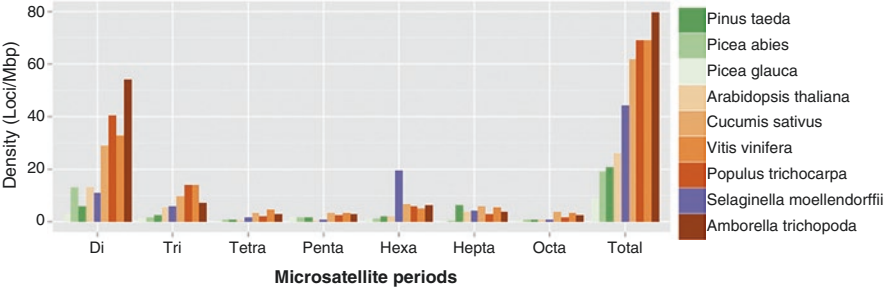


Fig. 4.3 Microsatellite density for three conifer genomes (green), one clubmoss genome (purple), and five angiosperm genomes (orange), (loci per megabase). (From Wegrzyn et al. 2014)

published until 1999 (Kossack and Kinlaw 1999). This early report, and those that followed prior to sequencing of BAC clones or entire genomes, shared common experimental approaches. To begin, a fragment of presumed very high-copy DNA would be cloned from conifer DNA either by cutting out a high-copy band from a gel of restriction-enzyme-digested genomic DNA or by PCR cloning using primers from a transposon or retrotransposon from an angiosperm. This cloned conifer DNA

fragment would then be sequenced and its identity determined by database queries. Additionally, the cloned DNA fragment might be used as a Southern hybridization probe to conifer DNA to estimate abundance and copy number or used as a FISH probe to determine distribution on chromosomes.

The first published report of a high-copy retrotransposon in a conifer was in *Pinus elliottii* (Kamm et al. 1996). A high-copy DNA fragment was cloned and sequenced, and its identity was similar to the *Ty1-copia* retrotransposon found in other plants. This cloned DNA fragment was used as a Southern hybridization probe to other conifers to infer that it existed in very high-copy numbers in other conifer taxa (Fig. 4.4). Furthermore, it was used as a FISH probe to show that it was widely distributed across all chromosomes. In a subsequent study by this group, it was

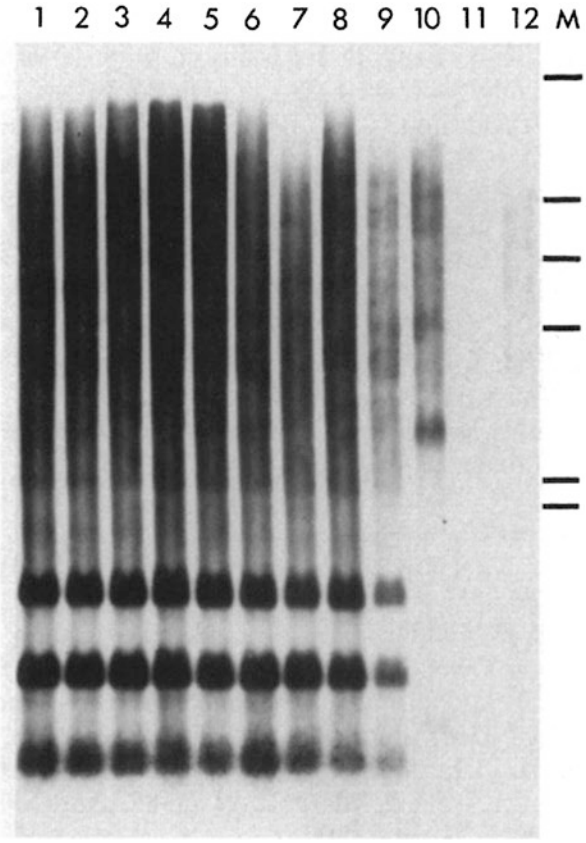


Fig. 4.4 Distribution of *TPE*1 in several species of *Pinus* and other gymnosperms. Southern blot of *Dra* I-digested genomic DNA of *P. echinata* (lane 1), *P. elliottii* var. *elliottii* (lane 2), *P. palustris* (lane 3), *P. caribaea* (lane 4), *P. oocarpa* (lane 5), *P. banksiana* (lane 6), *P. massoniana* (lane 7), *P. resinosa* (lane 8), *P. strobus* (lane 9), *Picea abies* and *P. glauca*, mixed (lane 10), *Taxodium distichum* (lane 11), and *Ginkgo biloba* (lane 12) was hybridized with *TPE*1. (From Kamm et al. 1996. Copyright (1996) National Academy of Sciences, USA)

shown that this *Ty1-copia*-like element is found widely in the plant kingdom and is dispersed throughout the genome (Brandes et al. 1997). The high-copy DNA fragment first reported by Kossack et al. (1990) was determined to be a different retrotransposon, a *Ty3-gypsy*-like retrotransposon that was known from other plant and animal systems (Kossack and Kinlaw 1999). Southern hybridizations with this cloned DNA showed that it is also found in high-copy number in other conifers. It would later be shown following full genome sequencing that these two retrotransposons types, the *Ty1-copia*-like and the *Ty3-gypsy*-like, make up a very large proportion of noncoding and repetitive DNA in conifers and are largely responsible for the large genome sizes of conifers. The mystery of what these ancient and diverged sequences in conifer genomes proposed by Kriebel (1985) might now has been solved.

A series of publications over the next 20 years or so, all before genome sequencing, firmly established the abundance of the *Ty1-copia*-like and the *Ty3-gypsy*-like retrotransposons in conifers (Elsik and Williams 2000; L'Homme et al. 2000; Friesen et al. 2001; Stuart-Rogers and Flavell 2001; Rocheta et al. 2007; Fan et al. 2013, 2014). For example, Elsik and Williams (2000) showed that the high proportion of apparent single-copy sequences in *Pinus* reported by Kriebel (1985) was in part made up of retrotransposons but grouped with protein-coding genes because the retrotransposons had diverged in sequence over evolutionary time. Other studies in *Pinus* showed that *Ty1-copia*-like and the *Ty3-gypsy*-like retrotransposons were abundant (Rocheta et al. 2007; Fan et al. 2013, 2014). L'Homme et al. (2000) and Stuart-Rogers and Flavell (2001) showed that these retrotransposons were also abundant in *Picea*. Finally, Friesen et al. (2001) showed that these transposons were present and abundant not only in conifers but also across other gymnosperms.

Several other retrotransposons were also discovered in the pre-genome-sequencing era. Miguel et al. (2008) discovered an *envelope*-like retrotransposon in *Pinus pinaster* that may have originated as a retrovirus. However, no experiments were done to estimate the copy number of this repetitive element in *P. pinaster*. Another interesting repetitive element, called *Gymny*, was found in *Pinus taeda* (Morse et al. 2009). It is related to repetitive elements previously discovered in angiosperms, but, interestingly, it appears to be unique to the genus *Pinus*. It was estimated that this *Gymny* retrotransposon might make up 1.26% of the *P. taeda* genome.

The studies described to this point in this section collectively reveal the great importance of transposons and retrotransposons in conifer genomes and their role in producing the very large size of conifer genomes. However, because these studies all employed a somewhat restrictive search approach (PCR cloning with heterologous primers, cloning high-copy DNA fragments from conifer DNA, etc.), it could not be known what the exact transposon and retrotransposon content of conifer genomes might actually be. This could only become known once very large pieces of conifer DNA, or entire genomes, were sequenced. We briefly presented findings of these studies in Chaps. 2 and 3, but we will describe these studies in greater detail as they relate to transposon and retrotransposon discovery in conifers.

In the first study in conifers reporting complete sequencing of a BAC clone, Hamberger et al. (2009) reported that the *Picea glauca* genome might be made up of approximately 40% transposons and retrotransposons. In a larger BAC sequencing project in *Pinus taeda*, Kovach et al. (2010) estimated that at least 80% of the genome was made up of repetitive elements. Furthermore, they were able to partition the total transposon and retrotransposon content in different classes (Table 4.2). The long terminal repeat (LTR) retrotransposons (*gypsy*-like and *copia*-like retrotransposons) make up the largest category. Liu et al. (2011a) estimated that 90% of the *Taxodium distichum* var. *distichum* genome is made up of repetitive elements based on Cot analysis of a BAC library.

The most comprehensive study of repetitive DNA content in a conifer, prior to complete genome sequences, was done in *Pinus taeda* (Wegrzyn et al. 2013). The repetitive DNA content was characterized from 103 fully sequenced BAC clones and 90,954 fully sequenced fosmid clones. This study used a combination of homology-based and de novo bioinformatic approaches to search for repetitive elements. The discovered repetitive sequences were then annotated using databases of repetitive DNA sequences. From these data, it was estimated that 86% of the genome would be comprised of repetitive elements and furthermore 60% of the total repetitive DNA is LTR retrotransposons. The various types of LTR and non-LTR retrotransposons discovered are shown in Table 4.3. *Gypsy*-like and *copia*-like repetitive elements, as expected, make up the largest portion, but elements such as LINE, SINE, TIR, Helitron, and other elements known from angiosperms were also found. In addition, several new *gypsy*-like and *copia*-like repetitive elements were also discovered and named (Table 4.4). In summary, by 2013, we were beginning to understand the very large transposon and retrotransposon content of at least a one *Pinus* and one *Picea* genome.

The next phase in the understanding of the repetitive DNA content in conifers came with the genome sequencing of the first few conifer species (Chap. 3). In addition to sequencing the *Picea abies* genome, Nystedt et al. (2013) also did partial sequencing in four other conifers and one other gymnosperm. The largest fraction of transposable elements was *Ty3-gypsy*-like retrotransposons, followed by *Ty1-copia*-like retrotransposons, and then LINEs and DNA transposable elements (Fig. 4.5). They also performed a phylogenetic analysis with these transposable elements to infer the timing of expansion of these elements in different clades. An important conclusion that emerged from these analyses is that conifers (and maybe all gymnosperms) might lack the functional ability to purge these accumulating sequences from their genomes, and this might explain how these genomes have gotten to be so large.

Table 4.2 Transposon and retrotransposon sequences in the genome of *Pinus taeda*

Type of noncoding DNA	Proportion of genome (%)
DNA transposons	32
Endogenous retroviruses	4
Non-LTR retroelement	22
LTR retroelement	42

From Kovach et al. (2010)

Table 4.3 Summary of the proportion of categories of repetitive DNA content of the *Pinus taeda* genome as represented by a set of sequenced BAC clones, a set of sequenced fosmid clones, and both. "Similarity" and "De novo" refer to the bioinformatic methods used to determine identity

	BAC		Fosmid		Combined	
	Similarity	De novo	Similarity	De novo	Similarity	De novo
Category	(%)	(%)	(%)	(%)	(%)	(%)
Class 1	2.9376	33.682	1.3394	21.295	1.3879	21.8222
LTR retrotransposon	2.9376	25.1368	1.3388	14.8972	1.3873	15.3333
Gypsy	2.0772	2.4524	0.918	1.1181	0.9477	1.175
IFG7	0.5793	1.1027	0.2189	0.3109	0.2184	0.3447
Gymny	0.1541	0.3491	0.0897	0.0942	0.0924	0.1051
Corky	0.9716	0.0543	0.3894	0.1743	0.4143	0.1692
PGGYPSYX1	0.3722	0	0.22	0	0.2226	0
Other	0	0.9462	0	0.5387	0	0.556
Copia	0.8604	1.6113	0.4208	0.5343	0.4396	0.5803
TPE1	0.8422	0.8066	0.4072	0.3699	0.4258	0.3885
TY1_PE	0	0.5255	0	0	0	0.0225
Copia4-PTR	0	0	0.0073	0.0096	0.007	0.0092
Copia ES	0	0	0.0002	0	0.0002	0
RT_GB	0.002	0	0.0003	0	0.0003	0
RT_PT	0.0162	0	0.0058	0	0.0063	0
Other	0	0.2792	0	0.1549	0	0.1602
DIRS	0	0.2988	0	0.2213	0	0.2246
PLE	0	0.1665	0	0.1266	0	0.1283
LINE	0	1.4017	0.0003	0.6752	0.0003	0.7062
PILN1_PT	0	0	0.0003	0	0.0003	0
SINE	0	0.0162	0	0.0025	0	0.003
Class II	0	1.2173	0	0.5023	0	0.5328
TIR	0	1.0497	0	0.3276	0	0.3584

(continued)

Table 4.3 (continued)

	BAC		Fosmid		Combined	
	Similarity	De novo	Similarity	De novo	Similarity	De novo
Category	(%)	(%)	(%)	(%)	(%)	(%)
Helitron	0	0.0221	0	0.038	0	0.0373
Annotated	100	2.8383	100	0.9588	1	1.0391
Total interspersed content	2.9376	38.8427	1.3394	25.4083	1.3879	25.9827
Tandem repeats	0.9292		0.5224		0.5398	
Total repetitive content	**3.8668**	**39.7719**	**1.8618**	**25.9307**	**1.9277**	**26.5225**

From Wegrzyn et al. (2013)

Table 4.4 High-coverage LTR families in *Pinus taeda* identified with the de novo methodology

Repeat family	Full-length copy (no.)	Length (bp)	Percent of sequence set
TPE1	159	1,077,598	0.39
PtPiedmont (93122)	133	969,109	0.35
IFG7	162	956,018	0.34
PtOuachita (B4244)	47	576,871	0.21
Corky	78	469,286	0.17
PtCumberland (B4704)	67	431,492	0.16
PtBastrop (82005)	38	378,631	0.14
PtOzark (100900)	32	378,020	0.14
PtAppalachian (212735)	67	367,653	0.13
PtPineywoods (B6735)	68	322,632	0.12
PtAngelina (217426)	24	309,248	0.11
Gymny	24	291,479	0.11
PtConagree (B3341)	50	285,850	0.1
PtTalladega (215311)	33	274,826	0.1
Total	**982**	**7,088,713**	**2.56**

From Wegrzyn et al. (2013)

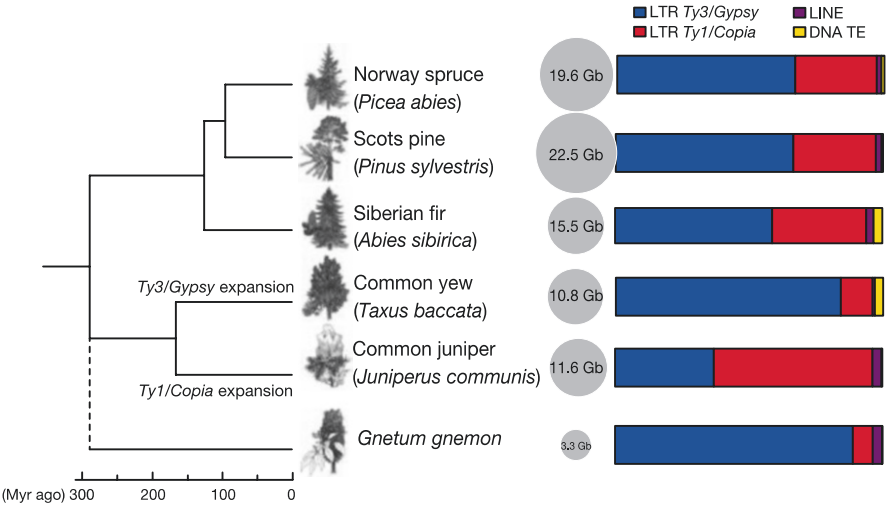


Fig. 4.5 Conifer genomes contain expansions of a diverse set of LTR-RTs. Distribution of different classes of transposable elements from six gymnosperm species. The figure is based on the total fraction of transposable elements (TE) identified and grouped into different classes from the different species. Genome sizes of the six species are given in circles and their phylogenetic relationship is shown at the left, with tentative dating of divergence times (*x*-axis) based on 64 chloroplast genes over 39 species and five fossil calibration points. (From Nystedt et al. 2013)

The genome sequencing and annotation of *Pinus taeda* (Neale et al. 2014; Wegrzyn et al. 2014) confirms the abundance and distribution of transposons and retrotransposons inferred from earlier studies of BAC sequencing (Kovach et al. 2010; Wegrzyn et al. 2013). Class I (LTR) repeats are more abundant than Class II (non-LTR) repeats and *gypsy*-like and *copia*-like retrotransposon elements are by far the most abundant (Fig. 4.6). The type, abundance, and distribution of transposon and retrotransposon elements in the second *Pinus* genome to be sequenced, *P. lambertiana* (Stevens et al. 2016), was very similar to that found in *P. taeda* with the exception of the *gypsy*-like class being 35% greater in *P. lambertiana* (Fig. 4.7). It was also shown that these *gypsy*-like elements are much younger and specific to *P. lambertiana* and thus may be responsible for the much larger size of this genome (Chap. 2). Finally, the type, abundance, and distribution of repetitive elements discovered from the sequenced *Pseudotsuga menziesii* genome (Neale et al. 2017a) were also quite similar to that found in *Pinus* and *Picea*. There is now a fairly clear understanding of the repetitive DNA content in the Pinaceae and the role of retrotransposons in producing such large genomes. It will be interesting to learn about the role of retrotransposons in some of the other conifers of smaller genome size once these genomes have been sequenced and annotated.

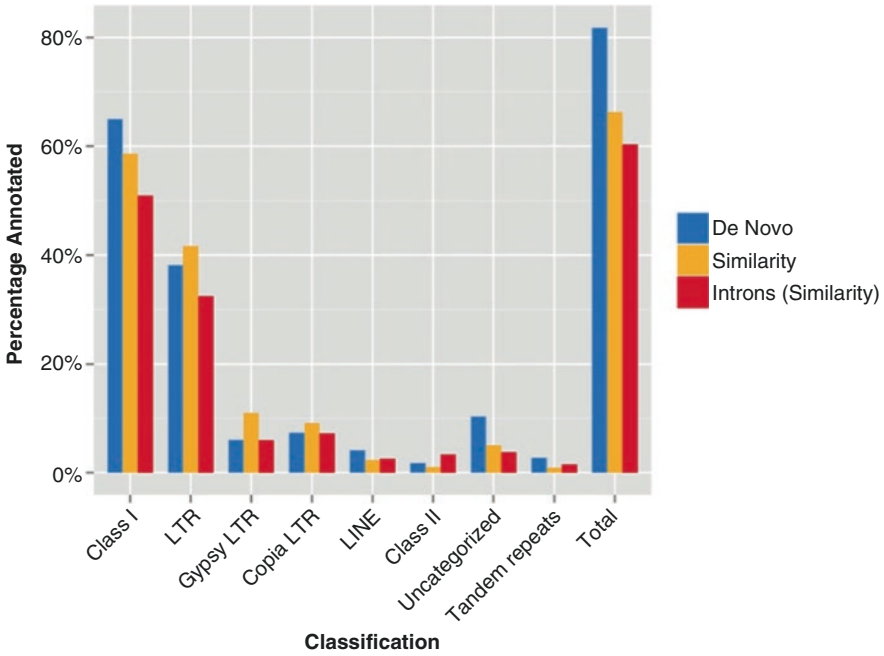


Fig. 4.6 Interspersed and tandem repetitive content in *Pinus taeda*. Overview of repetitive content for de novo (blue) and similarity (yellow) approaches. (From Neale et al. 2014)

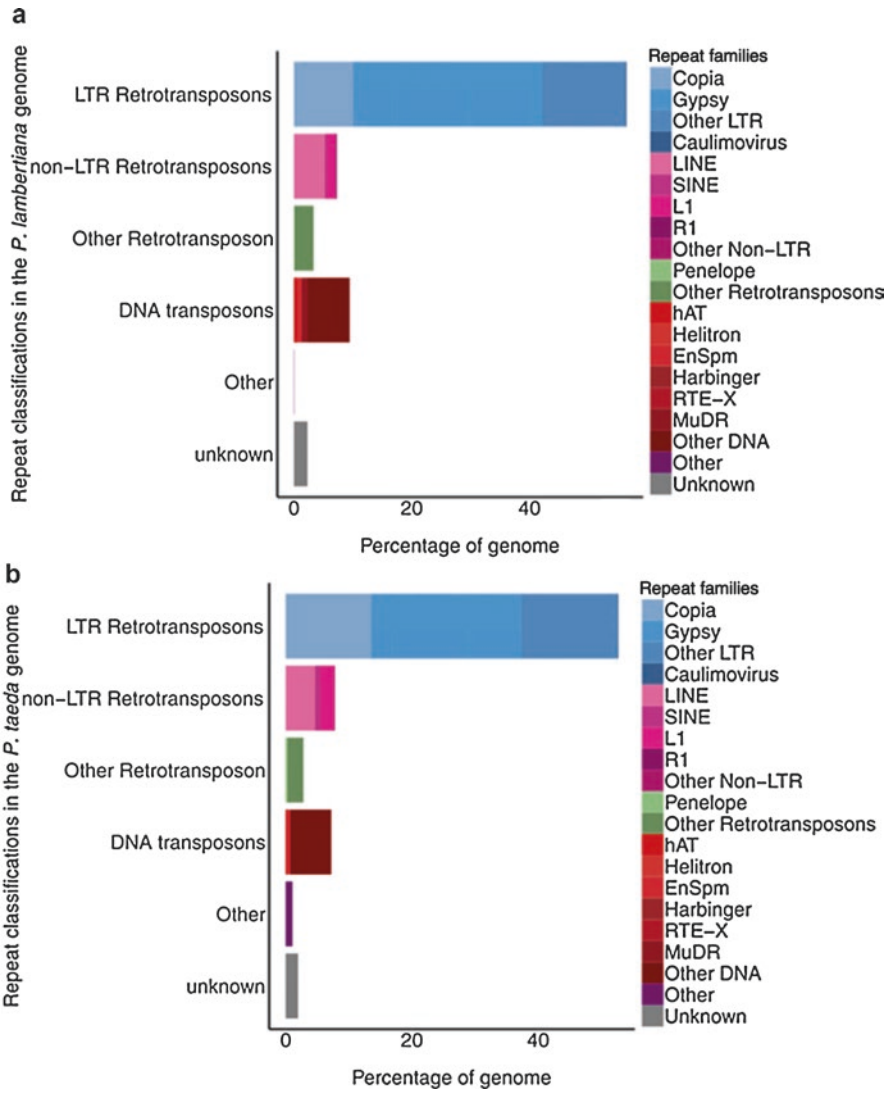


Fig. 4.7 Comparison of repetitive content between transposable element repeat families in *Pinus lambertiana* (**a**) and P. *taeda* (**b**). (From Stevens et al. 2016)

Pseudogenes

It was speculated very early on that conifers might have much larger gene families than angiosperms based on comparative Southern hybridization patterns using protein-coding gene probes (cDNA probes) (Kinlaw and Neale 1997; Chap. 2). Years later, following large volumes of EST sequencing and full genome sequencing, it has become clear that the complex Southern hybridization patterns were

likely due to mostly nonfunctional pseudogene copies and to a lesser degree to functional gene copies. In a recent review, Prunier et al. (2016) calculated that the gene-like fraction of the *Picea abies* and *Pinus taeda* genomes was estimated to be 2.4 and 2.9%, respectively, but this proportion of the genome was far in excess of what would be needed for 50,000 protein-coding genes. The assertion then was that all this additional gene-like sequence must be pseudogenes. However, there is yet to be a comprehensive analysis of pseudogene content in conifers that would bring a deeper understanding to what portion of conifer genomes are made up of noncoding but not repetitive DNA. This question has only been addressed in the context of just a few gene families in a couple of species. Following manual annotation of the terpene synthase and P450 gene families and genes in the mevalonate, methylerythritol phosphate, and phenylpropanoid pathways, Warren et al. (2015) estimated that approximately 50% of the gene-like sequences were actually pseudogenes. In a study of genes coding for dicer-like proteins in *Pinus lambertiana*, González-Ibeas et al. (2016) concluded that many of the gene models were pseudogenes. In summary, studies to date suggest that pseudogenes may only make up 1% or more of the total genome content of conifers, but the functional role and evolutionary potential of the pseudogene component are not yet known.

Summary

The reason conifer genomes are so large is that they are made up of vast amounts of noncoding and repetitive DNA. The number of protein-coding genes does not differ significantly from that in all other plant and animal species. The noncoding and repetitive DNA can be assigned to four major types: (1) ribosomal RNA genes (rRNA genes), (2) tandemly duplicated DNA, (3) transposons and retrotransposons, and (4) pseudogenes. These four classes of noncoding and repetitive DNA are found in all higher plants and animals; however, the transposon and retrotransposon class in conifers has been amplified to a very large extent over evolutionary time and accounts for the large size of current conifer genomes. There is preliminary evidence that mechanisms active in angiosperms to purge genomes of excess transposon and retrotransposon sequence are missing in conifers. The evolutionary explanation as to why this might be true is yet to be understood.

Gene Structure and Gene Families

5

A Short History of Early Conifer Gene Sequencing

Approaches for the discovery of protein-coding genes were described in detail in Chap. 3. These included EST sequencing, RNA-seq, and full genome sequencing. Raw DNA sequences generated by these approaches are assembled into sets that are assumed to come from a single genetic locus. These are often called *unigene sets*. Currently, it is believed that there may be as many as 50,000 unique genes (*unigenes*) in conifer genomes, although this number will likely decline as more sequence data and better bioinformatics tools reveal that the number of *unigenes* has been overestimated. This can result from allelic or alternative splicing differences at a single locus. In this chapter, our goal is to discuss what is known about the structure (introns, exons, promoters, other regulatory regions) of conifer genes and their relationship to one another in gene families. The discussion is organized around functional classes of genes that have been of most interest in conifers.

In the era before complete genome sequencing, approaches to clone and sequence conifer genes were tedious and time-consuming and mostly focused on transcripts and cDNAs, which offered limited insights to interpret gene structure. In the period before EST resources were available, two general approaches were used. The first, a multi-step process, began by purifying a protein product of a gene of interest, determining its complete or partial amino acid sequence, then designing PCR primers or oligonucleotide probes from the reverse-translated amino acid sequence, and finally screening cDNA libraries. This approach is technically demanding, takes a very long time, and was used only a few times to clone a gene and determine its structure (see the sections below on wood-forming genes and defense-related genes). The second approach, following the discovery of PCR, was designing primers based on sequences of the target gene previously cloned and sequenced in other species. This approach is called *cloning by homology*. It was quite easy and fast to apply and was widely used. The great limitation of this approach is that it would only yield genes previously cloned and sequenced in other species, and furthermore

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_5

the sequence divergence between those species and the conifer could not be too large or else PCR would fail. Therefore, genes unique to conifers, and thus potentially of greatest interest, would not be found when using this approach.

Once EST sequencing arrived, many more genes were suddenly available for study. But as we noted in Chap. 3, this approach begins with isolating mRNA from one or more tissue types, so in general, genes that are abundantly expressed in multiple tissues are more easily found (often referred to as *housekeeping genes*) than are genes expressed in a tissue-specific or developmental-specific manner. The latter, again, are potentially unique and more interesting. Furthermore, because it was cDNA that was sequenced there was limited opportunity to learn about gene structure. Finally, once complete genomes could be sequenced, the complete gene catalog of a conifer could be obtained in an unbiased manner. Because the first conifer genomes were published beginning only in 2013, there has been little time for researchers to study the conifer gene structure in a comprehensive and comparative manner. This chapter will then chronicle what was learned about conifer gene structure before complete genome sequencing.

Wood-Forming Genes

Genes involved in the formation of wood, specifically secondary xylem, have been a topic of great interest in forestry for obvious reasons. Genes in the pathways leading to the three major components of xylem (lignin, cellulose, and hemicellulose) have been studied extensively, although primarily in the angiosperms poplar and eucalyptus. There is a much smaller body of work in conifers. This is likely due to the ability to easily transform and regenerate some of these angiosperm tree species which facilitates studies of gene function. Nevertheless, we will summarize what is known about genes involved in wood formation in conifers. Genes in the lignin biosynthetic pathway have been of greatest interest with the eventual application of down-regulating the pathway and reducing the lignin content in xylem. This is a desirable property for paper making.

Genes in the synthesis of monolignols in the phenylpropanoid pathway (Fig. 5.1) were first cloned and sequenced in the early 1990s (Peter and Neale 2004). The first genes cloned were phenylalanine ammonia-lyase (PAL) from *Pinus taeda* (Whetten and Sederoff 1992) and cinnamyl alcohol dehydrogenase (CAD) from *P. taeda* (O'Malley et al. 1992) and from *Picea abies* (Galliano et al. 1993). In all three cases, the extent of sequence similarity between the conifer gene and angiosperm homologs was on the order of 60% (Table 5.1), thus providing an early indication of the amount of sequence divergence between an angiosperm and a gymnosperm. All that was obtained in these early studies was a full-length or partial cDNA clone, so information about intron-exon structures or regulatory sequences could not be obtained.

The work to more fully characterize the CAD gene from *Pinus taeda* continued for several years. Mackay et al. (1995) showed that CAD was controlled by a single gene based on segregation analysis and obtained a full-length cDNA that was used in comparative sequence analysis (Fig. 5.2). Soon after, Mackay et al. (1997)

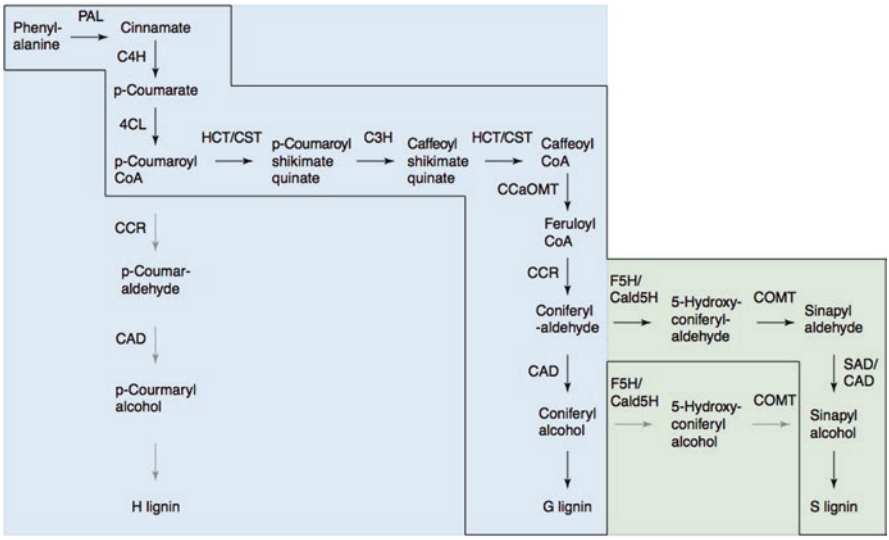


Fig. 5.1 The predominant pathway for monolignol biosynthesis in xylem cells is outlined in black, with the dark arrows showing the primary substrates and products and the gray arrows showing the minor substrates and products. The blue shading indicates the pathway that is conserved between angiosperms and gymnosperms, whereas the green shading indicates the angiosperm-specific pathway. (From Peter and Neale 2004)

Table 5.1 Comparison of *Pinus taeda* CAD and alcohol dehydrogenase N-terminal protein sequences with those from two angiosperm species

Sequence	Similarity (%)	Identity (%)
Pine CAD N-terminus to bean λCAD4	51	19
Pine CAD N-terminus to maize ADH1F	57	24
Pine ADH N-terminus to bean λCAD4	43	30
Pine ADH N-terminus to maize ADHIF	84	80

From O'Malley et al. (1992)

discovered that an elite *Pinus taeda* tree carried a null allele at the CAD locus and that this allele was associated with reduced lignin content in this tree (Ralph et al. 1997) (Table 5.2). It was not until several years later that the exact sequence basis of this null-allele mutation was discovered. Gill et al. (2003) showed that the null allele was caused by a two-base AA insertion into the fifth exon of the CAD gene that caused a premature termination of the mature protein (Fig. 5.3).

In 2009, Bedon et al. (2009) first reported on the sequence analysis of the 5′ promoter region of the CAD gene in both *Pinus taeda* and *Picea glauca*. This study showed that the CAD gene promoter has the binding sites for the MYB, WRKY, and bHLH transcription factors. This opened the door to additional studies of transcription factor binding.

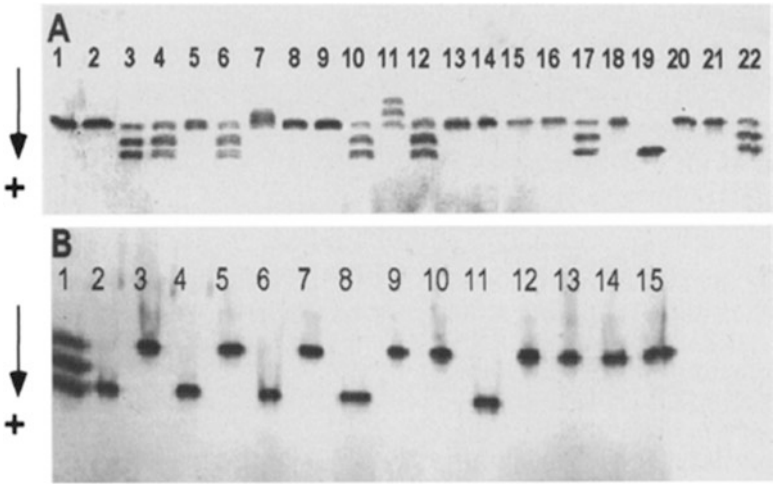


Fig. 5.2 Genetic variation and inheritance of CAD allozymes, as detected on nondenaturing polyacrylamide gel electrophoresis (PAGE) by enzyme-activity staining. (**a**) CAD electrophoretic patterns in crude xylem samples from a population of 22 *Pinus taeda* trees (lanes 1–22) and (**b**) segregation analysis of one heterozygous tree. Xylem tissue (lane 1) and the haploid megagametophyte tissue from 14 different seeds collected from the same tree (lanes 2–15). (From Mackay et al. 1995)

Table 5.2 Effect of the *cad-n1* allele on lignin content (% dry weight of cell wall residue (CWR)) as determined by two methods. ND means not determined

	Klason method			AcBr method		
cad genotype	Acid insoluble	Acid soluble	Total	AcBr	AF	Total
Cad/Cad	31.0 ± 0.7	0.71 ± 0.1	31.7 ± 0.8	30.7 ± 0.9	0.9 ± 0.1	31.6 ± 1.0
Cad/cad-n1	30.8 ± 0.7	0.69 ± 0.1	31.5 ± 0.8	ND	ND	ND
Cad-n1/ cad-n1	28.1 ± 0.4	0.82 ± 0.1	28.9 ± 0.5	19.2 ± 0.7	8.1 ± 0.2	27.3 ± 0.9
Relative content	91%		91%			86%

From Mackay et al. (1997). Copyright (1997) National Academy of Sciences, USA

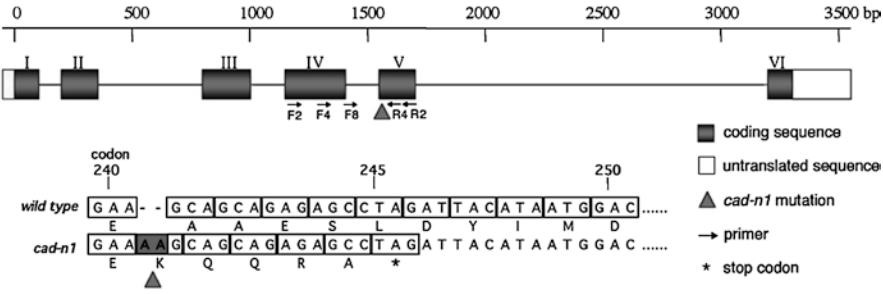


Fig. 5.3 Position of the *cad-n1* sequence mutation within the *cad* gene in *Pinus taeda* and effect of the frame shift on amino acid sequence. (From Gill et al. 2003)

The first report of the cloning and sequencing of a transcription factor controlling the expression of a gene in the lignin biosynthetic pathway was that of Patzlaff et al. (2003a) who cloned a member of the R2R3-MYB family from *Pinus taeda*. Electrophoretic mobility shift assays showed that the MYB transcription factor bound the expected DNA elements and that this would control the expression of genes having these elements in their promoter sequences. In a related study, Patzlaff et al. (2003b) cloned a second R2R3-MYB family transcription factor and showed that it could bind a PAL promoter and affect its expression. Likewise, R2R3-MYB family transcription factors were also discovered in *Picea marina* (Xue et al. 2003) and *P. glauca* (Bedon et al. 2007). Bedon et al. (2010) completed a more comprehensive study in both *Picea glauca* and *Pinus taeda*, discovering at least ten different R2R3-MYBs and characterizing gene expression controlling differences among genes. Another class of transcription factors, HD-Zip II genes, were also found in *Picea glauca* (Côté et al. 2010). Collectively, these studies show the important functional role of these transcription factors in conifers that were known from angiosperms and how these transcription factor families may have evolved to control secondary xylem development in conifers. Finally, a few other genes involved in lignin and cellulose biosynthesis were cloned and sequenced during this period including laccases (Bao et al. 1993; Sato et al. 2001), cellulose synthases (Nairn and Haselkorn 2005), expansins (Sampedro et al. 2006), and a 4-coumarate:coenzyme A ligase (Zhang and Chiang 1997).

Vegetative Growth Genes

Genes in pathways controlling vegetative growth patterns in conifers are of obvious interest to forest geneticists. There are many pathways and genes leading to meristematic tissues of shoots and roots, but only a small number of genes have been the subject of detailed investigation. Two general types of genes have been the focus of most research: (1) genes controlling the cell cycle in meristematic tissues and (2) homeobox transcription factors controlling the expression of genes in meristematic pathways.

The first gene to be cloned and sequenced was the *cdc2* gene from *Picea abies* (Kvarnheden et al. 1995). A cdc gene was first cloned from yeast and its role in controlling the cell cycle was established. The *cdc2* gene from *Picea abies* was cloned by homology and found to be highly identical (85–90%) to other plant *cdc2* genes, and its intron-exon structure was identical to that of *Arabidopsis thaliana*. Furthermore, it was estimated that the *cdc2* gene was a member of a gene family of about ten genes, most of which were pseudogenes (Table 5.3). In a follow-on study, Kvarnheden et al. (1998) found the same genes in two other spruces, *Pinus engelmannii* and *Picea sitchensis*, although the number of pseudogenes in P. abies was greater. A *cdc2* gene was also found in *Pinus contorta*, but the large gene family did not appear to be present. This early work was all conducted by the same group in Sweden but was not continued beyond this early stage.

Table 5.3 Features of the *Picea abies cdc2Pa* gene and related processed retropseudogenes

Sequence	Length (bp)	Nucleotide identity to *cdc2Pa* (%)	Specific mutation
cdc2Pa complete	1328		
cdc2Pa coding	882		
cdc2Pa Ψ1	623	89.1	Insertion of 1 bp
			Insertion of 4 bp
cdc2Pa Ψ2	197	92.4	Deletion of 1 bp
cdc2Pa Ψ3	548	89.2	Insertion of 1 bp
cdc2Pa Ψ4	85	85.9	Deletion of 1 bp
			Deletion of 4 bp
cdc2Pa Ψ5	90	85.6	Amino acid D146 in *cdc2Pa* changed to N
cdc2Pa Ψ6	443	90.6	Internal stop-codon

From Kvarnheden et al. (1995)

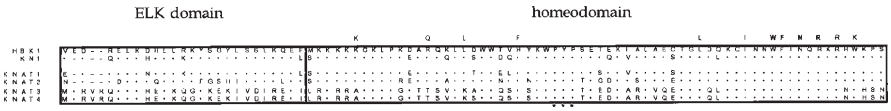


Fig. 5.4 Alignment of the amino acid sequences deduced from *Picea abies HBK1*, maize *Kn1,* and *Arabidopsis KNAT1–4* genes. *Kn1*, *KNAT1,* and *KNAT2* represent KNOX class I genes and *KNAT3* and *KNAT4* represent KNOX class II genes. ELK and the homeodomain regions are boxed. The three amino acids inserted in all TALE homeodomains are indicated by asterisks. Amino acids conserved among most homeodomains are indicated above the sequence, and invariant amino acids are shown in bold. (From Sundås-Larsson et al. (1998). Copyright (1998) National Academy of Sciences, USA)

The Swedish group then switched their attention to homeobox transcription factor genes. The first gene of this type to be cloned, again by homology, was HBK1 from *Picea abies* (Sundås-Larsson et al. 1998). It was shown to be similar to the KNOX homeobox genes from angiosperms and specifically similar to the *Knotted1* gene from maize (Fig. 5.4) which was shown as a key gene in meristem differentiation. Later, a Canadian group cloned several knox-I genes from several species of *Picea* and *Pinus* and from *Cryptomeria japonica* with the goal of understanding mutations and functional divergence among taxa (Guillet-Claude et al. 2004).

Subsequently two HD-GL2-like homeobox genes were cloned in *Picea abies* (Ingouff et al. 2001, 2003). These genes were cloned from somatic embryos and their expression was studied during embryogenesis. This research was conducted in the context of understanding the genetic control of embryogenesis.

In the context of a research program to understand the genetic control of adventitious root formation in pine, Goldfarb et al. (2003) cloned by homology and sequenced many genes from the Aux/IAA gene family from *Pinus taeda*. The pine Aux/IAA genes had a regulatory domain structure similar to that of angiosperm homologs, suggesting that their regulatory function in controlling rooting would be similar.

The final class of homeobox transcription factor genes related to meristem development that were studied in conifers were the Wuschel-related homeobox (WUS/WOX) genes. In angiosperms, the WUS genes are expressed in shoot meristems and the WOX genes are expressed in root meristems. Nardmann et al. (2009) first reported that there was just a single WUS/WOX gene in *Pinus sylvestris*, similar to that of other plants. Later, Hedman et al. (2013) cloned and sequenced 11 WUS/WOX genes from *Picea abies* and showed the large differences in intron-exon structure among these genes (Fig. 5.5). They argued that there were unique genes expressed in shoot and root meristems. Early studies show that there appears to be quite a bit of conservation in genes, gene structure, and gene function between conifers and angiosperms related to vegetative meristem development, but because all these studies used a cloning by homology approach, it remains unclear which genes unique to meristem development in conifers await discovery.

Floral Genes

Genes and pathways controlling floral development in plants, in general, are topics of great interest. In the model species *Arabidopsis thaliana*, floral identity genes have been characterized in great detail. Likewise, in angiosperm forest trees, such as poplars, these genes have been the focus of many studies. One reason for the interest in these genes is the potential opportunity to genetically engineer trees for

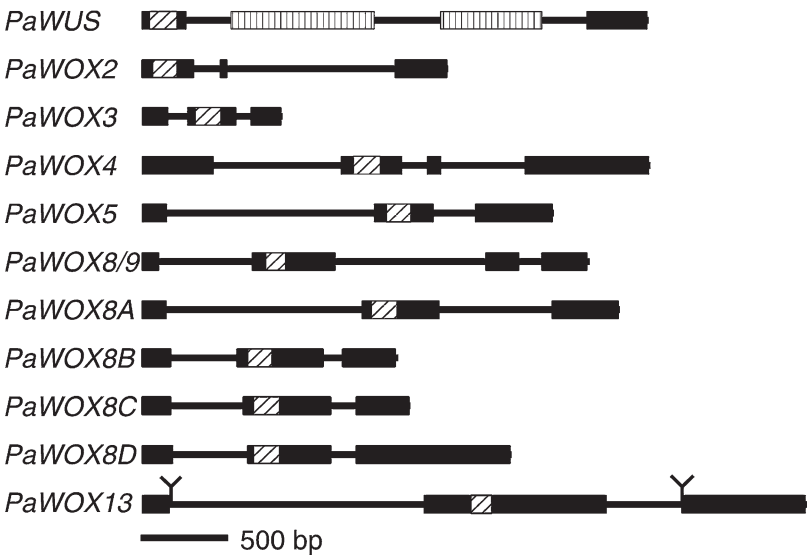


Fig. 5.5 Gene structure of *Picea abies* WOX genes. Exons are thick and black, introns are thin lines. The conserved homeobox region is striped diagonally. The repeat regions in *PaWUS* are striped vertically. The gene structure of *PaWOX13* is only putative. The ends of known intron sequences are marked with lines. (From Hedman et al. 2013)

sterility so that other transgenes, such as herbicide resistance, do not spread into wild populations. In conifers, however, there has been much less research activity on this topic, presumably due to the great difficulty in transforming and regenerating most conifer species.

The group in Sweden that did the early work on genes controlling vegetative meristem development was also the first to study genes in floral development of conifers. They again employed the cloning by homology approach using PCR primers designed from genes sequenced in *Arabidopsis thaliana*. The target genes were MADS-box transcription factor genes which control floral development in angiosperms. Tandre et al. (1995) cloned and sequenced three genes from *Picea abies* with similarities to the DEFIECIENS and AGAMOUS genes in *Arabidopsis thaliana*, which they called *deficiens-agamous-like* (*dal*) genes. The *P. abies* genes had the signature MADS box and K box domains. Subsequently, Tandre et al. (1998) performed gene expression experiments to show that one of the *dal* genes, *dal2*, was functionally similar to AGAMOUS in *Arabidopsis thaliana*. Later, Carlsbecker et al. (2003) found one *dal* gene, *dal10*, that appeared to be unique in conifers. Also, using a cloning by homology approach in *Cryptomeria japonica*, Fukui et al. (2001) found two MADS-box genes with sequence homology to the PISTILATA and APETALA3 genes in *Arabidopsis thaliana*. Like the other studies with genes controlling vegetative meristem development, these studies with genes involved in floral development suggest conservation in genes and gene function between angiosperms and gymnosperms, but again this inference is confounded by the cloning by homology approach that was used.

In one of the very first examples of using the newly produced conifer reference genome sequences, Gramzow et al. (2014) compared the complete set of MADS-box genes in conifers to the complete set in angiosperms. They found one type, Type I, to be deficient in conifers, whereas Type II was more common in conifers than in angiosperms (Table 5.4). This study demonstrates clearly the power of having reference genome sequences in conifers and the limitation and bias associated with the earlier studies that used the cloning by homology approach.

Light-Regulated Genes

The study of genes whose expression is controlled by light was a topic of great interest beginning in the late 1970s, primarily in the model species *Arabidopsis thaliania* but also in a small number of other herbaceous annual plant species. This class of genes was of interest for conifers because of their perennial habit and dormancy requirements. One of the first light-regulated genes to be cloned and sequenced was the ribulose-1,5-bisphosphate carboxylase (RBCS) gene from *Larix laricina* (Hutchison et al. 1990). Both a genomic clone and several cDNA clones were sequenced so intron-exon structure could be inferred and upstream regulatory regions observed. Sequence comparisons with angiosperm homologs reached the familiar conclusion of conservation of gene function (Hutchison et al. 1990).

Table 5.4 Number of MADS sequences identified from the gymnosperm genome and transcriptome data. For the three species for which whole-genome information is available, the numbers are given as follows: the number of MADS sequences identified from genome data (and supported by transcriptome data) plus the number of MADS sequences identified from transcriptome data for which the genomic locus could not be identified

Family	Species	Abbreviation	Total	Type	
				I	II
Order **Gnetales**					
Gnetaceae	*Gnetum gnemon*	GgMADS	41	0	41
Order **Coniferales**					
Pinaceae	*Cedrus atlantica*	CaMADS	13	0	13
	Picea abies	PaMADS	253 (41) + 8	12	249
	Picea glauca	PgMADS	107 (9) + 14	3	118
	Picea sitchensis	PsMADS	17	1	16
	Pinus banksiana	PbMADS	2	0	2
	Pinus contorta	PcMADS	14	0	14
	Pinus lambertiana	P1MADS	41	0	41
	Pinus palustris	PpaMADS	21	0	21
	Pinus pinaster	PpiMADS	10	0	10
	Pinus taeda	PtaMADS	346 (39) + 21	17	350
	Pseudotsuga menziesii	PmeMADS	40	0	40
Podocarpaceae	*Podocarpus macrophyllus*	PmaMADS	16	1	15
Araucariaceae	*Wollemia nobilis*	WnMADS	11	0	11
Sciadopityaceae	*Sciadopitys verticillata*	SvMADS	22	1	21
Taxaceae	*Taxus baccata*	TbMADS	3	0	3
Cephalotaxaceae	*Cephalotaxus harringtonia*	ChMADS	35	0	35
	Cryptomerica japonica	CjMADS	10	0	10
Cupressaceae	*Sequoia sempervirens*	SsMADS	19	0	19

From Gramzow et al. (2014)

Another early light-regulated family of genes to be cloned in a conifer were the chlorophyll a/b binding genes (CAB genes), that were later renamed light harvesting complex (LHC) genes. Jansson and Gustafsson (1990, 1991) cloned by homology and sequenced two cab-II (LHC II) genes and three cab-I (LHC I) genes from *Pinus sylvestris*. Sequence homology among LHC I and LHC II genes in *P. sylvestris* was high (>80%) and also with homologs from tomato. This again argues for conservation in function between angiosperms and gymnosperms. Yamamoto et al. (1993) also cloned and sequenced both LHC I and LHC II from *P. thunbergii* and showed sequence similarity to the *P. sylvestris* genes. Finally, Barrett et al. (1994) cloned and sequenced a cab-II gene from *P. contorta* and showed that it also had high similarity to the *P. sylvestris* homolog, but because this gene was cloned from genomic DNA, and not cDNA as was the *P. sylvestris* gene, they were able to show that cab-II lacked introns. Further, they showed by sequence comparison to angiosperm genes that there was a lack of sequence similarity in the regulatory regions. So here we see researchers beginning to note differences versus similarities between angiosperms and conifers.

Because it was known for some time that conifer seedlings which germinated and grew in the dark produced chlorophyll and were green (Oku et al. 1974), as opposed to being etiolated like most angiosperms, researchers wondered if the LHC genes in conifers were light regulated. Initially, in *Pseudotsuga menziesii* (Alosi et al. 1990) and in *Pinus thunbergii* (Yamamoto et al. 1991) and then later in a suite of conifers (Mukai et al. 1992), it was shown that *cab* and *rbcs* were light-independently regulated. These observations stimulated research to understand, at the DNA sequence level, what might be responsible for light-independent regulation. Using a reporter gene approach, Yamamoto et al. (1994) showed that it was the 5′ promoter region that was responsible for the light-independent expression, although the exact DNA sequence differences with light-dependent angiosperm promoters were not reported.

Another family of light-regulated genes of great interest in plants are those coding for the photoreceptor phytochrome. It had been established in angiosperms that the light regulation of LHC and RBCS genes was in fact phytochrome regulated. Alosi and Neale (1992) showed that even though the LHC and RBCS genes were light-independently regulated in *Pseudotsuga menziesii*, their expression levels could still be affected by red and far-red light treatment. Elmlinger et al. (1994) showed that the glutamine synthase gene was both light and phytochrome regulated. The first report of cloning and sequencing of phytochrome genes from a conifer was in *Picea abies* and *Pinus sylvestris* (Clapham et al. 1999). There are five known phytochrome genes in angiosperms, PHYA-PHYE. The *Picea abies* gene was similar to the PHYA gene and the *Pinus sylvestris* gene was similar to PHYB. In a pair of papers to understand population genetics and evolutionary divergence of phytochrome genes in *Pinus sylvestris*, García-Gil et al. (2003) found that the number of non-synonymous mutations in PHYA and PHYB genes was quite low, suggesting there may be selective and functional constraints on diversity in these genes in conifer populations. However, when diversity was estimated in phytochrome pseudogenes, higher levels of diversity were found (García-Gil 2008). Although the work in light-regulated genes in conifers is not as comprehensive as the other classes of genes discussed so far, it is shown that important differences in genes, gene structure, and gene expression were notable between angiosperms and conifers at the DNA sequence level.

Defense-Related Genes

A signature characteristic of conifers is their production of oleoresins for defense against insect herbivores and fungal pathogens. A study of the chemical components of oleoresins (monoterpenes, sesquiterpenes, and diterpenes, Fig. 5.6) was conducted but it was not until the 1990s that researchers began to clone, sequence, and study the genes in the biosynthetic pathways for these compounds. As discussed in this and previous chapters, primary discovery studies in conifer genomics have most often been done in *Pinus* and *Picea,* only occasionally in other genera. However, in the case of cloning genes coding for terpene synthases (Tpsd), the

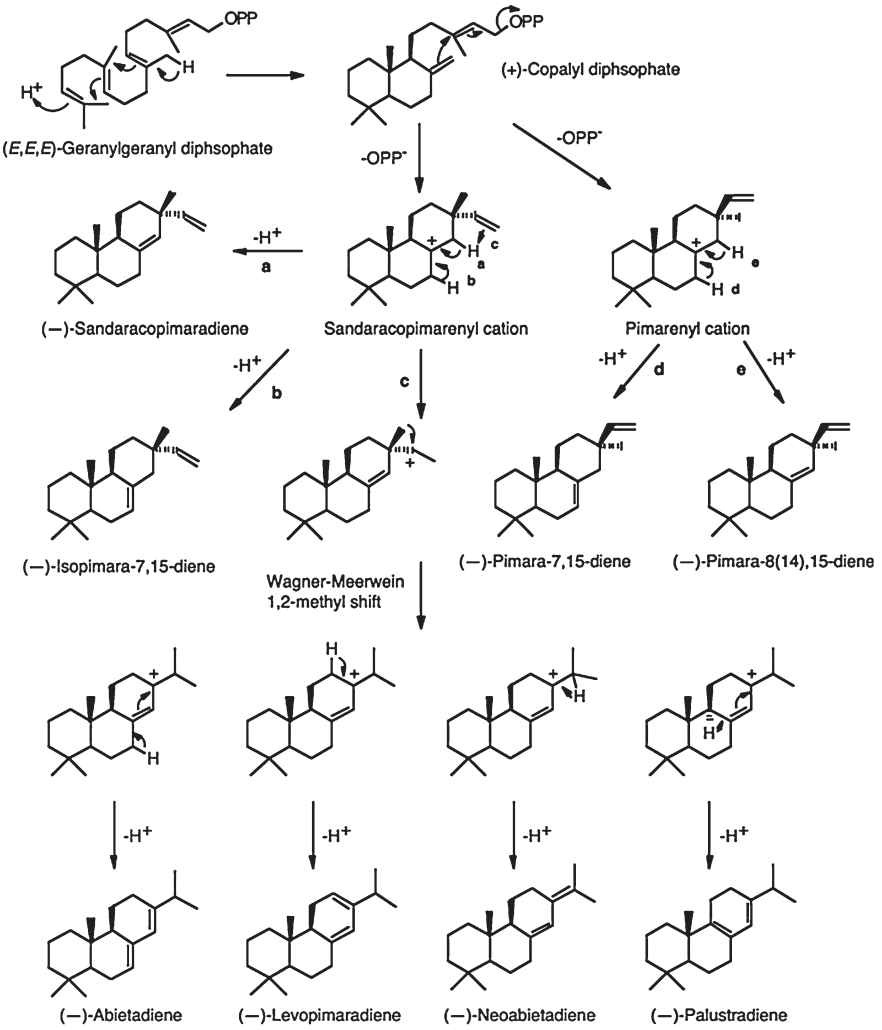


Fig. 5.6 Proposed reaction mechanism for the biosynthesis of some diterpenes from geranylgeranyl diphosphate. OPP represents a diphosphate group. (From Keeling and Bohlmann 2006a)

pioneering work was done in *Abies grandis*. Seven monoterpene synthase genes from the terpene synthase (Tpsd) gene family were all cloned by homology and sequenced from *Abies grandis* (Bohlmann et al. 1997, 1999). In parallel, both diterpene synthase genes (abietadiene synthase) (Vogel et al. 1996) and sesquiterpene synthase genes (Bohlmann et al. 1998a; Steele et al. 1998) were also cloned and sequenced from *Abies grandis*. Sequence comparison analysis among all the *Abies grandis* monoterpene, diterpene, and sesquiterpene synthase genes shows the close evolutionary relationship among these genes but is quite different from angiosperm terpene synthases, which is in keeping with the well-developed and specialized oleoresin chemistry of conifers (Bohlmann et al. 1998b).

Terpene synthase (TPS) gene discovery in *Picea* provides an outstanding example of the use of developing technologies to ultimately produce a complete catalog of all genes, which then enables a more thorough understanding of the functional role of these genes in oleoresin synthesis and chemical defense. To begin, Martin et al. (2004) used a cloning by homology strategy to obtain nine TPS gene family members from *Picea abies* (Table 5.5). These genes were all members of one TPS gene family, *Tsp-d.* In the next phase of discovery, Keeling et al. (2011) screened a large *Picea* EST resource and found 69 TPS genes. The number discovered here exceeded the number previously found in angiosperms and is consistent with the great importance of terpene synthesis for oleoresin production in conifers. In an intermediate step in advance of a complete genome sequence, a *Picea glauca* BAC library was screened for clones containing either the TPS or cytochrome P450 genes, or both (see Chap. 3), and two complete genes were discovered. For the first time the complete gene structure (intron-exon, regulatory regions) of genes of oleoresin biosynthesis could be determined. Analysis of the 5′ regulatory sequences gave an indication of the transcription factors binding these genes and an indication of the molecular basis of the expression of these genes in conifers. Once the *Picea glauca* genome was sequenced, Warren et al. (2015) found 83 TPS genes and many pseudogenes. From this, the number and proportion of monoterpene synthases, sesquiterpene synthases, and diterpene synthases in *Picea glauca* were determined (Fig. 5.7). The University of British Columbia research group now has annotated a near-complete set of TPS genes with which to conduct functional studies in oleoresin production and insect defense. In a similar fashion, the same group also discovered and annotated cytochrome P450 genes in *Picea glauca* (Warren et al. 2015) and has begun to establish functions of members of this large gene family in oleoresin formation in different *Pinus* species and in *Picea sitchensis* (Ro et al. 2005; Hamberger et al. 2011; Geisler et al. 2016).

Table 5.5 Gene name, accession numbers, and functional annotation of *Picea abies* TPS genes

Gene	Clone	Accession	TPS class	Functional annotation
PaTPS-Car	PaJF67	AF461459	Mono-TPS	(+)-3-Carene synthase
PaTPS-Lim	PaDM743	AY473624	Mono-TPS	(−)-Limonene synthase
PaTPS-Myr	PaJBl6	AY473626	Mono-TPS	Myrcene synthase
PaTPS-Lin	PaJF39	AY473623	Mono-TPS	(−)-Linalool synthase
PaTPS-Pin	PaJF104	AY473622	Mono-TPS	(−)-α/β-Pinene synthase
PaTPS-Far	PaJF71	AY473627	Sesqui-TPS	E,E-α-Farnesene synthase
PaTPS-Bis	PaDM03	AY473619	Sesqui-TPS	E-α-Bisabolene synthase
PaTPS-Lon	PaDM486	AY473625	Sesqui-TPS	Longifolene synthase
PaTPS-LAS	PaDM2420	AY473621	Di-TPS	Levopimaradiene/abietadiene s.
PaTPS-lso	PaDM2425	AY473620	Di-TPS	Lsopimara-7,1 5-diene synthase

From Martin et al. (2004)

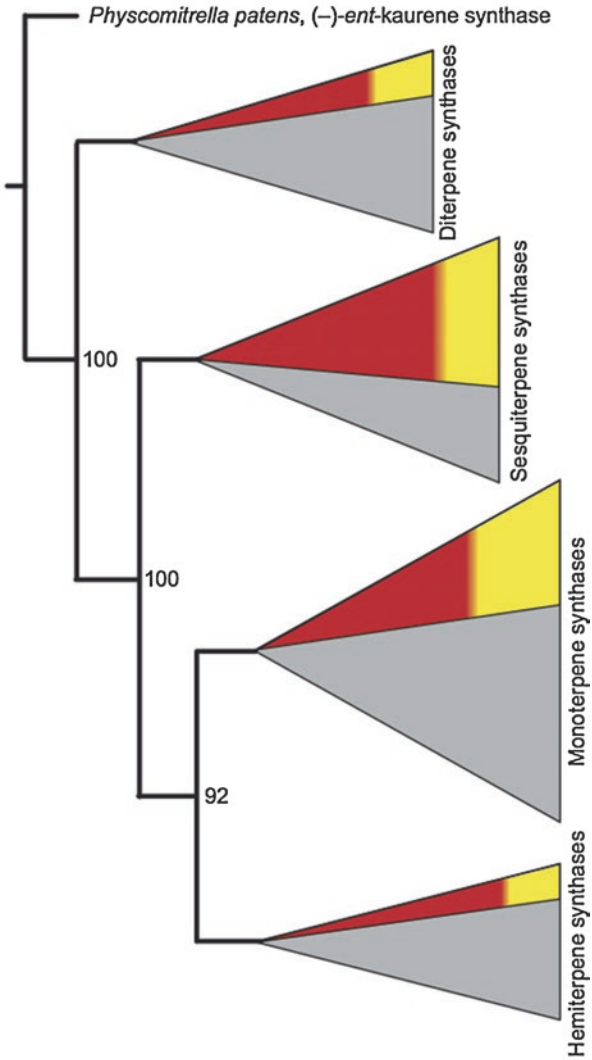


Fig. 5.7 Schematic phylogenetic tree of gymnosperm and *Picea glauca* terpene synthases. A schematic phylogenetic tree of the large family of gymnosperm terpene synthase proteins (≥400 amino acids) is shown, with *Physcomitrella patens* entkaurene synthase as the root. Relative areas show the proportion of distinct clades of diterpene, sesquiterpene, monoterpene, and hemiterpene synthases. The proportions of each synthase type originating from *Picea glauca* are shown in red and yellow, with yellow being the proportion of putative *P. glauca* pseudogenes. Bootstrap values are indicated at the nodes. (From Warren et al. 2015)

Disease-Resistant Genes

Mechanisms of resistance to pathogens and breeding for resistance have long been active areas of research in conifers (Chap. 14). Two forms of resistance are generally recognized: (1) qualitatively inherited (vertical) and (2) quantitatively inherited (horizontal). In Chap. 11, we will discuss genetic mapping approaches to begin identifying genes for each of these resistance types. In this chapter, however, we will discuss what is known about disease-resistance-related genes in conifers that have previously been discovered and characterized in angiosperms. The cloning by homology approach was used almost exclusively to find these genes in conifers.

One of the first disease-resistance-related genes to be cloned and sequenced in a conifer was a chitinase gene from *Pinus taeda* (Chang et al. 1996). A cDNA clone, *pLP6*, was isolated from a differential screen of cDNA made from mRNA of root tissues from watered and unwatered seedlings. Following sequencing, *pLP6* was shown to be a class I chitinase gene which is associated with stress or wounding. Comparison to angiosperm chitinases revealed that the domain structure and 5′ untranslated region of this gene were quite different and therefore novel. In *Pinus strobus*, Wu et al. (1997) cloned by homology a class II chitinase gene. This chitinase gene had an intron-exon structure identical to other class II chitinase genes from angiosperms. An endochitinase gene from *Pinus monticola* was also cloned by Liu and Xiang (2011) and later it was shown that there was considerable population diversity in this gene that might be associated with resistance (Liu et al. 2014).

Another class of disease-resistance-related genes that has been studied extensively in angiosperms are the nucleotide-binding-site and leucine-rich-repeat (NBS-LRR) genes. This class of genes is divided into two major groups, the Toll-like receptor (TIR) and the coiled-coil (CC) classes. Cloning by homology has found both types in conifers. A TIR-NBS-LLR gene was found in *Pinus monticola* (Liu et al. 2003b) and a CC-NBS-LLR gene was found in *Pinus lambertiana* (Jermstad et al. 2006). Both of these species are infected by white pine blister rust (*Cronartium ribicola*) and each has both qualitative and quantitative resistance mechanisms, but the genes' underlying resistance was not known. Both reports reach the conclusion that these genes may be involved in pathogen resistance in conifers and be functionally like those of angiosperms. However, they observed contrasting results in terms of population diversity in these genes. In *Pinus monticola*, a large amount of variation in the TIR-NBS-LLR genes was found and the K_a/K_s ratio (see Chap. 11) indicated diversifying selection. In *P. lambertiana*, in contrast, the CC-NBS-LLR genes were completely monomorphic (Fig. 5.8) in population samples, suggesting this gene is under purifying selection. The recently completed sequence of the *P. lambertiana* genome suggests that the dominant gene for resistance to white pine blister rust (*Cr1*) may be an NBS-LLR gene (Stevens et al. 2016).

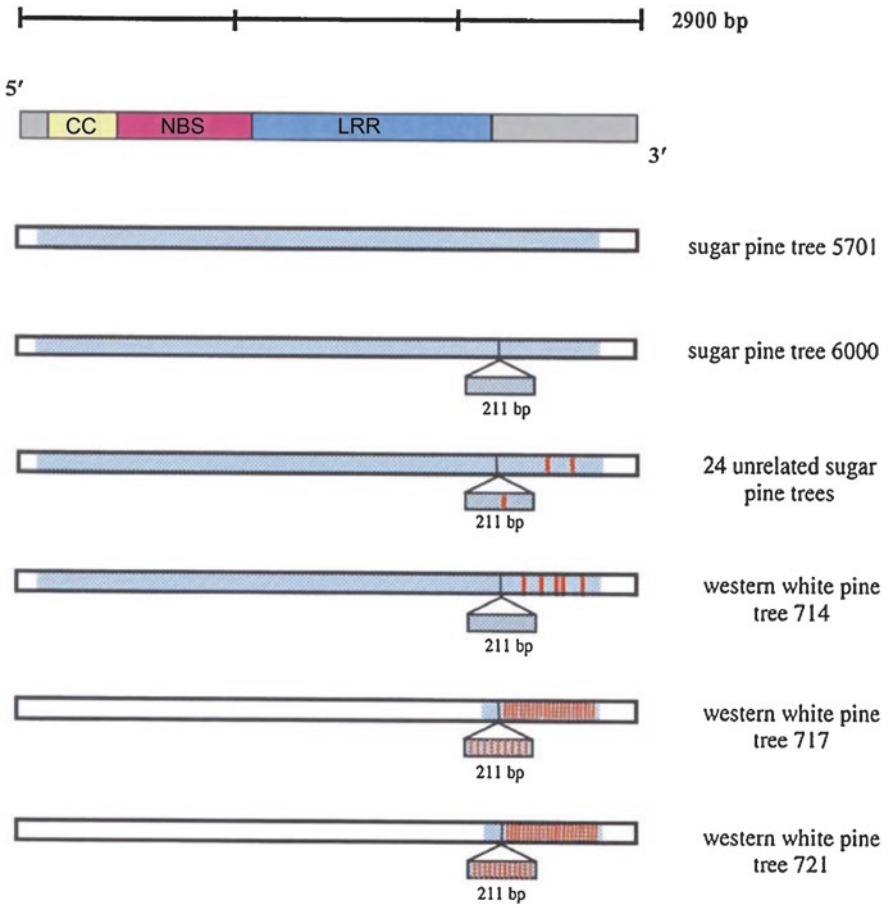


Fig. 5.8 Characterization of the protein predicted by Pl_RGC-CC_1 in *Pinus lambertiana* (sugar pine) and *P. monticola* (Western white pine). A start codon is located at position 55 and a stop codon at position 2388. Gray-shaded areas represent the noncoding regions. Blue-shaded areas represent the region of RGC-CC_1 that was analyzed by sequence in each of the sample sets. Red vertical lines represent SNPs that were detected. Only sugar pine tree 5701 contained a 211-bp deletion at position 2428. (From Jermstad et al. 2006)

Summary

Developing a comprehensive understanding of the number of protein-coding genes, the organization of these genes in gene families, and their underlying structure (intron, exon, regulatory regions) is paramount for understanding the expression of genes (Chap. 6) and the functional differences due to genetic variation among genes (Chap. 10). Conifer geneticists began this work immediately following the ability to clone and sequence individual genes. The first approach to cloning a gene from a conifer involved first isolating and sequencing the protein and then screening cDNA libraries. This approach was extremely laborious and only a couple of genes were cloned in this way. Following the invention of PCR, conifer genes could be cloned based on sequence homology to genes previously cloned in other plant or even animal systems. This approach was widely used and much was learned about gene number, gene families, and gene structure. However, because cloning was based on homology to non-conifers, genes unique to conifers based on sequence escaped detection. EST libraries were the next advancement that led to large numbers of cloned genes, but because these libraries were derived from cDNA, information on gene structure remained limited. It was not until the first conifer genomes were sequenced beginning in 2013 (Chap. 3) that a comprehensive and unbiased inventory of conifer genes and the nature of their structure were revealed. Comparative studies of sequenced conifer genomes (Chap. 17) have just begun, but once this line of research advances, a much deeper understanding of the importance of gene structure will emerge.

Gene Expression and the Transcriptome

6

A Short History of Gene Expression Studies in Conifers

The central dogma of molecular biology, first described by Francis Crick, states that the heritable transmission of information moves in the direction of the DNA to RNA (transcription) and from RNA to protein (translation). In Chaps. 2, 3, 4, and 5 of Part I of this volume, we have focused almost entirely on DNA, the genomes of conifers, and the genes encoded in the genome. In this chapter, we focus on RNA or more specifically the messenger RNA (mRNA) which collectively is called the *transcriptome*. The amount of mRNA transcribed from an individual gene at any particular time point or within any tissue type is one measure of *gene expression*. In the final chapter of this section, we will discuss the *proteome* and *metabolome*, which are additional measures of the expression of genes.

The study of the expression of genes has been of great interest in all areas of biology because the phenotypic differences observed among individuals might not only be determined by allelic differences among individuals, but also by the timing and the magnitude of expression of genes. For example, the genetic basis of why one tree might be resistant to some pathogen and another tree susceptible might be due entirely to the differences in the level of expression of one or more genes and not due to any genetic difference in the protein that the gene codes for. The heritable difference between individuals responsible for differences in levels or timing of gene expression might be found in promoter regions or in transcription factors. For all these reasons, conifer geneticists have made gene expression studies a high priority.

Gene expression in conifers has been studied for a very long time, beginning in the 1980s when methods to isolate and measure the quantity of mRNA first became available. The first technique to be widely used was the Northern blot technique where mRNA was bound to a membrane, hybridized with a radioactive DNA probe, and mRNA abundance measured by signal intensity. In fact, many of the papers described in the previous chapter on gene structure (Chap. 5) included a section on gene expression using the Northern blot technique. Only a brief mention of the gene

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_6

expression results of these papers was given there since the focus of Chap. 5 was on gene structure (DNA) and not expression (RNA). We will not try to summarize the results from the literature from the era where the Northern blot and other advanced techniques such as quantitative real-time PCR (qRT-PCR) were used since these techniques were applied to just one or a small number of genes in individual studies. This work has largely been eclipsed by modern high-throughput technologies which we will summarize.

The two primary techniques for high-throughput gene expression analysis are (1) microarrays and (2) RNA-seq. In an excellent review by Parent et al. (2015), a fairly complete review of the application of these techniques in forest trees is given. They discuss the limitations of the early microarray techniques (cDNA and oligo arrays) which have now been replaced by the RNA-seq technique. In this chapter we will complete and update this review specifically for conifers (Table 6.1). Furthermore, we will follow an organizational structure similar to that used by Parent et al. (2015) by giving sections on (1) wood formation, (2) biotic factors, (3) abiotic factors, and (4) seasonal patterns. Nearly all studies follow a very similar approach to finding genes that may be uniquely expressed, over-expressed, or under-expressed in a specific tissue, at a specific time point, or with a specific inductive condition. This requires comparing the mRNA pool in the sample of interest against a control. These contrasts are done in different ways among the studies shown in Table 6.1 and gene expression technologies also vary. Nevertheless, taken as a whole, a first approximation of the underlying role of gene expression in determining phenotypic differences begins to emerge. These types of gene expression studies will surely continue in the years ahead with much improved and comprehensive resources such as well-annotated reference genome sequences, very-high-throughput RNA-seq technologies, and more precision on mRNA sampling, such as specific cell types or even individual cells. In the final section of this chapter, we will briefly discuss epigenetic factors controlling gene expression, a new and exciting field of discovery.

Wood Formation

Forest geneticists and tree breeders have long been interested in the genetic basis of wood formation. In Chap. 11, we discuss how genetic mapping approaches have been used to discover genes underlying wood formation. In this chapter we review how gene expression studies have also been used to discover these genes. There are many physical and chemical wood property traits of interest. Examples include wood specific gravity, microfibril angle and length, stiffness, percent earlywood and latewood, compression wood, lignin content, and cellulose content. All of these traits are known to have a heritable component and are quantitatively inherited.

The study of gene expression in the transition from earlywood to latewood has been conducted in a number of species (Table 6.1). In a study that predated microarray and RNA-seq technologies, Le Provost et al. (2003) used the cDNA AFLP technique to find 100 differentially expressed genes (DEGs) between earlywood and

Table 6.1 Summary of conifer gene expression studies (MA is microarray)

Species	Method	Comparison	No. of analyzed genes	No. of differentially expressed genes (%)	Reference
Wood formation					
Pinus pinaster	cDNAAFLP	Compression wood sampled in time course after bending	5500 (fragments)	100 (2)	Le Provost et al. (2003)
Cryptomeria japonica	Oligo MA	Earlywood vs latewood	18,082	10,380 (57)	Mishima et al. (2014)
Cunninghamia lanceolata	RNA-seq	Cambial tissues at the active vs dormant stages	59,669	4415 (7.3)	Qiu et al. (2013)
		Cambial tissues at the active vs reactivating stages		883 (1.5)	
		Cambial tissues at the reactivating vs dormant stages		4018 (6.7)	
Picea glauca	cDNA MA	Xylem versus phloem and needles	10,400	360 (xylem specific)	Pavy et al. (2008)
	Oligo MA	Comparison of seven vegetative tissue types from aerial and below-ground organs	23,853	18,052 (76)	Raherison et al. (2015)
Picea sitchensis	cDNA MA	Cambium and xylem from high and low wood density trees	21,840	25	Stephenson et al. (2011)
Pinus pinaster	cDNA MA	Xylem at five time points within a growing season	3512	667 (19)	Paiva et al. (2008)
	cDNA MA	Compression vs normal wood	6841	496 (7.2)	Villalobos et al. (2012)
Pinus radiata	cDNA MA	Earlywood vs latewood at the juvenile stage (5 years)	3320	687 (21)	Li et al. (2011b)
		Earlywood vs latewood at the transition stage (9 years)		995 (30)	
		Earlywood vs latewood at the mature stage (30 years)		381 (12)	

(continued)

Table 6.1 (continued)

Species	Method	Comparison	No. of analyzed genes	No. of differentially expressed genes (%)	Reference
	cDNA MA	High-stiffness and low-stiffness wood	3320	481 (14)	Li et al. (2011a)
	cDNA MA	Spring and autumn earlywood and latewood	3320	319 (10)	Li et al. (2010)
	cDNA MA	Compression wood and opposite wood	3320	963 (29)	Li et al. (2013)
Pinus taeda	SAGE	Lignifying xylem	42,641	N/A	Lorenz and Dean (2002)
	cDNA MA	Earlywood vs latewood	2171	110 (5)	Yang and Loopstra (2005)
	cDNA MA	Earlywood vs latewood	350	71 (20)	Egertsdotter et al. (2004)
	RT-qPCR	500 genotypes	111	106 (95)	Palle et al. (2011)
Biotic factors					
Larix gmelinii	RNA-seq	Needles of control vs jasmonic acid-treated trees	51,157	2383 (4.7)	Men et al. (2013)
Picea glauca x *P. engelmannii*	cDNA MA	Barks of trees that are susceptible vs resistant to the white pine weevil (*Pissodes strobi*)	17,825	191 (1)	Verne et al. (2011)
Picea glauca	RNA-seq	Needles of resistant trees; uninfected vs infected with white pine-blister rust (WP-BR) (*Cronartium ribicola*)	23,000	789 (3.4)	Liu et al. (2013a)
	Oligo MA	Needles of trees that are susceptible vs resistant to the spruce budworm (*Choristoneura occidentalis*)	23,853	486 (2.1)	Mageroy et al. (2015)
Picea sitchensis	cDNA MA	Bark of control vs *Pissodes strobi* treated trees	9720	2382 (24.5)	Ralph et al. (2006)
		Bark of control vs mechanically wounded trees		3089 (31.8)	
		Shoot tips of control vs Western spruce budworm (*Choristoneura occidentalis*) treated trees		358 (3.7)	
		Shoot tips of control vs *Choristoneura occidentalis*-treated trees, 3 h post-treatment, 52 h post-treatment		3490 (35.9)	

	cDNA MA	Apical shoots with vs without removing bark	16,700	610 (0.4)	Friedmann et al. (2007)
Pinus radiata	*Pinus* oligo MA	Mucilaginous xylem of control vs ethephon-treated trees, 8 weeks post-treatment	175,614	23,084 (13)	Dubouzet et al. (2014)
Pinus sylvestris	cDNA MA	Roots of control vs saprotrophic fungus (*Trichoderma aureoviride*) inoculated, 15 days post-inoculation	2109	10 (0.5)	Adomas et al. (2008)
	cDNA MA	Roots infected with *Laccaria bicolor*	2109	236 (11)	Heller et al. (2008)
	RNA-454RT-qPCR	*Phlebiopsis gigantea*-treated seedlings	Many		Sun et al. (2011)
Thuja plicata	RT-qPCR	Foliar glands from wild type and mutants	33,202	604 (0.02)	Foster et al. (2013)
Abiotic factors					
Picea abies	RNA-seq	Embryonic callus generated at cold (18 °C) vs warm (30 °C) temperature	143,723	1608 (1.1)	Yakovlev et al. (2014)
Pinus contorta	RNA-seq	Needles of trees grown under seven treatments varying in temperature, humidity, and day length	23,889	11,658 (48.8)	Yeaman et al. (2014)
Pinus sylvestris	cDNA MA	Freeze tolerance in roots and shoots	1500	Many	Joosen et al. (2006)
	cDNA MA	Hypocotyls which were grown under continuous red vs far-red light	12,523	644 (5.1)	Ranade et al. (2013)
Pinus taeda	cDNA MA	Normal and drought-stressed roots	26,496	2445 (9)	Lorenz et al. (2011)
Seasonal and diurnal patterns					
Picea sitchensis	cDNA MA	Needle tissues in autumn during cold acclimation	21,840	2224 (10)	Holliday et al. (2008)
	Subtractive cDNA	Autumn bud tissue	550,000	11,121 (2.0)	Reid et al. (2013)
Pseudotsuga menziesii	RNA-seq	Daily cycle	41,382	12,042 (29)	Cronn et al. (2017)
		Seasonal cycle	36,145	21,225 (58.7)	

Modified and updated from Parent et al. (2015)

Table 6.2 Proportion of genes differentially expressed within the functional classes in earlywood and latewood ($P < 0.05$) in *Pinus taeda*

	% of differentially expressed genes in each group				
Functional class	Mar.–Oct.	Apr.–Oct.	Jun.–Oct.	Jul.–Oct.	Sep.–Oct.
Cytoskeleton	–	33	22	11	–
Gene replication, gene expression	17	39	28	11	6
Growth regulators	28	54	62	23	31
Lignin biosynthesis	22	35	26	26	17
Membrane	33	29	33	10	14
Others	18	32	21	25	4
Protein synthesis, protein degradation	31	44	19	25	–
Signal transduction	12	19	8	19	6
Stress related	71	71	71	43	57
Unknown	15	22	22	12	14
Wall carbohydrates	21	39	21	18	12
Wall protein	43	52	52	52	19

From Egertsdotter et al. (2004)

latewood and between compression wood and opposite wood in *Pinus pinaster*. Egertsdotter et al. (2004) used a cDNA microarray of 350 genes in *Pinus taeda* and found 71 differentially expressed genes between earlywood and latewood. These genes were further categorized by functional classes (Table 6.2). Many of these genes are involved in cell wall formation. Yang and Loopstra (2005) used a cDNA microarray with more genes (2171) and found 110 genes differentially expressed between earlywood and latewood in *P. taeda* (Table 6.3). Comparison of the lists of differentially expressed genes in these early studies reveals a small number of genes common to all studies, but for the most part they were different sets of genes in each study. This is likely a function of the small number of genes on microarrays at that time and the imprecision in gene annotations. In *Pinus radiata*, Li et al. (2011b) also used a 3320 gene cDNA microarray and found many genes differentially expressed between earlywood and latewood in trees of four different ages. They found 147 genes differentially expressed between juvenile wood and mature wood (Li et al. 2011a), 481 differentially expressed genes between high- and low-stiffness wood trees (Fig. 6.1), and 963 differentially expressed genes between compression wood and opposite wood (Fig. 6.2) with this same microarray. Like the *Pinus taeda* studies, many of these genes are involved in cell wall biosynthesis. In a much more recent study in *Cryptomeria japonica* using an oligo array of 18,380 genes, Mishima et al. (2014) found 10,380 differentially expressed genes. What proportion of the genes are functionally related to the transition from earlywood to latewood versus simply associated is not clear, but it is clear that this transition is complex in terms of the genes involved and the differential expression of these genes.

Table 6.3 Examples of genes expressed preferentially in latewood compared with earlywood in several *Pinus taeda* clones

	p-value	
Putative function	SAR	SLA
Putative dehydrin	<0.01 (2.5)	<0.01 (1.7)
AGP5	<0.01 (1.8)	<0.01 (1.1)
Glycine-rich protein	<0.01 (1.8)	<0.01 (0.1)
Adenylate kinase	<0.01 (1.8)	<0.01 (1.0)
No hit	<0.01 (1.7)	<0.01 (1.6)
Cellulose synthase-1	<0.01 (1.7)	<0.01 (1.3)
Fructose-bisphosphate aldolase	<0.01 (1.5)	<0.01 (1.0)
Tubulin alpha-1 chain	<0.01 (1.4)	<0.01 (1.0)
Lp6 protein – Loblolly pine	<0.01 (1.4)	<0.01 (0.9)
AGP5	<0.01 (1.2)	<0.01 (1.7)
Expansin9 precursor	<0.01 (1.1)	<0.01 (0.7)
Putative importin	<0.01 (1.1)	<0.01 (0.6)
Transcription factor Hap5a	<0.01 (1.1)	<0.01 (0.4)
Tubulin beta-2 chain	<0.01 (1.0)	<0.01 (0.7)
Cellulose synthase	<0.01 (1.0)	<0.01 (0.5)
Putative SF16 protein	<0.01 (1.0)	<0.01 (1.0)
1,4-Benzoquinone reductase	<0.01 (1.0)	<0.01 (0.7)
adoMet synthetase 2	<0.01 (0.9)	<0.01 (0.6)
Isoflavone reductase homolog	<0.01 (0.9)	<0.01 (0.6)
MADS-box protein AGL2	<0.01 (0.8)	<0.01 (0.9)
Pinoresinol-lariciresinol reductase	<0.01 (0.7)	<0.01 (0.3)
Phenylcoumaran benzylic ether reductase	<0.01 (0.6)	<0.01 (0.4)
Aluminum-induced protein	<0.01 (0.6)	<0.01 (0.2)
No hit	<0.01 (2.2)	
Sucrose synthase	<0.01 (1.7)	
Polyubiquitin	<0.01 (1.7)	
Beta tubulin	<0.01 (1.4)	
No hit	<0.01 (1.4)	
Cinnamyl-alcohol dehydrogenase	<0.01 (1.2)	
adoMet synthetase	<0.01 (1.2)	
Putative adoMet synthetase	<0.01 (1.1)	
No hit	<0.01 (1.1)	
14A9	<0.01 (1.0)	
Cellulose synthase	<0.01 (0.9)	
Trans-cinnamate 4-hydroxylase	<0.01 (0.4)	
Laccase		<0.01 (0.9)

From Yang and Loopstra (2005)
Abbreviations: *SAR* South Arkansas, and *SLA* South Louisiana
Significance levels and log ratios (values in parentheses) from the microarray analyses are given for each putative function

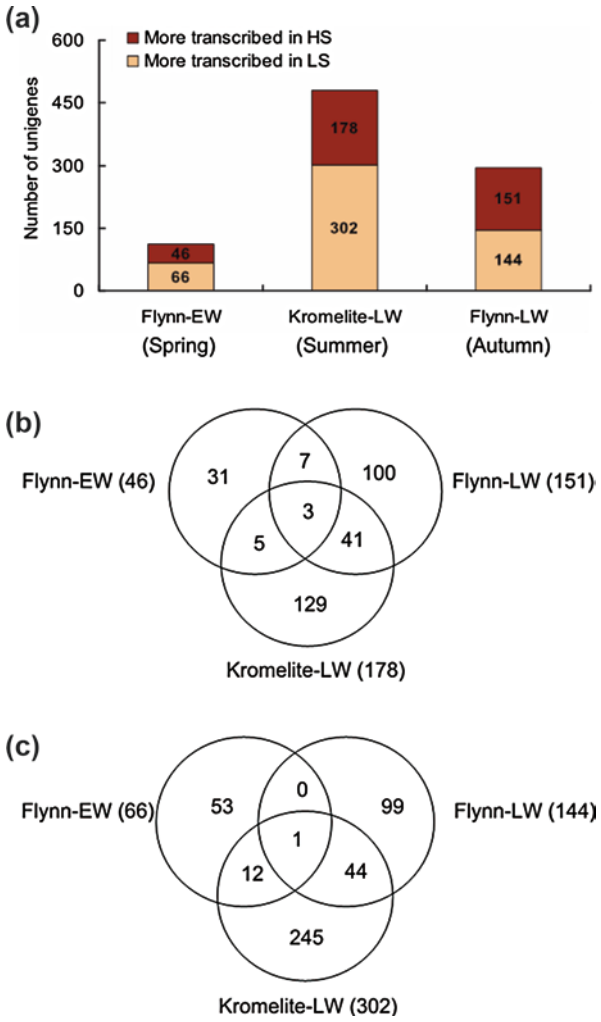


Fig. 6.1 Differentially transcribed unigenes identified in *Pinus radiata*. Unigenes differentially transcribed in wood with contrasting stiffness (HS high stiffness, LS low stiffness) and microfibril angle were identified using cDNA microarrays containing 3320 xylem unigenes. Comparisons made were (**a**) the number of unigenes identified at different sampling seasons in the two trials, (**b**) unigenes preferentially transcribed in the highest stiffness (lower microfibril angle), and (**c**) lowest stiffness wood (higher microfibril angle). Identified unigenes common to developing xylem tissues collected in different seasons are also indicated in the figures (EW earlywood, LW latewood). (From Li et al. 2011a)

Gene expression studies have been done to observe differences over a time course in developing xylem. Using a technology that predated microarrays and RNA-seq called SAGE profiling, Lorenz and Dean (2002) observed differences in expression between the upper and lower crown in a 10-year-old *Pinus taeda* tree (Fig. 6.3). Paiva et al. (2008) used a 3512 cDNA microarray to find 667

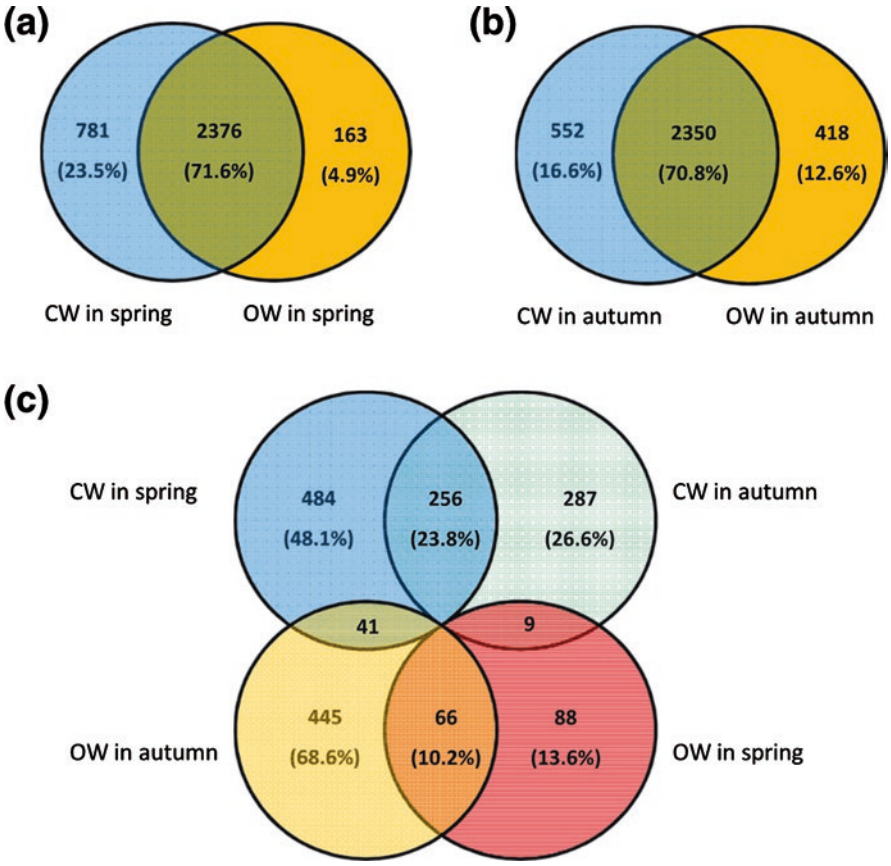


Fig. 6.2 Transcriptome comparisons between compression (CW) and opposite wood (OW) formed in branches in *Pinus radiata*. Genes differentially transcribed in CW and OW sampled in spring and autumn were identified using cDNA microarrays, respectively. Numbers of preferentially transcribed genes identified from developing xylem sampled in spring (**a**) and autumn (**b**) were present. Differentially transcribed genes were further compared between the two seasons. A number of genes showed consistently differential transcription in the two wood tissues across the two seasons (**c**). (From Li et al. 2013)

differentially expressed genes in developing xylem in *Pinus pinaster* (Table 6.4). Two studies in *Picea glauca* used gene expression studies to find genes uniquely expressed in xylem tissue. Pavy et al. (2008) used a 10,400 cDNA microarray to find 360 genes uniquely expressed in xylem (Table 6.5) whereas Raherison et al. (2015) found a very high proportion of genes (76%) that were unique to at least one tissue type, including xylem, across several tissue types of *Picea glauca* (Fig. 6.4). Gene expression has also been studied in developing cambial tissues. In one of the first reports using RNA-seq, Qiu et al. (2013) found many genes unique to different stages of cambial activity in *Cunninghamia lanceolata* (Fig. 6.5).

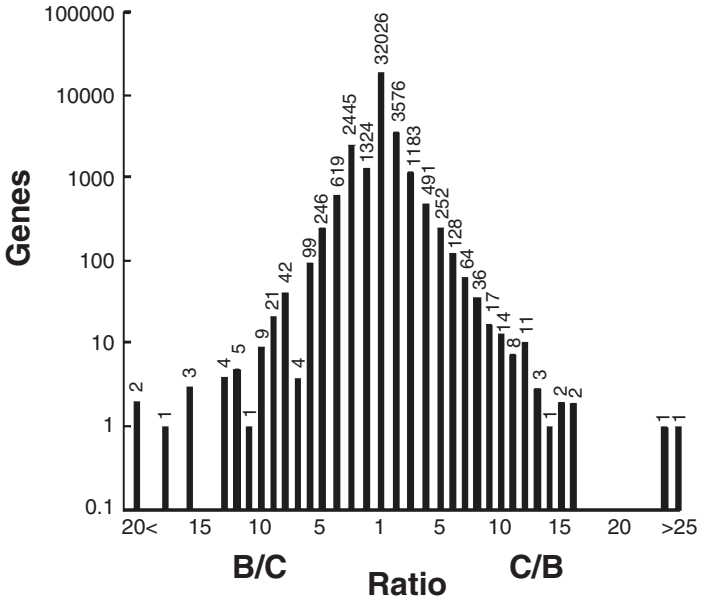


Fig. 6.3 Differential gene expression patterns along the axial gradient of lignifying xylem in *Pinus taeda* displayed as the ratio between crown (C) and base (B) library counts for each tag. The number of genes displaying each ratio is plotted against a log scale on the ordinate with the exact number displayed above each bar. (From Lorenz and Dean 2002)

Table 6.4 Number of differentially expressed genes in developing xylem of *Pinus ponderosa* according to their cluster and functional category

Functional category	% of EST spotted onto the array	Cluster					Total	G-test(*p*-value)
		1	2	3	4	5		
Communication/signal transduction	2.7	16	1	0	0	3	20	0.00
Cell division and growth	2.3	9	3	0	0	1	13	0.16
Protein fate	3.0	13	2	4	0	1	20	0.01
Energy	2.6	13	4	0	1	1	19	0.00
Metabolism	10.2	55	15	2	1	4	77	0.00
Cellular organization	2.1	10	3	1	1	1	16	0.58
Stress response	2.8	12	2	4	1	3	22	0.65
Protein synthesis	4.5	39	5	1	7	0	52	0.00
Intracellular traffic	1.5	10	1	0	1	0	12	0.01
Transcription	4.0	15	6	0	1	1	23	0.00
Transport	2.3	10	3	0	0	0	13	0.00
Not classified								
Putative protein	36.9	81	82	6	2	18	198	0.00
Unknown protein	21.6	67	52	3	3	9	134	0.00
Others	3.4	17	5	1	0	2	25	0.00
Total	**100**	**367**	**184**	**22**	**18**	**44**	**635**	

From Paiva et al. (2008)

Table 6.5 Molecular functions of the 360 xylem-preferential *Picea glauca* sequences annotated based on protein matches with gene databases

Function	Gene match (no.)
Unknown	118
Transcription factor	18
Cell wall	42
Kinase	27
Nodulin-like	8
Kinesin	7
Protein binding	12
Protease	9
Peroxidase	4
Oxidoreductase	7
Transporter	11
DNA binding	3
Lipid binding	3
Catalytic activity	7
Transferase	9
Hydrolase	5
Calmodulin binding	4
Others	66
Total	**360**

From Pavy et al. (2008)

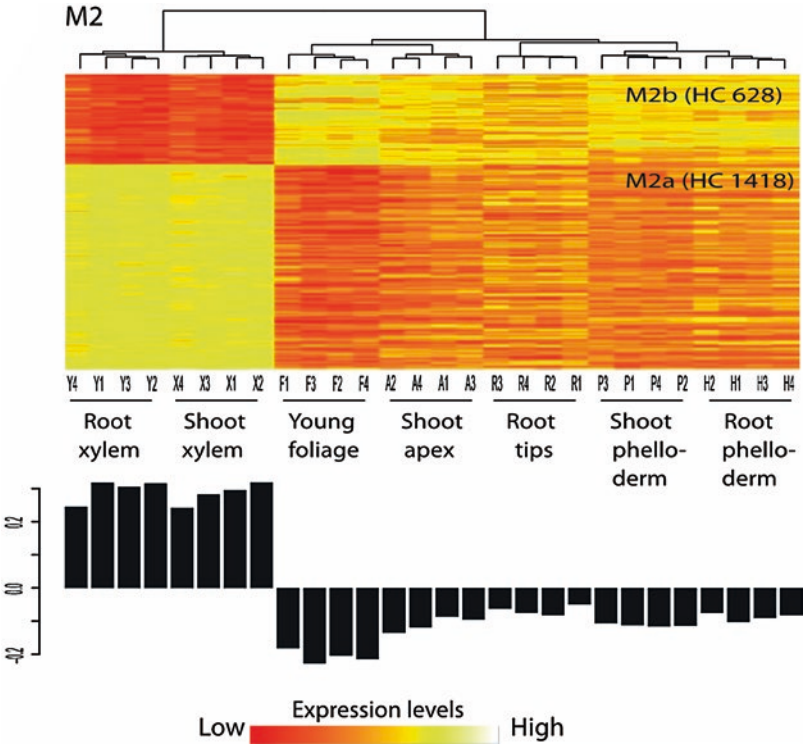


Fig. 6.4 Relative patterns in gene expression across different tissue types in *Picea glauca* as shown with heat maps. (From Raherison et al. 2015)

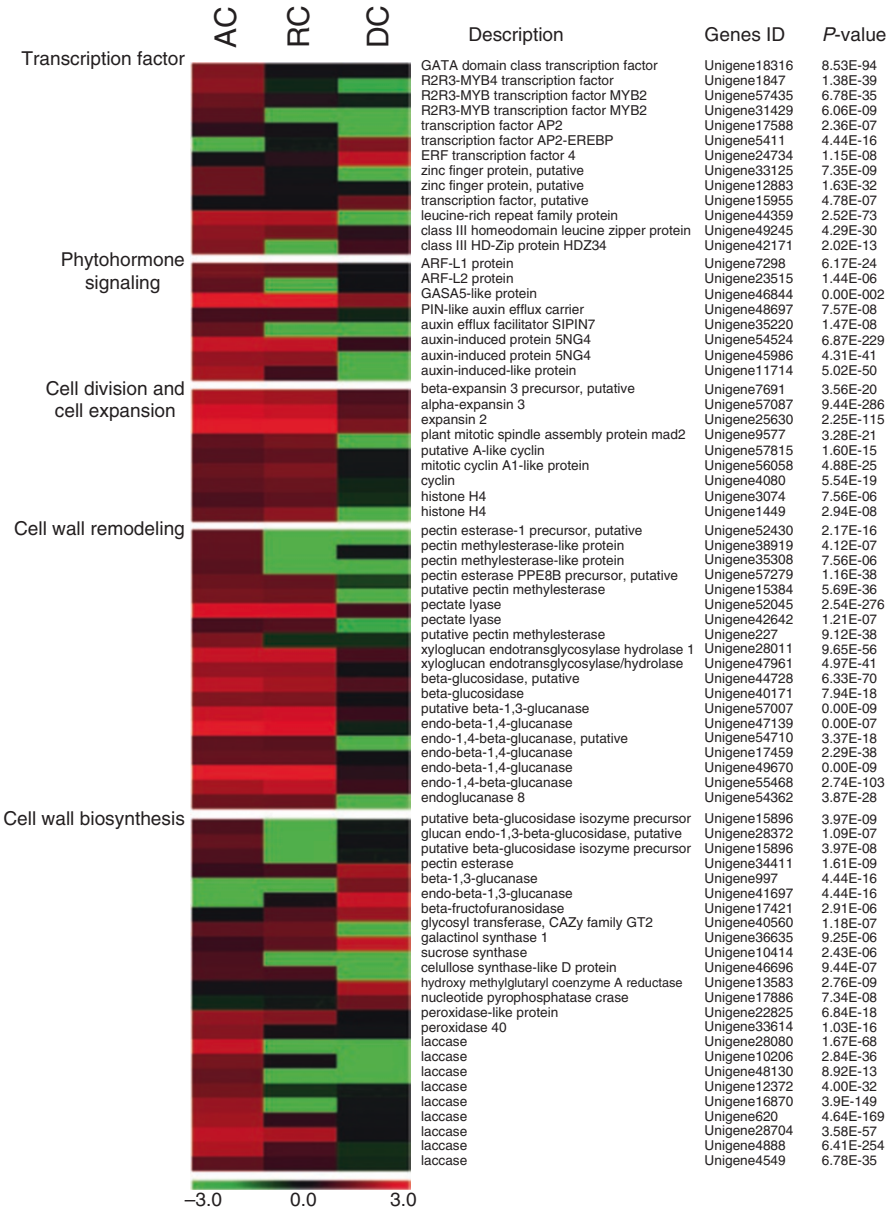


Fig. 6.5 Expression profile of 122 differentially expressed genes involved in transcription regulation, phytohormone signaling, cell division and cell expansion, cell wall remodeling, cell wall biosynthesis, defense/stress response, transport, kinase, cytoskeleton, and others in the dormant, reactivating, and active cambia of *Cunninghamia lanceolata*. (From Qiu et al. 2013)

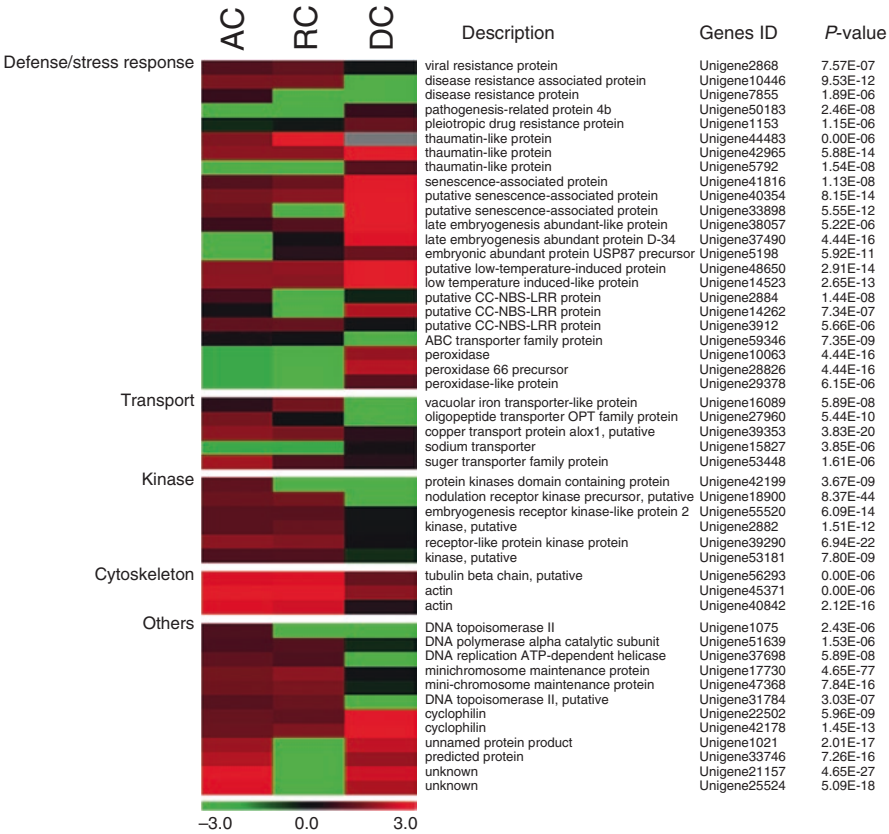


Fig. 6.5 (continued)

Other studies have investigated the role of gene expression in determining important physical properties of wood. Stephenson et al. (2011) used a 21,840 cDNA microarray to find a small number of genes differentially expressed between high and low specific gravity wood in *Picea sitchensis* (Table 6.6). In *Pinus pinaster*, Villalobos et al. (2012) found 496 genes differentially expressed between compression wood and normal wood using a 6841 cDNA microarray (Fig. 6.6).

Finally, expression differences in a suite of 111 wood-property candidate genes among 600 unique *Pinus taeda* genotypes were tested using qRT-PCR (Palle et al. 2011). Nearly all the genes (106) had some amount of expression difference among some of the genotypes (Fig. 6.7). This study is important because unlike the other studies where only one or a small number of unique genotypes were used, this study demonstrated that there is significant population-level variation in gene expression and that this trait could potentially be used in breeding.

Table 6.6 Gene ontology (GO) functional categories significantly over-represented in differentially expressed genes from three microarray experiments with *Picea sitchensis*

Analysis	Upregulated	Downregulated	Up- and downregulated
Exp. 1	None	Lipid metabolism	Lipid metabolism
Exp. 2	None	None	None
Exp. 3	Jacalin lectin myrosinases	Peroxidases	Receptor kinases
Combined	Protein degradation	Flavonol synthesis	Receptor kinase signaling
	Jacalin lectin myrosinases		Cell organization
	Cell wall synthesis/modification		Jacalin lectin myrosinases
	Metal handling		Glycolysis

From Stephenson et al. (2011)

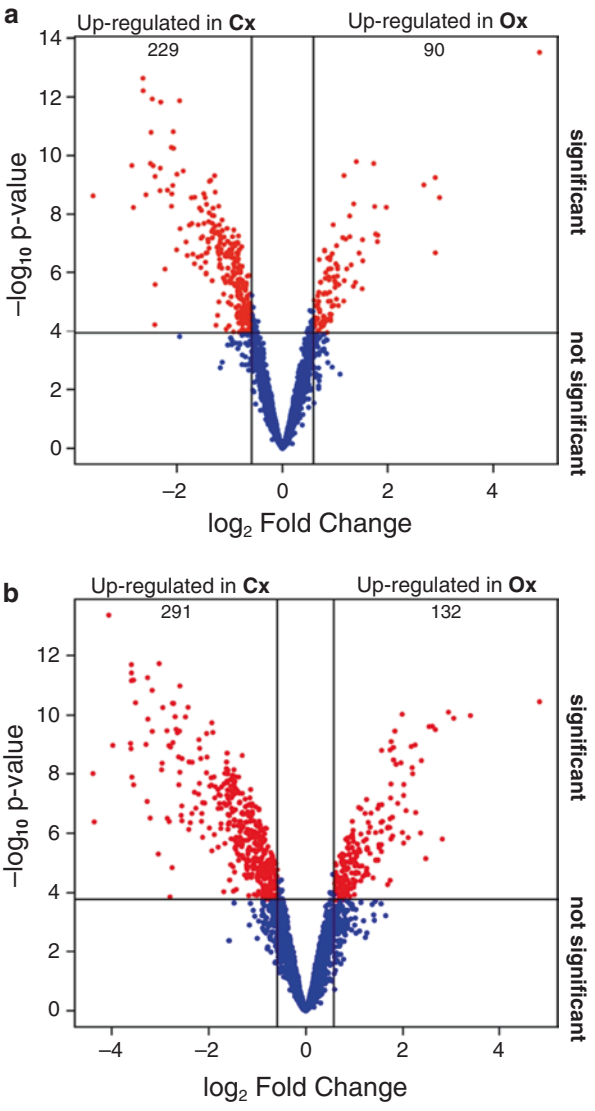


Fig. 6.6 Volcano plots of microarray analyses to identify genes differentially expressed during compression and opposite wood formation in *Pinus pinaster*. (**a**) Results from the analysis of microarray 1 constructed with cDNA clones from the composite library. (**b**) Results from the analysis of microarray 2 constructed with cDNA clones from subtractive libraries. (From Villalobos et al. 2012)

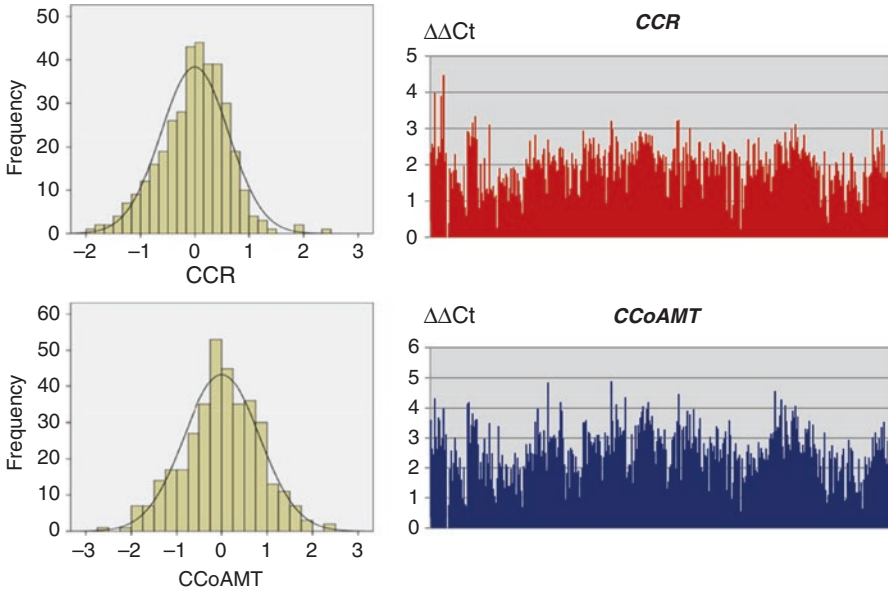


Fig. 6.7 Normal distribution plots and barplots showing the range of ΔΔCt values for the CCR and CCoAMT genes among different *Pinus taeda* clones in the population. The mean is zero for the normal distribution plots. The ΔΔCt of the clone with the highest expression was considered zero when constructing the barplots. (From Palle et al. 2011)

In summary, and as we will conclude throughout this chapter, gene expression studies in wood formation reveal large differences in the specific genes and the levels of expression of genes among developmental states. This complexity revealed in just one or a small number of trees adds to the complexity associated with genetic variation among trees (Chap. 11).

Biotic Factors

High-throughput gene expression studies have also been conducted to discover genes induced by exposure to diseases (primarily fungal) and insects. The 2109 cDNA microarray constructed from *Pinus taeda* ESTs was used in two studies with *Pinus sylvestris* roots exposed to two different fungal pathogens. Adomas et al. (2008) found just ten genes with induced expression in root tissues when inoculated with the saprotrophic fungus *Trichoderma aureoviride* (Fig. 6.8), whereas Heller et al. (2008) found 236 genes induced when inoculated with the pathogenic fungus *Laccaria bicolor*. Further, there were no induced genes in common suggesting different response mechanisms to these different classes of fungi. This study also established that a microarray constructed from cDNAs of one pine species could be used with another related pine species. In a more recent study with the saprotrophic

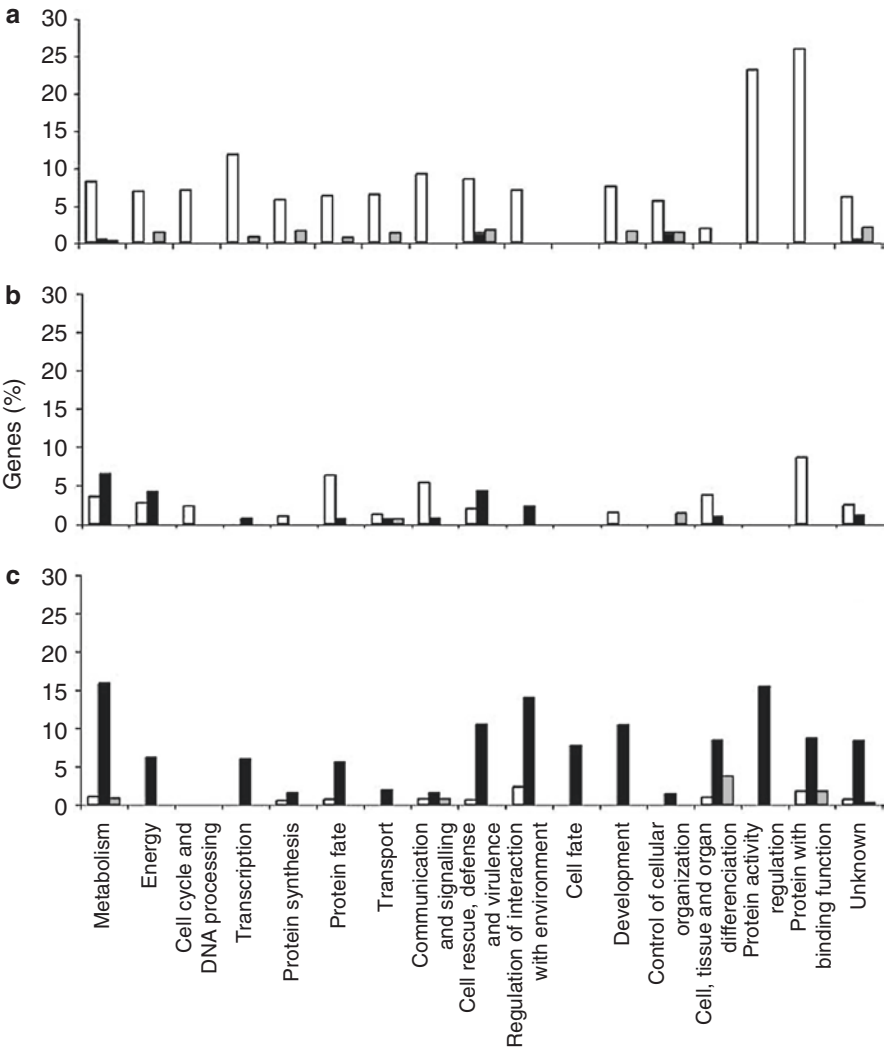


Fig. 6.8 Percentage of genes, by functional category, upregulated in *Pinus sylvestris* in response to challenge with a pathogen (*Heterobasidion annosum*; black bars), an ectomycorrhizal symbiont (*Laccaria bicolor*; white bars), or a saprotroph (*Trichoderma aureoviride*; gray bars) at (**a**) 1, (**b**) 5, and (**c**) 15 days post-inoculation. (From Adomas et al. 2008)

fungus *Phlebiopsis gigantea*, Sun et al. (2011) found a large number of genes were upregulated following exposure to the fungus.

RNA-seq was used to discover differences between resistant and susceptible *Pinus monticola* genotypes to the white pine-blister rust pathogen *Cronartium ribicola* (Liu et al. 2013a). A total of 979 differentially expressed genes were detected (Fig. 6.9). Functional classification of these genes revealed both enzyme-coding and regulatory (transcription factor) genes.

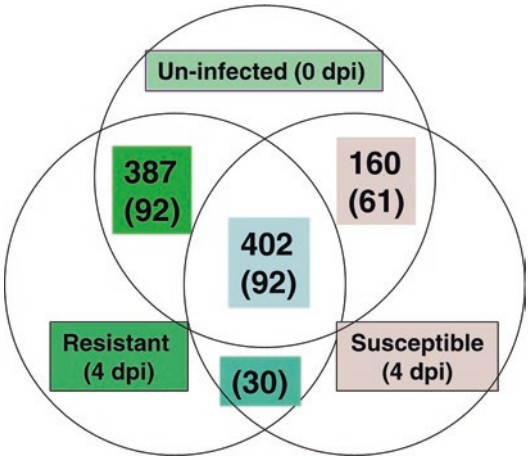


Fig. 6.9 Venn diagrams of differentially expressed genes (DEGs) in white pine-blister rust (WP-BR) interactions for illustrating the relationship of DEGs in two Western white pine genotypes (*Cr2*/− and *cr2*/*cr2*) post rust infection. The numbers of DEGs detected between resistant (*Cr2*/−) and susceptible (*cr2*/*cr2*) seedlings are shown in parentheses. (From Liu et al. 2013a)

Gene expression studies have also been successfully used to discover genes expressed in response to insect attack. Friedmann et al. (2007) first examined gene expression in developing *Picea sitchensis* shoots with a 16,700 cDNA microarray. They found 610 genes differentially expressed between the vegetative shoot tips and the woody basal region, but also many defense-related genes in the apical shoot tissues that might be responsible for constitutive defense mechanisms. In a series of studies that followed, gene expression differences were studied in resistant versus susceptible trees to insect attack. Ralph et al. (2006) constructed a 9720 cDNA microarray in *Picea sitchensis* and found a large number of differentially expressed genes in trees exposed to spruce budworm (*Choristoneura occidentalis*) and also to white pine weevil (*Pissodes strobi*). There was considerable overlap in the genes with induced expression to the two insects (Fig. 6.10). Mageroy et al. (2015) produced a 25,853-feature oligo microarray and found 486 differentially expressed genes between *Picea glauca* trees resistant and susceptible to spruce budworm attack. One of these genes, which was then functionally identified as a beta-glucosidase gene involved in acetophenone biosynthesis, showed a 773x level difference in expression, demonstrating its importance in the resistance response (Table 6.7). Verne et al. (2011) also studied gene expression differences in white pine weevil resistant versus susceptible trees but this time in interior spruce (*Picea glauca* x *engelmanii*). They used a 17,825 cDNA microarray and found 191 differentially expressed genes (Table 6.8).

Predation due to browsing can be a big problem in reforestation after harvesting. Some species, such as *Thuja plicata*, are generally not affected by disease or insects but can be heavily browsed. Foster et al. (2013) compared gene expression differences between trees with foliar glands containing terpenoids versus a natural mutant without glands. Using an RNA-seq approach they found 604 of a total of 33,207 genes that were unique to the trees with foliar glands, suggesting that the expression of these genes might underlie resistance to browsing. In fact, among the differentially expressed genes observed by RNA-seq were a terpene synthase (Foster et al.

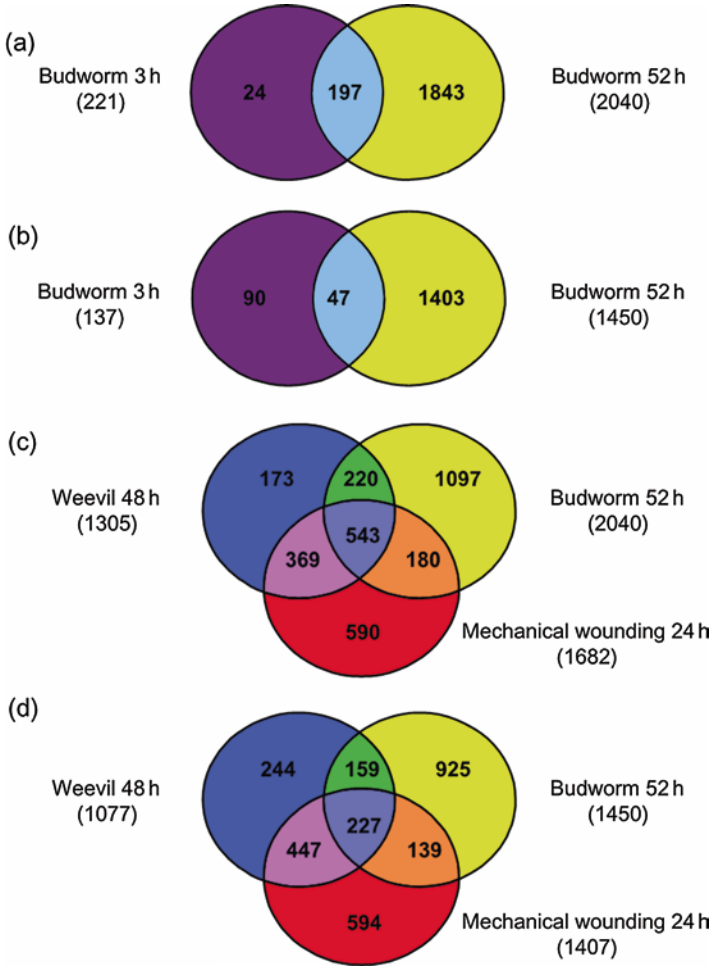


Fig. 6.10 Overall changes of gene expression induced by weevils, budworms, or mechanical wounding in *Picea sitchensis*. Venn diagrams showing distinct and overlapping patterns of genes induced or repressed by weevil feeding (48 h), budworm feeding (3 and 52 h), or mechanical wounding (24 h). (**a** and **b**) Intersection of genes that were upregulated (**a**) or downregulated (**b**) between budworm feeding after 3 and 52 h. (**c** and **d**) Intersection of genes that were upregulated (**c**) or downregulated (**d**) between weevil feeding, budworm feeding (52 h), and mechanical wounding. (From Ralph et al. 2006)

2013) and a cytochrome P450 (Gesell et al. 2015) with functions shown in the biosynthesis of thujone, a metabolite that deters deer browsing.

Application of compounds that are known to stimulate defense or resistance responses has also been used in gene expression profiling experiments. Men et al. (2013) compared control needles with those following application of jasmonate in *Larix gmelini* using RNA-seq to find a large number of differentially expressed

Table 6.7 Experimental system and expression of the *Pgbglu-1* gene in *Picea glauca*. Genes with largest differences in transcript abundance in the foliage of resistant ($n = 7$) and nonresistant trees ($n = 7$) identified using an oligonucleotide microarray

Overexpressed genes	GenBank accession no.	Predicted function	Fold change
Resistant trees	BT114253	Beta glucosidase 40	773
	BT108582	NA	23
	BT111304	No exine formation 1	15
	DR554713	3-oxo-5-alpha-steroid 4-dehydrogenase protein	15
	BT117306	NA	13
Nonresistant trees	C0236779	NA	18
	DR591433	NA	14
	BT106714	NA	10
	EX432133	Subtilisin-like serine endopeptidase protein	9
	BT103518	UBX domain-containing protein	7

From Mageroy et al. (2015)

Table 6.8 Summary of *t*-test comparisons between resistant and susceptible trees in *Picea glauca* x *engelmanii*

18,725 analyzed genes	Upregulated	Downregulated
Genes with p value <0.05	1225	1274
Genes with FC > 1.5	60	151
Maximum FC	2.24	3.91
Significant genes	54	137

From Verne et al. (2011)

genes. In *Pinus radiata*, Dubouzet et al. (2014) found many genes upregulated following Ethephon application to seedlings using an oligo microarray.

What becomes quite clear from all these gene-profiling experiments related to disease and insect defense response is that the mechanisms of defense are very complicated and 100s, if not 1000s, of genes are involved. In time, comparative genomic studies will likely reveal which sets of genes and pathways might be common to many different defenses and which may be unique to specific disease or insects.

Abiotic Factors

The relationship between a tree's genotype and abiotic factors (moisture, temperature, light, soil type, etc.) is a topic we return to throughout this book (see Chaps. 8, 10, 11, and 12). Not surprisingly, gene expression studies have also been used to discover genes responding to abiotic factors in the environment. Lorenz et al. (2011) used a 26,496 cDNA microarray in *Pinus taeda* to discover 2445 genes

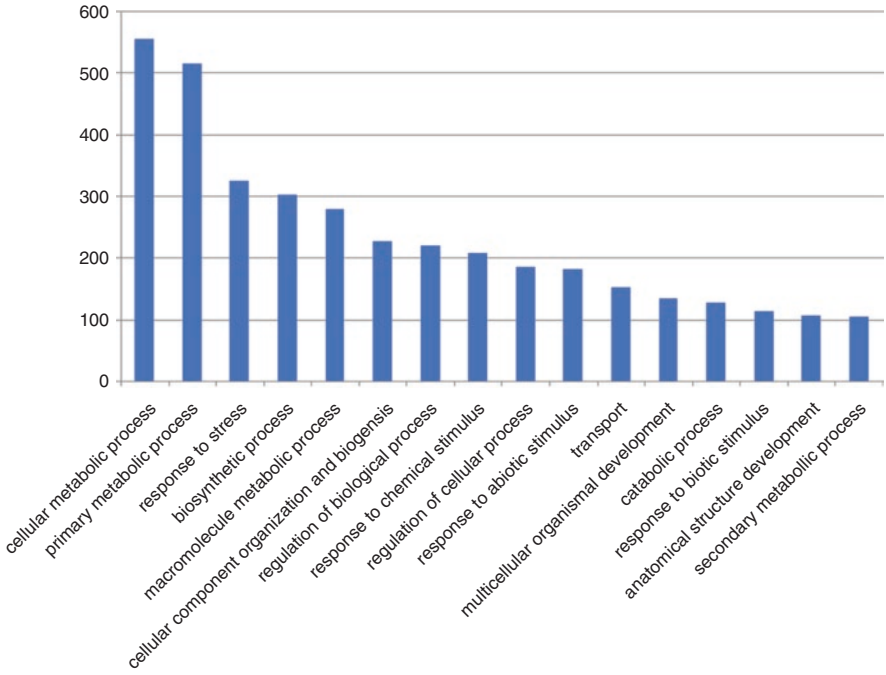


Fig. 6.11 Gene ontology (GO) analysis following a *Pinus taeda* gene expression study. Comparison of GO terms identified from the 2445 differentially expressed genes identified in the well-watered and drought-stressed comparison. (From Lorenz et al. 2011)

differentially expressed between control and drought-stressed roots (Fig. 6.11). It can be seen that these genes come from a diverse array of biological functions suggesting diverse and complex mechanisms for response to drought. Response to both high and low temperatures has been studied. Joosen et al. (2006) used a 1500 cDNA microarray in *Pinus sylvestris* and found 49 differentially expressed genes in response to cold acclimation in roots and shoots. Many of these genes had previously been implicated in different stress responses (Table 6.9). Yakovlev et al. (2014) performed an RNA-seq experiment in *Picea abies* embryogenic cultures grown under warm and cold temperatures and found 1608 differentially expressed genes (Fig. 6.12). The effect of light treatment has been studied in *Pinus sylvestris*. Ranade et al. (2013) discovered 644 genes differentially expressed between continuous red-light and continuous far-red-light treated seedlings using a 12,523 cDNA microarray. These genes are involved in a variety of biological processes (Fig. 6.13). Finally, in a very comprehensive RNA-seq experiment to discover genes responding to temperature, moisture, and light in both *Pinus contorta* and *Picea glauca* x *engelmanni*, Yeaman et al. (2014) found nearly half the genes were differentially expressed following treatments (Fig. 6.14). This study clearly illustrates

Table 6.9 Genes potentially involved in cold hardiness in *Pinus sylvestris* identified by gene expression studies

Putative function	Functional category
LP-3 like A	Abiotic stress
LP-3 like B	
LP 3 like C	
LP-3 like D	
Aluminum-induced protein	
Dehydrin dhn3 or dhn 7	
Dehydrin dhn5	
Dehydrin like	
Class IV chitinase	Biotic stress
Thaumatin-like	
Defensin (gamma thionin family)	
Intracellular PR protein	
PR10 protein	
Beta-1,3-glucanase	
LEA protein	Development/stress
LEA protein	
LEA protein	
Metallothionein	
Purple acid phosphatase	Secondary metabolism/stress
Pinosylvin synthase	
Malate dehydrogenase	Primary metabolism
Epoxide hydrolase	
Sucrose synthase	
Transcription factor E2F	Transcription
CAAT-binding factor subunit	
ELIP	Photosynthesis
Set proteinase inhibitor	Protein degradation
Proline-rich protein	Cell wall
P. taeda EST set C_CF388200	Unknown
No BLAST match, contig 52	
No BLAST match, contig 50	
P. taeda EST set C_CF472479	
Acid phosphatase	Primary metabolism
Xyloglucan endotransglycosylase	Cell wall
Calmodulin	Signal transduction
Alpha tubulin 1	Structure
Aquaporin	Transport
14-3-3 protein	Regulators
P. taeda EST CF668373	Unknown
P. taeda EST BE241143	
P. taeda EST CF389725	
Ubiquitin	Protein degradation
Ubiquitin/ribosomal protein 27a	
60S ribosomal protein L14	Translation
40S ribosomal protein S5	

(continued)

Table 6.9 (continued)

Putative function	Functional category
Ubiquitin/ribosomal protein L40	
Ribosomal protein S7	
Ribosomal S26	
Ribosomal S11	

From Joosen et al. (2006)

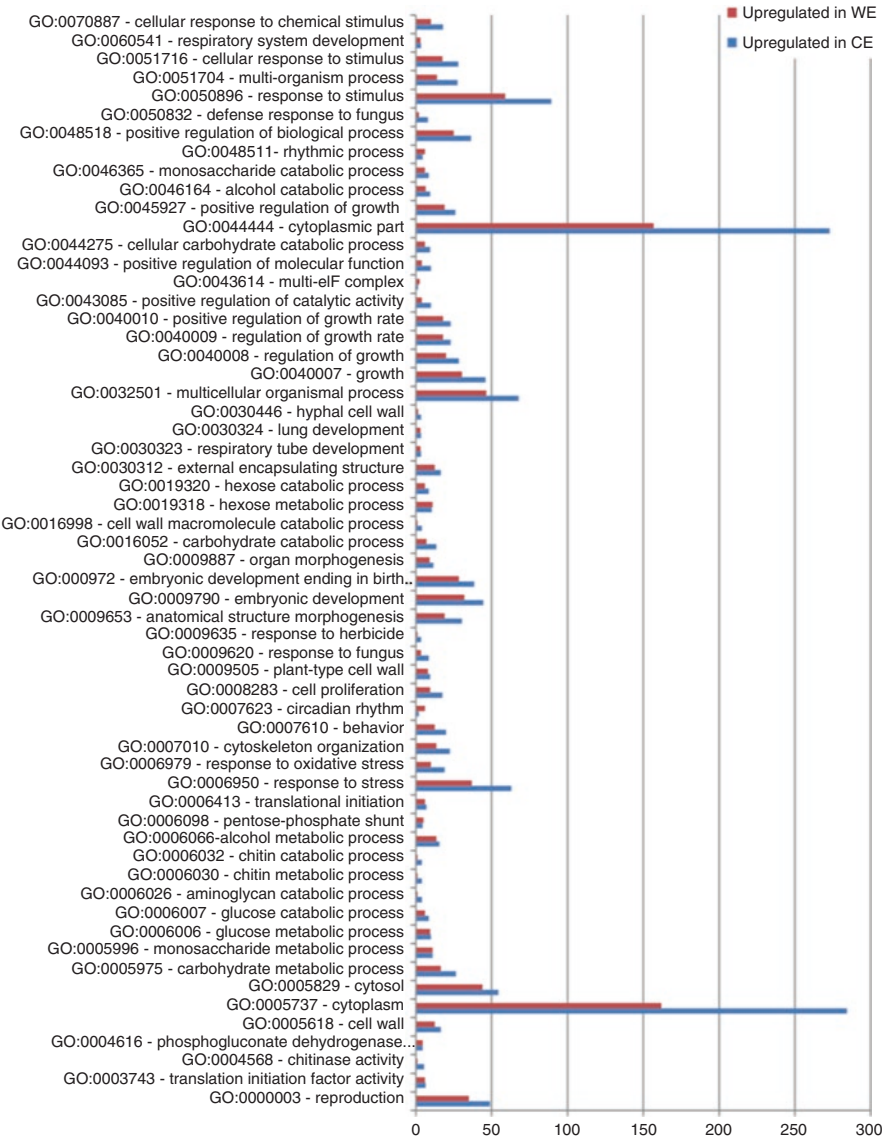


Fig. 6.12 List of the gene ontology categories most enriched in differentially expressed unique transcripts during embryogenesis in *Picea abies* grown under 18 °C cold exposure (CE) or 30 °C warm exposure (WE) epitype-inducing embryogenesis conditions. (From Yakovlev et al. 2014)

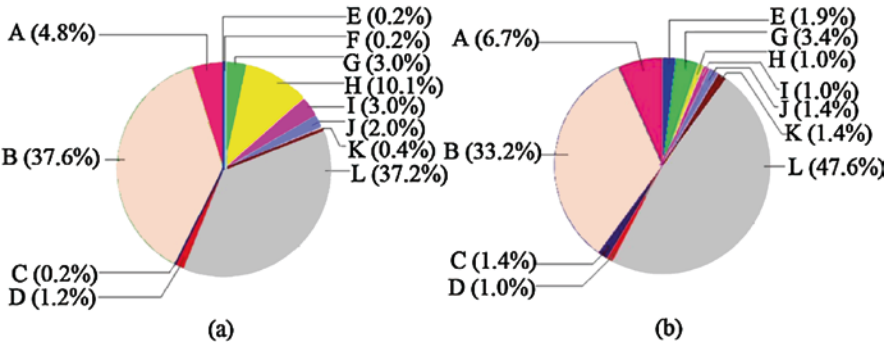


Fig. 6.13 Pie charts showing proportion of expressed genes in *Pinus sylvestris* of different biological processes following light treatments. (**a**) Higher expression under cR light; (**b**) higher expression under cFR light. A: multicellular organism process. B: Localization. C: Multi-organism process. D: Locomotion. E: Growth. F: Reproduction. G: Biological regulation. H: Developmental process. I: Anatomical structure formation. J: Cellular process. K: Immune system process. L: Response to stimulus. (From Ranade et al. 2013)

		CW	HD	MD	MW18	MW	MWbs	MWh
CW	P		4745	1623	379	853	3743	2991
	S		4336	1912	338	1058	3497	2788
HD		2051		1201	1906	1766	1407	961
		2739		1330	2051	1781	1678	915
MD		527	103		0	1	1012	699
		754	45		0	1	1514	222
MW18		119	663	0		3	1282	988
		300	234	0		0	1770	559
MW		376	146	2	1		1311	794
		491	77	1	1		1836	269
MWbs		951	121	59	143	55		1755
		1150	124	77	176	141		1404
MWh		1590	564	190	236	130	371	
		2245	375	234	447	192	252	

Fig. 6.14 Number of genes with differential expression in pairwise contrasts among treatments in *Pinus contorta* (above diagonal) and hybrid *Picea* (below diagonal), at a false discovery rate (FDR) < 0.01. Upper numbers in each cell indicate the number of genes that are upregulated in the treatment listed in the column; lower numbers indicate the number of genes that are downregulated for the treatment listed in the column. Darker shades of orange indicate greater numbers of differentially expressed genes. Experimental treatments are indicated by CW, cold wet; HD, hot dry; MD, mild dry; MW18, mild wet 18 h; MW, mild wet; MWbs, mild wet budset; and MWh, mild wet heat. (From Yeaman et al. 2014)

the very complex role of gene expression in determining phenotypic differences but in some ways also reveals the limitation of gene-expression profiling in ultimately getting to the root and fundamental causes of phenotypic variation.

Seasonal Patterns

All organisms that are exposed to daily light and dark cycles show diurnal patterns of gene expression. In Chap. 5 we introduced some of the genes whose expression is light-regulated. In perennial plants, seasonal patterns of gene expression are of great interest since they are often controlled by long-day and short-day cycles. Understanding how the expression of these genes relates to the onset of dormancy, hardening, and breaking of dormancy has motivated a small number of gene expression studies over seasonal cycles in conifers. In a first study of this type, Holliday et al. (2008) observed 1257 upregulated and 967 downregulated genes from a total of 21,840 genes expressed between late summer and early winter in seedlings of *Pinus sitchensis*. Many of these differentially regulated genes were the same as those observed in response to temperature and moisture stress in other studies (see section on abiotic factors). Reid et al. (2013) used a subtractive approach to find ESTs unique to autumn bud tissues in *Picea sitchensis* versus the complete *Picea* EST catalog which is comprised from ESTs of many different tissues. A surprisingly large number of unigenes, 11,121, were unique to autumn buds. In the most comprehensive study to date, Cronn et al. (2017) used RNA-seq in *Pseudotsuga menziesii* to discover genes unique to diurnal and seasonal cycles. They found 29% of the genes were expressed in a daily cycle and 58.7% were expressed over an annual cycle. These results illustrate the importance of light/dark cycles and day length in determining patterns of gene expression.

Epigenetic Control of Gene Expression

Epigenetic control of gene expression is a fairly new and intriguing area of research in all of biology and has just barely begun in conifers. Some very early work in Norway in *Picea abies*, labelled as *conditioning*, was the precursor to recent genomic research in epigenetics (Kohmann and Johnsen 1994; Johnsen et al. 1996, 2005). Epigenetic mechanisms include DNA methylation, histone modification, and small RNAs. Yakovlev et al. (2011) followed these earlier observations based on whole plant phenotypes with a study of differential gene expression in *Picea abies* seedlings that were derived from seed that had been given cold versus warm treatment under long- and short-day conditions. The idea was to see if there was an epigenetic memory of these conditions in embryos that followed through to seedlings. Comparison of the transcriptomes of seedlings whose seed was cold versus warm treated revealed differences in several coding genes and small RNAs (Fig. 6.15). Although correlative, these data suggest an epigenetic mechanism affecting gene expression.

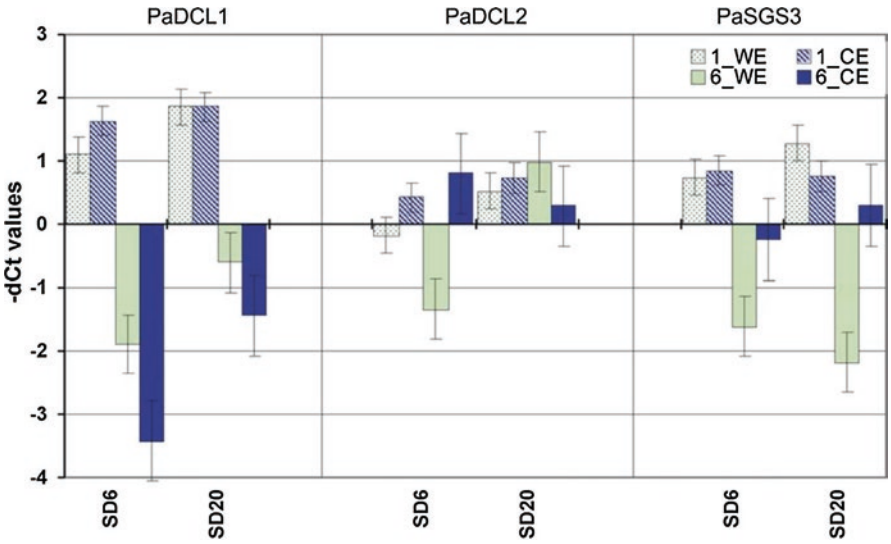


Fig. 6.15 Transcript profiles of selected *Picea abies* gene homologs from miRNA and ta-siRNA processing pathways in plants from families 1 and 6 with low and high epigenetic memory response after cold (CE cold exposure) and warm (WE warm exposure) treatment at days 6 and 20 under short-day (SD) photoperiodic conditions. (From Yakovlev et al. 2011)

The role of DNA methylation in accounting for epigenetic control has only been addressed in a few studies. Greenwood et al. (1989) did not find any differences in DNA methylation between juvenile and mature *Larix laricina* seedlings in an age maturation study. However, this study was done at a time of very limited genomics resources and tools and the power to detect DNA methylation differences was rather limited. Fraga et al. (2002a, b) did however observe DNA methylation differences between juvenile and mature meristematic tissues in *Pinus radiata.* In a more recent study in *Sequoia sempervirens*, Huang et al. (2012b) observed a greater amount of DNA methylation in adult versus juvenile and rejuvenated shoots (Fig. 6.16). This observation of greater methylation in mature tissues is consistent with much of the work in angiosperms.

Small RNAs are noncoding RNAs that are able to bind mRNA and restrict translation, thus affecting gene expression. Three classes of small RNAs are known, small-interfering RNAs (siRNAs), micro-RNAs (miRNAs), and Piwi-associated RNAs. Small RNAs were first discovered in *Caenorhabditus elegans* and later in many animal and plant (angiosperm) systems. The first report of small RNAs in conifers was that of Morin et al. (2008) who showed that *Pinus contorta* contained 21-nt small RNAs (miRNAs) but not siRNAs (Fig. 6.17). This finding was confirmed in additional work in *Pinus contorta* (Dolgosheina et al. 2008). Subsequently, Wan et al. (2012b) reported large and diverse families of miRNAs in *Pinus densata*. Wan et al. (2012a) also showed the presence of miRNAs in *Cunninghamia lanceolata* and these miRNAs bind some mRNAs. In addition, they also showed the presence of siRNAs

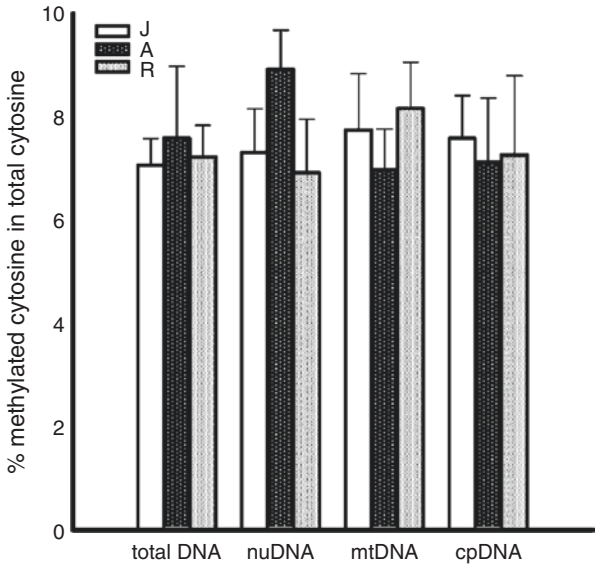


Fig. 6.16 Global DNA methylation in total, mitochondrial (mt), nuclear (nu), and chloroplast (cp) DNA from juvenile (J), adult (A), and rejuvenated shoots (R) in *Sequoia semervirens*. Percentages of methylcytosine were calculated by measuring the ratio of methylcytosine to total cytosine on HPLC. Results were calculated from three measurements (+SD). (From Huang et al. 2012b)

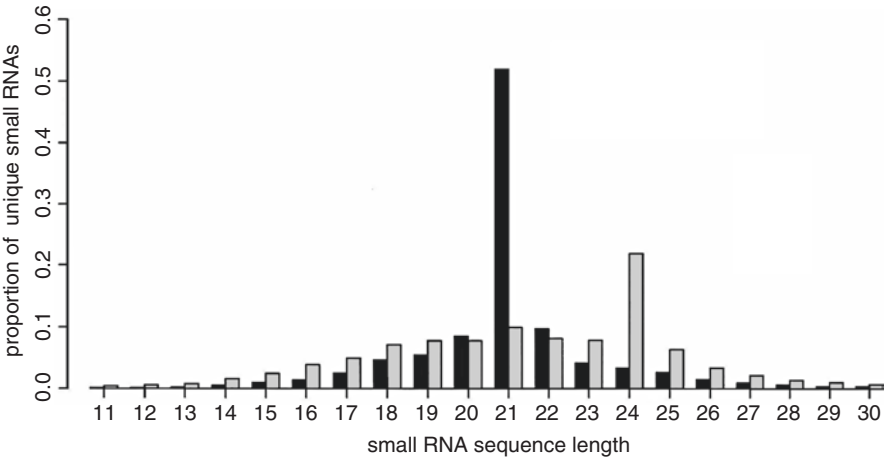


Fig. 6.17 Lengths of unique small RNA sequences from *Pinus contorta* (black bars, 58,466 sequences) and *Oryza sativa* (gray bars, 8615 sequences). The bulk of *P. contorta* small RNAs are 21 nt long with low variance ($\sigma = 8.1$). The *O. sativa* sequences have a major peak at 24 nt and a minor peak at 21 nt, yielding a median of 22 nt and a higher variance ($\sigma = 20.0$). (Morin et al. 2008)

in *Cunninghamia lanceolata* which were thought to be absent in conifers based on the earlier work in pine. In a recent study in *Larix leptolepis*, Zhang et al. (2013) confirmed the presence of 24-nt small RNAs (miRNAs). Although the research in small RNAs is just beginning in conifers, it is clear that these RNAs exist and can regulate gene expression. What fundamental differences in the types and mechanisms of controlling gene expression in conifers are yet to be determined.

Summary

The study of the expression of genes in conifers has been a very active area of research. Early studies involved just one or a small number of genes generally using the Northern blot technique which was difficult to perform and unable to generate accurate measures of gene expression. In the early 2000s, microarray technology became available whereby the expression of large numbers of genes could be measured in parallel. The microarray technique was soon followed by the RNA-seq approach which remains the primary approach for gene expression studies to this date. Gene expression studies were often focused on identifying genes underlying whole plant phenotypes such as wood formation, response to biotic factors, and response to abiotic factors and seasonal changes. Differentially expressed genes found in individual studies become strong candidates for determining differences in whole-plant phenotypes. The limitation of this body of work however lies in the fact that very little has been replicated and powerful database and bioinformatics tools are lacking to compare results across studies. Until such tools are developed and used in comprehensive comparative studies, the results from all primary studies remain quite preliminary and without clear application.

Proteomics and Metabolomics

7

A Short History of Proteomic and Metabolomic Research in Conifers

In this chapter we move to the final stage of the Central Dogma and beyond. Having previously discussed the *genome* (Chaps. 2, 3, 4, and 5) and the *transcriptome* (Chap. 6), we will now present the *proteome* and the *metabolome*. It has only been in the last 20 years that technologies have been developed to capture, characterize, and quantify the complete pool of proteins and metabolites from an animal or plant tissue sample. Prior to that, individual proteins and metabolites generally had to be assayed one at a time. This work is known as protein or metabolite profiling.

Just like the transcriptome, the proteome and metabolome make up the primary building blocks which work in complex biochemical pathways ultimately leading to the development of a tree and the nearly infinite variation we see among whole-tree phenotypes. Understanding which suite of genes (*genome*) underlie all this complex phenotypic variation is a daunting task. Therefore, atomizing the phenotypic variation down to discreet units such as the *transcriptome, proteome,* and *metabolome* can make it easier to identify such genes. Of course, all this reductionist research must then be synthesized and reassembled to completely understand the genetic control of whole-tree phenotypic variation. It is for these reasons that there has been such interest in proteomics and metabolomics in conifers.

As we did with the conifer literature on transcriptomics, we have summarized the literature on proteomics and metabolomics in two tables (Tables 7.1 and 7.2, respectively). As with transcriptomics, a variety of electrophoretic techniques have been used to characterize the proteome and metabolome. We refer the reader to the primary papers for the technical details on the techniques used and will instead focus our discussion on the discovery obtained from this research. Abril et al. (2011) list five different areas which characterize proteomic research: (1) descriptive proteomics, (2) differential expression proteomics, (3) post-translational modifications, (4) interactomics, and (5) proteinomics, with the first two being the primary focus of conifer proteomics research to date. This is akin to the focus in both

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_7

Table 7.1 A brief summary of proteomics projects in conifers

Species	Organ or tissue	Experimental workflow	Objectives of the study	Protein spots identified (and differential)	References
Conifer (multiple genera)	Pollen drop	2-DE, TOF MS/MS	Pollination drops		Prior et al. (2013)
Larix x eurolepis	Somatic embryos	2-DE, MS/MS	Somatic embryo development	1188 (147)	Teyssier et al. (2014)
Pinus armandii	Needles	2-DE, MALDI-TOF MS	Response to drought	343 (13)	He et al. (2012b)
Pinus armandii	Needles	2-DE	Response to drought and temperature	550 (27)	He et al. (2007)
Pinus bungeana	Nifedipine-treated pollen	2-DE, MS/MS	Pollen tube development	700 (34)	Wu et al. (2008)
Pinus massoniana	Seed	2-DE (DIGE), MS/MS	Seed development	N/A (61)	Zhen et al. (2012)
	2-yr-old saplings	2-DE (DIGE), MS/MS	Simulated acid rain	1000 + (65)	Wang et al. (2013)
Pinus nigra	Phloem	2-DE	Host-pathogen interaction	800 (19)	Wang et al. (2006b)
Pinus pinaster	Seeds (haploid megagametophytes)	2-DE	Genetic analysis	900	Gerber et al. (1993)
	Seeds (megagametophytes)		Population, tree variability, genetic background effect	968	Bahrman et al. (1994)
	Needle, bud, and pollen			902	Bahrman and Petit (1995)
	Needles (three-generation F-2 inbred pedigree)		Genetic analysis	900	Plomion et al. (1997), Costa and Plomion (1999)
	Needles	2-DE internal peptide micro-sequencing	Water stress	1000 (38)	Costa et al. (1998)

	Needles and xylem	2-DE internal peptide micro-sequencing	Descriptive	900 (needle) (324) 600 (xylem) (177)	Costa et al. (1999)
	Differentiating xylem from a 22-year-old tree at four heights up the tree and at four positions around the trunk		Wood formation	137 (26)	Plomion et al. (2000)
	Differentiating xylem tissue at the base and top of the stem	2-DE, MS/MS	Wood formation	1039 (160)	Gion et al. (2005)
	Differentiating xylem from juvenile and mature wood		Wood formation	1372 (33)	Paiva et al. (2008)
	Immature embryos	2-DE, MS/MS	Somatic embryo development	1428 (83)	Morel et al. (2014a)
	Somatic embryos	2-DE, MS/MS	Somatic embryo development	976 (52)	Morel et al. (2014b)
	Seedling stems	2-DE, MS/MS	Phototropism and gravitropism	486 (68)	Herrera et al. (2010)
Pinus radiata	Needles	2-DE, MS/MS	Organogenesis	850 (130)	Valledor et al. (2008, 2010)
	Xylem	LC MS/MS	Compression wood	175	Mast et al. (2010)
Picea abies	Seedlings: needles	2-DE, MS/MS	Response to drought	283 (5)	Blödner et al. (2007)
	Seedlings: needles and roots	2-DE, MS/MS 2-DE, MS/MS	Response to heat shock differentiation between 2 provenance samples	300–700 (14), several 100 (19)	Valcu et al. (2008a, b)
	Cell suspension cultures	iTRAQ	Elicitor (chitosan) induced changes in protein expression	1347 (35)	Lippert et al. (2009)
	Seedlings	MRM	Response to methyl jasmonate	19 (19)	Zulak et al. (2009)

(continued)

Table 7.1 (continued)

Species	Organ or tissue	Experimental workflow	Objectives of the study	Protein spots identified (and differential)	References
Picea glauca	Embryos	2-DE, ESTs	Embryo maturation up to 35 days after abscisic acid treatment	696–1250 (48)	Lippert et al. (2005)
Picea meyeri	Pollen tubes	2-DE, MS/MS 2-DE, MS/MS	Actin cytoskeleton (via LATB assay) Ca^{2+}-CaM signaling	600 (100)	Chen et al. (2006, 2009)
Picea obovata	Needles	2-DE (DIGE), MS/MS	Freeze tolerance: population variability	676 (43/252)	Kjellsen et al. (2010)
Picea sitchensis	Bark tissues	2-DE, MS/MS, cDNA microarray	Host response to insect feeding (*Pissodes strobi*)	1397 (104	Lippert et al. (2007)
Cunninghamia lanceolata	Seeds in six developmental stages	2-DE DIGE, MS/MS	Seed development	136 (71)	Shi et al. (2010)

Modified from Abril et al. (2011)

2-DE 2-dimensional electrophoresis, *TOF* time of flight, *MS/MS* tandem mass spectrometry, *MALDI* matrix assisted laser desorption/ionization, *DIGE* differential gel electrophoresis, *LC* liquid chromatography, *iTRAQ* isobaric tag for relative and absolute quantification, *MRM* multiple reaction monitoring

Table 7.2 A brief summary of metabolomics projects in conifers

Species	Study objectives	Technique	References
Pinus pinaster	Seasonal differences in wood formation	GC/MS	Paiva et al. (2008)
Pinus taeda	Somatic embryo development	GC/MS	Robinson et al. (2009)
Pinus taeda	Association study	GC/MS	Eckert et al. (2012a)
Picea abies	Somatic embryo development	GC/MS	Businge et al. (2012)
Picea abies	Somatic embryo development	GC/MS	Businge et al. (2013)
Picea glauca	Development of somatic embryos	NMR	Dowlatabadi et al. (2009)
Picea sitchensis	Cold acclimation	GC/MS	Dauwe et al. (2012)
Pseudotsuga menziesii	Variation among trees in xylem tissue	GC/MS	Robinson et al. (2007)
Taxus x media cv. *Hicksii*	Taxoid metabolites	HPLC/ NMR	Ketchum et al. (2003)

GC/MS gas chromatography-mass spectrometry, *NMR* nuclear magnetic resonance spectroscopy, *HPLC* high-performance liquid chromatography

transcriptomic and metabolomics research in conifers and will be the focus of the presentation in this chapter. Furthermore, proteomics and metabolomics research has generally focused on the same areas of biology as transcriptomics research, so four major topic areas will be presented: (1) wood formation, (2) biotic factors, (3) abiotic factors, and (4) seed development.

Wood Formation

Much of the work in conifer proteomics and metabolomics of wood formation has been done in *Pinus pinaster* by a group in France. Costa et al. (1999) characterized the proteome of needles and xylem. They found that 29% of the protein spots were unique to xylem. Plomion et al. (2000) followed with a differential proteomic study where they found that 19% of the protein spots were unique to compression wood versus controls. Gion et al. (2005) found 1039 protein spots from developing xylem of which 160 could be functionally classified (Fig. 7.1). Proteins from a diverse array of processes were found but surprisingly a greater proportion were defense related (19.4%) than were those involved in cell-wall biosynthesis (5.7%). To begin to understand the relative role of transcriptional versus translational gene regulation, Gion et al. (2005) also estimated the correlation between protein abundance and the homologous EST abundance ($r = 0.46$). This approach demonstrates a similarity between the mRNA pool and the protein pool in the same tissues at the same time but also points out that transcript abundance is not a perfect predictor of the active protein pool. Differences in the transcriptome and proteome in xylem tissue from crown wood versus base wood were also explored (Paiva et al. 2008). Not surprisingly, they found differential expression of both the transcriptome and proteome at different positions in the stem, and there was strong similarity in the

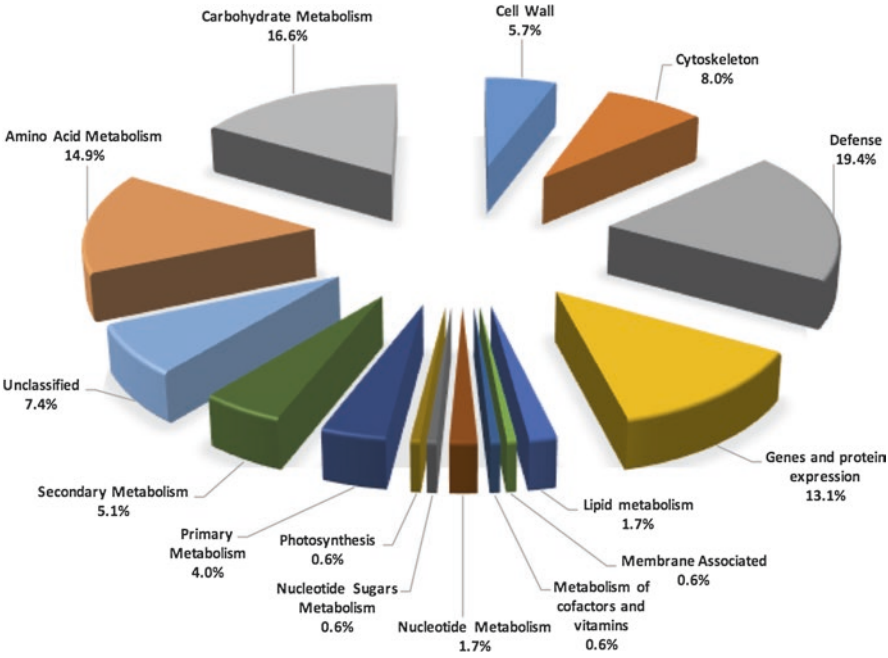


Fig. 7.1 Functional distribution of the major proteins in *Pinus pinaster* wood-forming tissue, separated by 2-DE. (From Gion et al. 2005)

functional categories of the transcriptome and the proteome. This same group of researchers also conducted a comparative transcriptomic and metabolomics study over a seasonal cycle. There are clear differences in the abundance of different metabolites over the season (Fig. 7.2) and also among different classes of genes in the transcriptome (Fig. 7.3). Studies of this type are highly descriptive but do illustrate the great molecular complexity that occurs in the transition from earlywood to latewood. Finally, Herrera et al. (2010) found 68 differentially abundant proteins that may be involved in phototropism and gravitropism in a study that compared control versus bent stems in *Pinus pinaster.*

Other than *Pinus pinaster*, only a couple of proteomic or metabolomic studies have been done in other conifers related to wood formation. Mast et al. (2010) were specifically interested in the membrane-bound proteins in *P. radiata* compression wood. They found 175 proteins from a number of functional groups, but as in the *P. pinaster* example, the largest proportion was in the defense-related category (Fig. 7.4). In *Pseudotsuga menziesii*, Robinson et al. (2007) found 139 metabolites common to a sample of 181 different trees and were able to correlate metabolomic profiles to economically important whole-tree phenotypes such as growth and physical and chemical wood properties. Studies of this type are again highly descriptive but nevertheless help identify the complex suite of molecular components that underlie complex traits of long-standing interest. Eckert et al. (2012a) conducted a metabolomics study in *Pinus taeda* somewhat similar to that of Robinson et al.

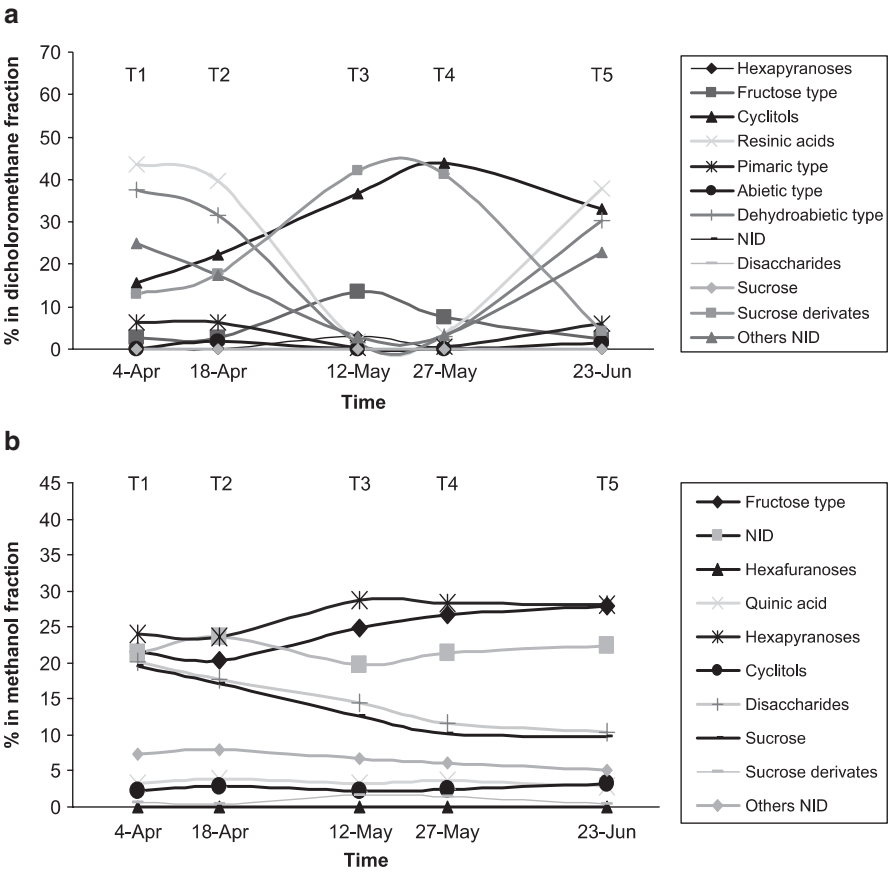


Fig. 7.2 Variation of metabolites from *Pinus pinaster* extracted in (**a**) dichloromethane and (**b**) methanol during the 2003 growing season. Metabolite quantification and identification were by gas chromatography/mass spectrometry (GC/MS). NID non-identified compound. (From Paiva et al. 2008)

(2007) in that both studies sought to discover genetic variation for differences in metabolite concentrations among individuals. Eckert et al. (2012a) found a large number of SNPs associated to differences in metabolite concentrations among individuals, thus providing an early look at the molecular genetic basis of metabolic differences among individuals.

Biotic Factors

Proteomic and metabolic approaches have been used in a small number of studies to discover proteins and metabolites that might be involved in response to fungal pathogens, insects, or chemical elicitors such as methyl jasmonate. Wang et al.

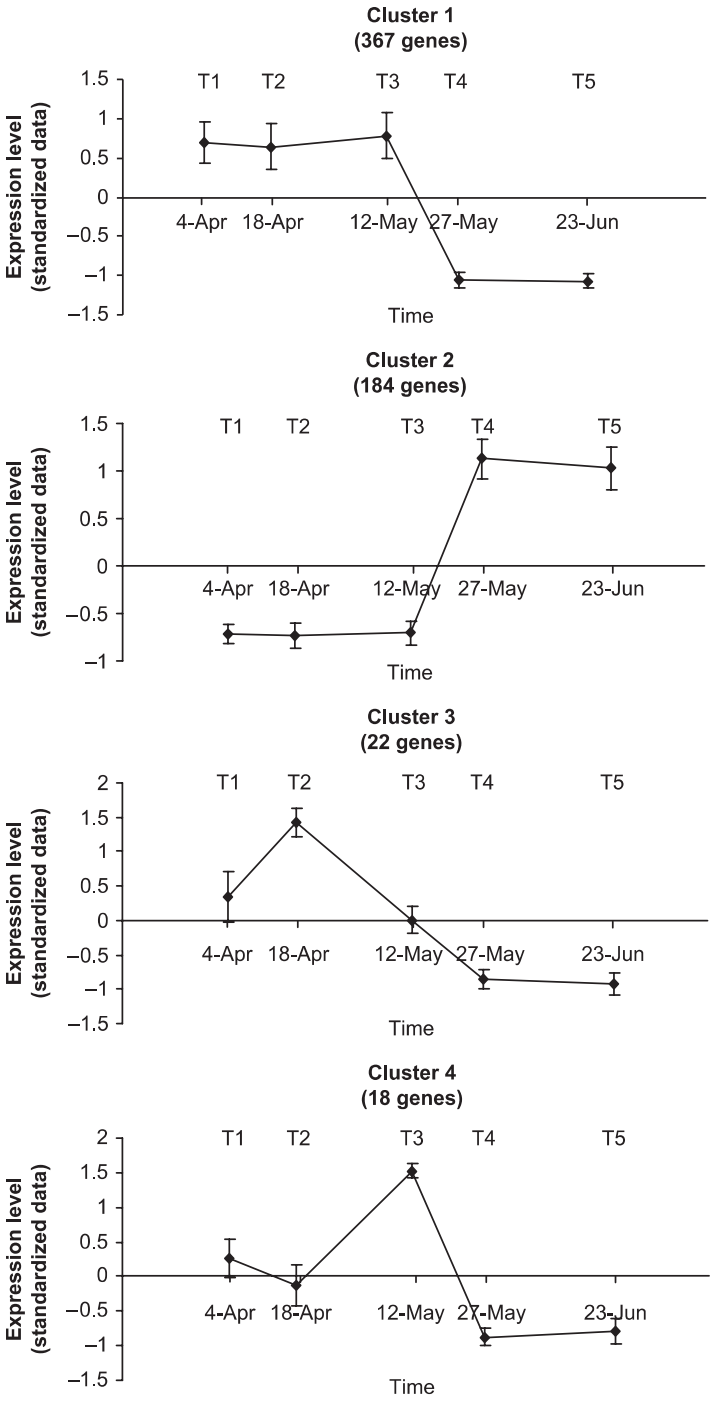


Fig. 7.3 Clustered mean expression profiles of differentially expressed genes during the 2003 growing season in *Pinus pinaster.* (From Paiva et al. 2008)

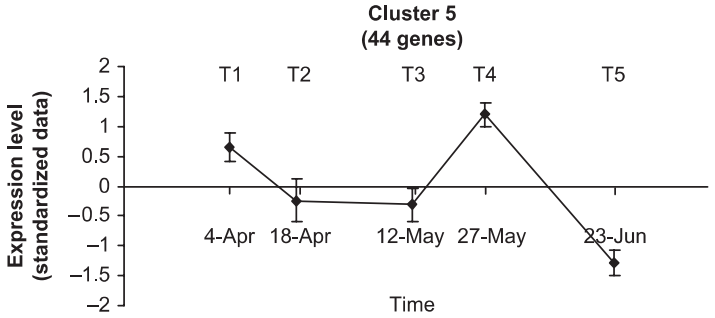


Fig. 7.3 (continued)

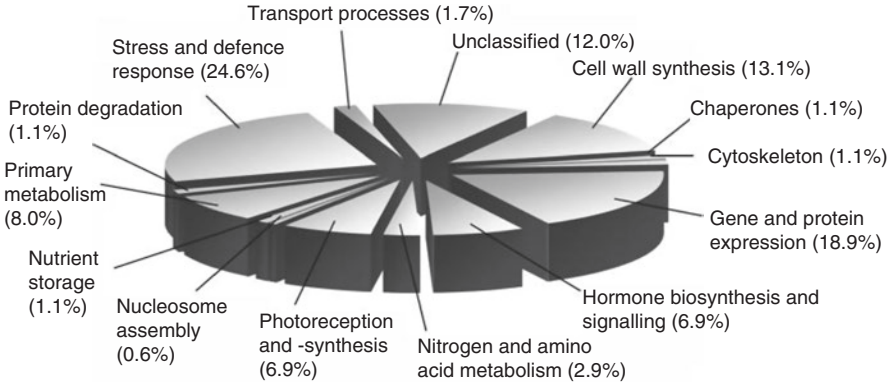


Fig. 7.4 Pie chart depicting the assignment of identified proteins to functional groups in *Pinus radiata* compression wood. (From Mast et al. 2010)

(2006b) searched for proteins that became abundant after challenge with two fungal pathogens, *Diplodia scrobiculata* and *Sphaeropsis sapinea,* in *Pinus nigra* seedlings. They found 19 proteins that were up- or downregulated, several of which were heat shock proteins. Surprisingly no other proteomic or metabolomics studies related to fungal attack in conifers were found in the literature.

One study has been done in *Picea* to determine what proteins might be involved in defense against insects. Lippert et al. (2007) compared the proteomes of *Picea sitchensis* shoot tissues before and after exposure to white pine weevils (*Pissodes strobi*). They detected 1397 protein spots in total, of which 104 were differentially expressed between the controls and insect-exposed tissue. Many of these proteins are known to be heat shock and stress related (Fig. 7.5). Furthermore, they compared the differential proteomic pool with the differential transcriptome from these same samples and observed products (mRNAs and proteins) resulting from different sets of genes, the conclusion again being that the transcriptome is not predictive of the proteome and they both need to be carefully characterized.

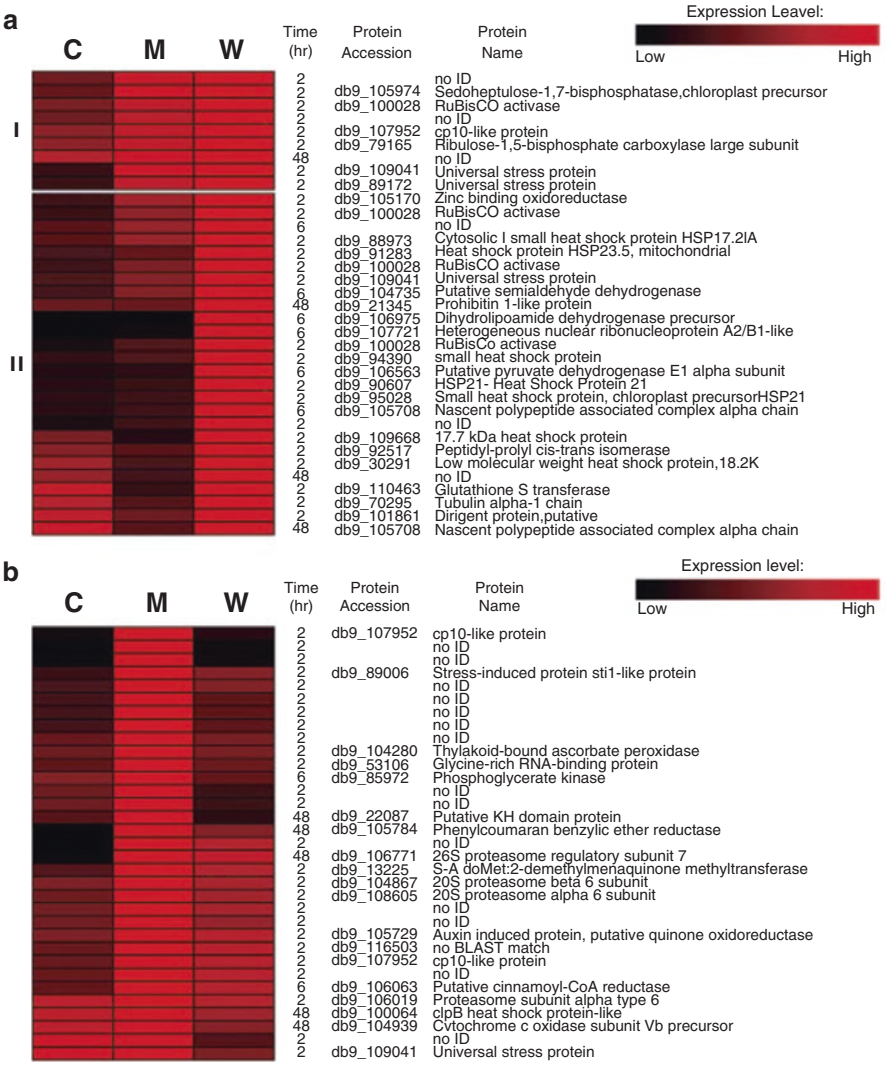


Fig. 7.5 Proteomic differences in *Picea sitchensis* shoots following exposure to white pine weevils. Clustering of differentially expressed proteins based on the pattern of relative expression among the three treatment groups: control (C), mechanically wounded (M), and weevil (W) treatments. (From Lippert et al. 2007)

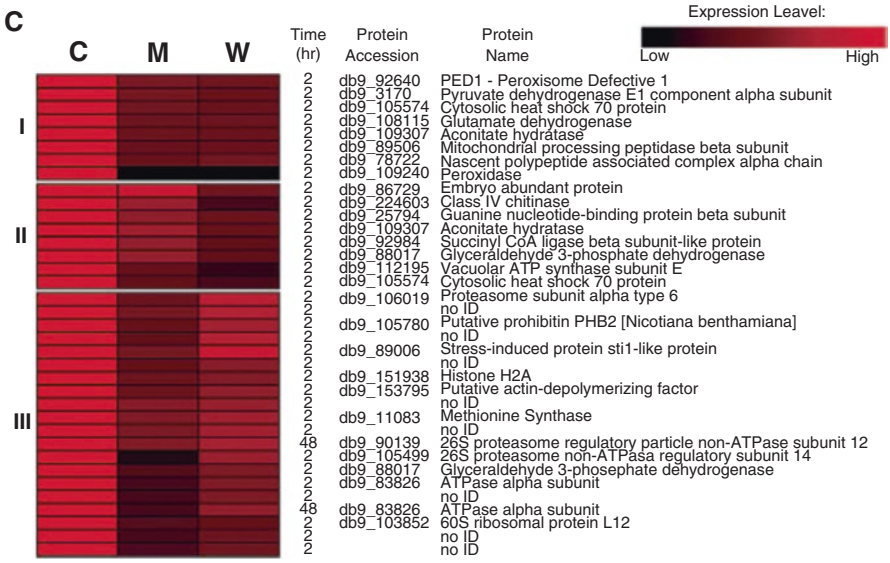


Fig. 7.5 (continued)

An indirect approach to studying responses to fungal disease or insect attack is to treat plants with some chemical that is known to elicit a similar response. One widely used chemical is methyl jasmonate (MeJA). Zulak et al. (2009) applied MeJA to *Picea abies* seedlings and observed the response in the transcriptome, proteome, and metabolome. This paper is an early example of an integrated "multi-omics" approach with focus on the products of terpene synthase genes. In total, 19 different polypeptides were monitored, all of which differ in abundance between control and MeJA-treated seedlings (Fig. 7.6). Furthermore, they observed a fairly high correlation with the corresponding transcripts and metabolites. Therefore, in the example of the terpene pathway and the terpene synthases, there seems to be a higher level of correspondence between the transcriptome, proteome, and metabolome that has not been observed in other studies. The selective reaction monitoring (SRM) method used for the proteome analysis by Zulak et al. (2009) allows for the protein-specific analysis of very closely related members of multigene families. Using the same method, Hall et al. (2011) identified a specific terpene synthase that contributes to the metabolome associated with weevil resistance in *Picea sitchensis*.

Another example of MeJA induction of metabolites is shown by Ketchum et al. (2003) where MeJA was applied to *Taxus x media* cv. *Hicksii* cell suspension cultures, and there was an increase in abundance of taxoid metabolites, the compounds leading to the synthesis of the cancer treatment drug Taxol. In a study using an elicitor to mimic a very early defense response, Lippert et al. (2009) applied a chitosan treatment to *Picea abies* cell suspension cultures and found 35 of 1347 proteins whose abundance increased following treatment. The advantage of elicitor experiments is that the experimental conditions can be tightly controlled relative to subjecting plants to diseases and insects, but the biological relevance of results will need to be confirmed.

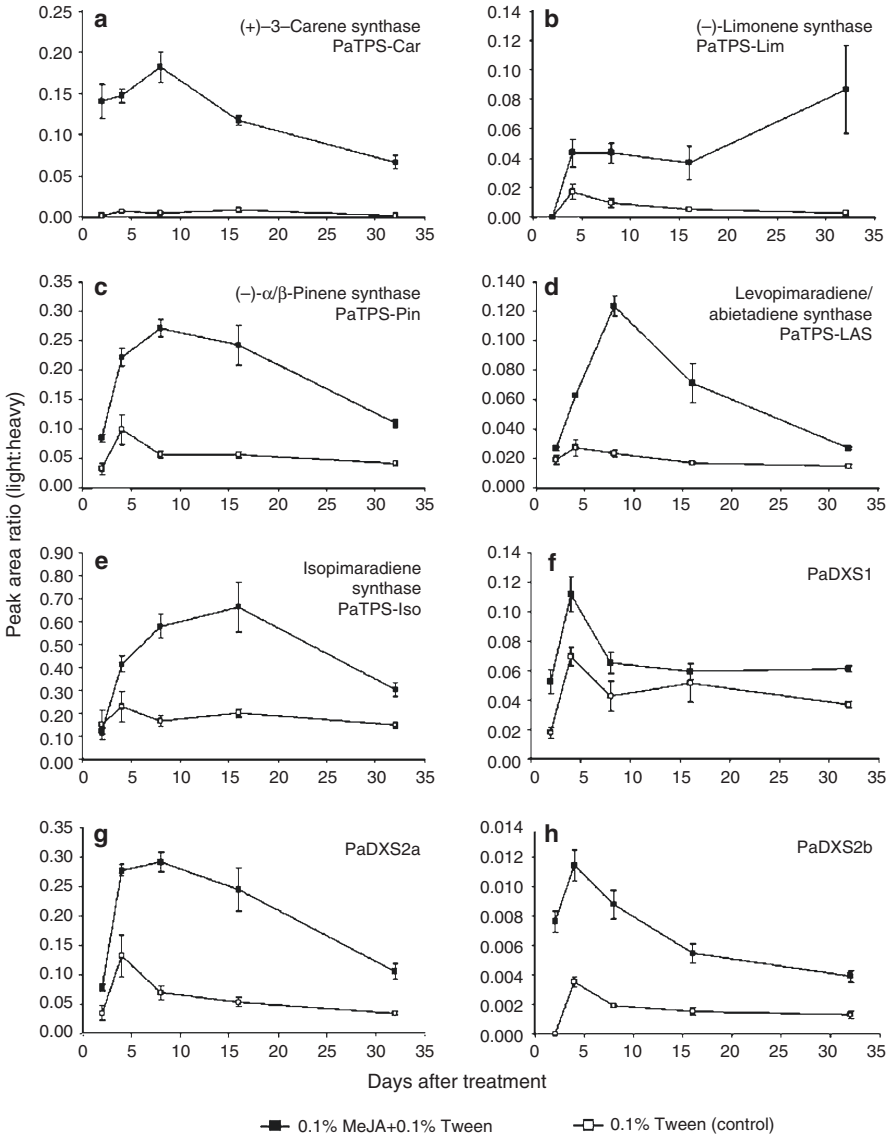


Fig. 7.6 Protein levels of terpene synthase (TPS) and 1-deoxy-D-xylulose5-phosphatesynthase (DXS) enzymes as detected by selective reaction monitoring (SRM) from a time course of methyl jasmonate (MeJA)-treated and control *Picea abies* bark protein extracts. The results are the peak area ratio of the endogenous peptide (light) to the isotopically labeled internal standard (heavy). Data presented are the mean of four biological replicates for each time point. (From Zulak et al. 2009)

Abiotic Factors

Changes in the proteome and metabolome in response to abiotic factors, primarily moisture and temperature, are an active area of research. One of the first studies of this type was by Costa et al. (1998) who observed differences in the proteome between control and drought-stressed *Pinus pinaster* seedlings. They observed a total of 1000 protein spots of which 38 (3.8%) were either upregulated or downregulated in response to drought. These proteins were from several and different functional categories including photosynthesis, cell elongation, antioxidant metabolism, and lignification. Blödner et al. (2007) performed a very similar experiment in *Picea abies* but found only 1.7% of the proteins were differentially expressed. The number of differentially expressed proteins in both the *Pinus pinaster* and the *Picea abies* experiments were quite small, although in both cases many were involved in photosynthetic processes. It is not clear if any of these proteins were in common between these studies and whether these are directly related to a drought response or simply associated. He et al. (2007) also performed a drought-stress experiment in *Pinus armandii* but also included a heat shock treatment. They found that 4.9% of the proteins were differentially expressed in response to both treatments and that these genes were essentially from the same functional categories as reported by Costa et al. (1998) in the *P. pinaster* study. Later, He et al. (2012b) expanded their drought-stress experiments to three additional pine species in addition to *P. armandii.* Results were very similar to the previous study in terms of differentially expressed proteins and functional categories; however, this time they showed that *P. armandii* had the strongest response in all physiological and proteomic measures. *Pinus armandii* is a member of the subgenus *Strobus*, whereas the others are members of subgenus *Pinus* (Chap. 16). Whether this difference is due to evolutionary differences between the subgenera or specific to *P. armandii* is not clear. The results from these proteomic and metabolomics studies are also consistent with the transcriptomic study reported in Chap. 6 (Lorenz et al. 2011) where response to drought is a very complex physiological and molecular process and breeding for drought adaptation will require taking a quantitative genetic approach (Chap. 11).

Results from metabolic studies in response to temperature are very much in line with proteomic and metabolomic studies for drought and transcriptomic studies for response to drought and temperature. Dauwe et al. (2012) conducted a metabolomics study in *Picea sitchensis* that used the same populations as the earlier transcriptomic study of Holliday et al. (2008) (Chap. 6). The results of these two studies were combined by generating a "metabolic map" (Fig. 7.7). This figure attempts to show correlations between the transcriptome and metabolome over the seasonal cycle but again reinforces the massive complexity of these processes.

One final proteomic study in conifers worth noting was that of Wang et al. (2013) who treated *Pinus massoniana* seedlings with simulated acid rain and found 65 spots (28 unique proteins) of more than a 1000 protein spots that were differentially regulated in response to the treatment. As with other studies, these proteins were assigned to many different functional categories.

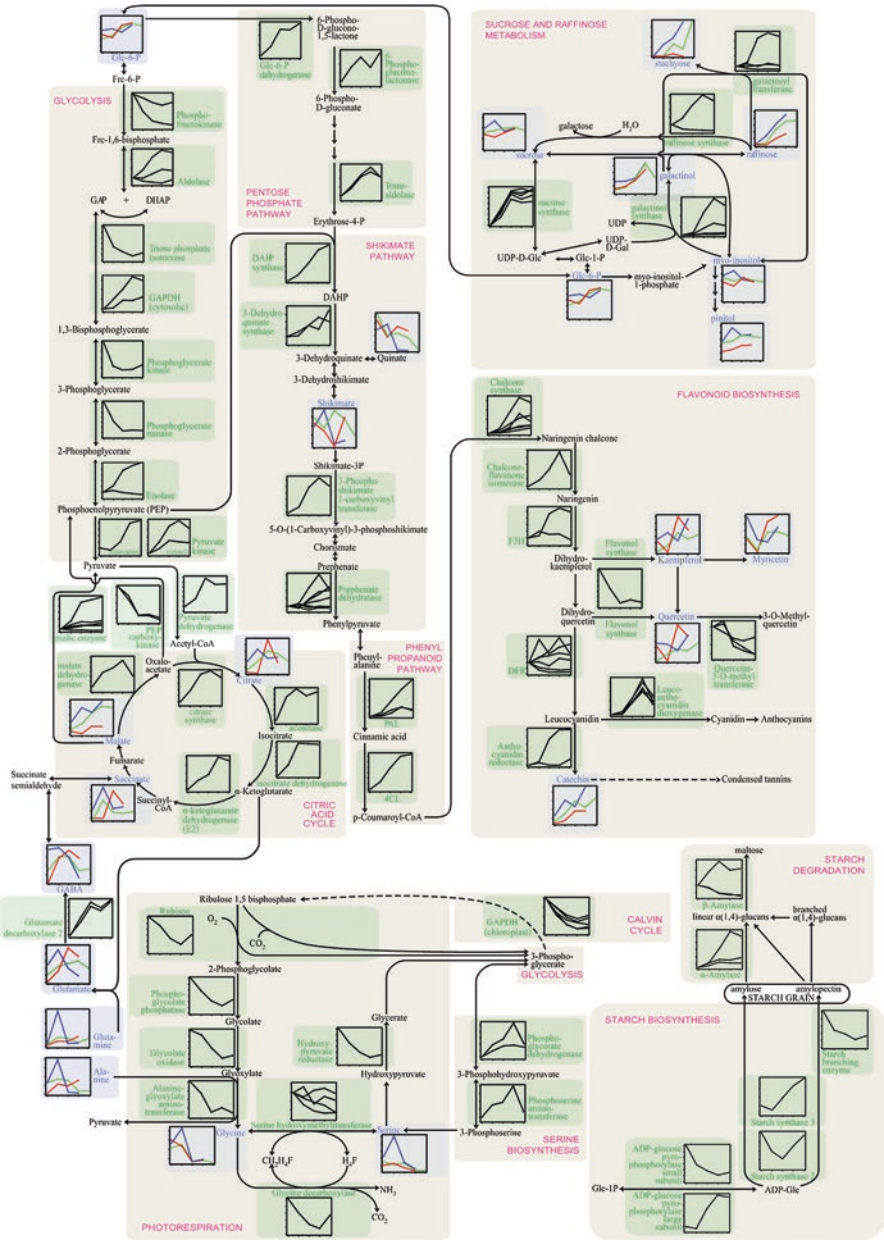


Fig. 7.7 Metabolic pathways in *Picea sitchensis*. Metabolites that were identified in the GC/MS spectra are shaded in blue. These metabolites are accompanied by graphs with the *x*-axis representing the time points 1 (30 August) to 5 (December 13), the *y*-axis representing the average peak intensity in the GC-MS chromatograms, and the color indicating the population: California (CA), red; British Columbia (BC), green; and Alaska (AK), blue. Enzymes for which the transcript levels during autumn cold acclimation have previously been monitored are shaded in green. These enzymes are accompanied by graphs in which the *x*-axis represents the time points and the *y*-axis represents the expression level fold-change as compared to time point one in the BC population, as reported in Holliday et al. (2008). (From Dauwe et al. 2012)

Seed Development and Somatic Embryogenesis

Interest in proteomic and metabolomic profiling during zygotic seed development and somatic embryogenesis is motivated by the desire to develop efficient means of cloning trees for commercial reforestation. Thus, research in this area has been focused largely on commercial conifers. The general approach taken to identify important proteins and metabolites during zygotic or somatic embryo development is to sample at different time points in development and search for differentially abundant proteins or metabolites. Lippert et al. (2005) used this approach with somatic embryo cultures of *Picea glauca* and found between 696 and 1250 protein spots across several sampling stages, of which 48 were differentially expressed. As always seems to be the case, the functional classification of these 48 proteins revealed that they were involved in many different cellular processes (Fig. 7.8). Subsequently, Dowlatabadi et al. (2009) conducted a metabolic profiling experiment in *Picea glauca* somatic embryo cultures with the specific goal of finding metabolites unique to somatic embryos grown on maintenance media versus maturation media. Clear differences were found and metabolites specific to somatic embryos grown on maturation media were identified. However, there was no attempt to correlate these results with the earlier proteomic profiling. Teyssier et al. (2014) conducted a proteomic profiling experiment in *Larix* × *eurolepsis* somatic embryo cultures and again found a large number (12%) of proteins that varied in abundance across development stages. A unique aspect of this study was that it was combined with assays to determine the extent of DNA methylation (Chap. 6) across the time course in an attempt to identify a cause of the differential abundance. Likewise, Morel et al. (2014a) combined proteomic with transcriptomic profiling in *Pinus pinaster* somatic embryos under different levels of moisture availability to better understand the molecular processes leading to embryo maturation. They found that 6% of the proteins were differentially abundant and from very diverse functional categories.

Proteomic profiling has also been conducted with zygotic embryo development. Shi et al. (2010) sampled at several stages in developing seeds of *Cunninghamia lanceolata* and found that 52% of the protein spots varied in abundance across the developmental stages. These proteins were involved in many cellular processes including carbon metabolism, methionine metabolism, energy production, protein storage, disease, defense, and embryo development. This long list again points out the great molecular complexity of all developmental processes. Zhen et al. (2012) also studied proteomic profiling in *Pinus massoniana* zygotic embryos and reported results very similar to those of the previous study. A direct comparison of proteomic profiles between developing somatic embryos and zygotic embryos was done in *P. pinaster* (Morel et al. 2014b). From a total of 976 protein spots common to both somatic and zygotic embryos, 52 were differential abundant, leading the authors to conclude that there is commonality to the proteomic profiles of these two embryo types.

Metabolomic profiling of somatic embryo development has also been studied in a couple of conifers. Robinson et al. (2009) sought to find a relationship between

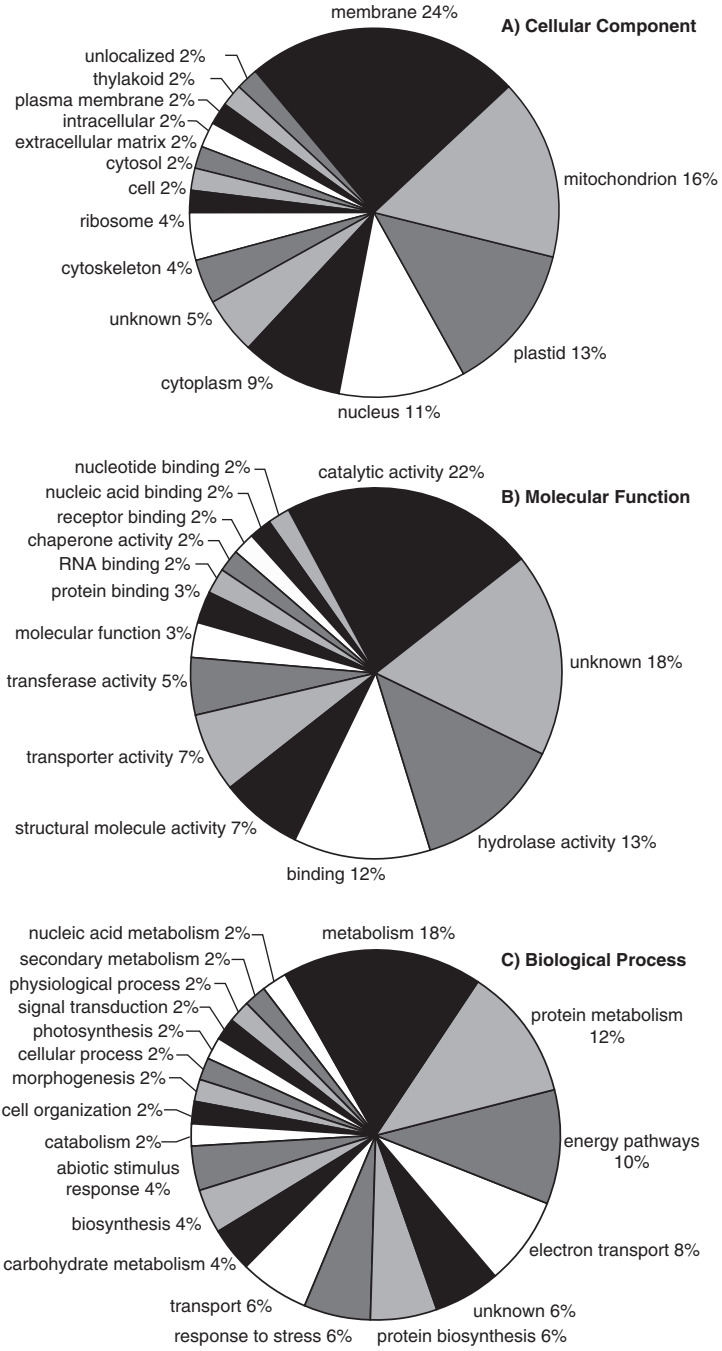


Fig. 7.8 Ontological classification of differentially regulated proteins in *Picea glauca* in terms of (**a**) cellular component, (**b**) molecular function, and (**c**) biological process. (From Lippert et al. 2005)

metabolic profiles and regenerative capacity in *Pinus taeda* somatic embryo cultures. Although the correlations of individual metabolite concentrations with regenerative capacity were quite low, they were able to use data from all the metabolites to build a very good predictive model (Fig. 7.9). Even though a causative relationship could not be inferred, it is clear that differences in the metabolome among individual genotypes are related to and predictive of regenerative capacity. Businge et al. (2012) conducted a similar study where they measured the metabolomics profiles at different developmental stages in three different *Picea abies* cell lines, one normal, one aberrant, and one blocked for somatic embryo development. Some very large differences in abundance of different metabolites were observed among the cell lines at different developmental stages. Results such as these are again quite descriptive but begin to point to the pathways that might be important in regenerative success. Businge et al. (2013) also performed an experiment to test the effect of carbohydrate and osmoticum on the developmental, proteomic, and metabolomic profiles of developing somatic embryos of *Picea abies*. A clear relationship was observed between storage reserves and eventual germination with abundance of specific proteins and metabolites resulting from carbohydrate and osmoticum treatments.

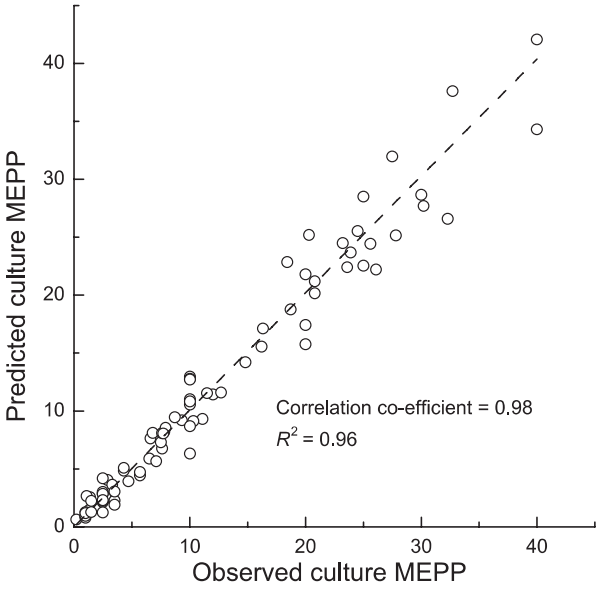


Fig. 7.9 Plot of measured versus predicted productivity of *Pinus taeda* somatic embryo cultures, as determined by complete cross-validation of the prediction model. MEPP (mean number of embryos produced per plate) is the mean number of mature embryos per culture plate, for a culture that was distributed across multiple plates prior to embryo maturation. Each marker represents a single culture. Dashed line represents fit of least squares regression. The correlation coefficient and coefficient of determination (R^2) are provided. (From Robinson et al. 2009)

Summary

The study of the proteome and metabolome in conifers is very new as the technologies to measure the abundance of large numbers of proteins and metabolites simultaneously have only recently been developed. Isolating and measuring proteins and metabolites is technically more demanding than working with DNA (genome) or RNA (transcriptome). Many research organizations have built core facilities with specialized equipment and trained staff to provide this research capacity. The biological functions studied in conifers using proteomic and metabolomic data are very similar to those used in transcriptomic studies—wood formation, biotic stress factors, abiotic stress factors, and development. And just as in transcriptomic studies, comparative and time-course studies have been conducted to identify proteins and metabolites uniquely or abundantly present at different developmental stages or in response to biotic and abiotic stress. The complexity of these differences is nearly overwhelming. It will take years of replicated study and sophisticated bioinformatics tools to ultimately and precisely understand the patterning and genetic control of the proteome and metabolome in conifers.

Part II

Variation

Phenotypic Variation in Natural Populations

8

Introduction

The recognition and subsequent study of genetic variation in conifers began with the observation that trees varied in observable or measurable (phenotypic) traits and that variation was often distributed geographically. Virtually everything we learned about genetic variation (genotypic) in forest trees before the development of allozymes, genetic markers, and genome sequences derived from the study of phenotypic variation among and within natural and domesticated populations of trees. The predominant research approach for such studies has been the common garden trial, which provides an objective means for dissecting observable phenotypic variation into its component effects, genotypic and environmental, as described by the simple equation P = G + E (phenotype = genotype + environment; Fig. 8.1). In forestry, common garden studies in which accessions from multiple natural populations of a single species are evaluated within the same test site are known as provenance trials. Common garden studies that evaluate pedigreed accessions from domesticated populations, usually across a narrower geographic area, are referred to as genetic tests or progeny tests. Here we are concerned only with the study of genetic variation in natural populations, though genetic testing has revealed a great deal about variation in phenotypic traits and the genetic basis of that variation. Historically treatments of this topic have frequently used the misleading term "geographic variation" to describe patterns of phenotypic variation across the landscape. Since we are not interested in variation in geography, but in biological variation shaped by natural selection driven by climatic factors associated with geography, the choice of "phenotypic variation" in the chapter title more accurately reflects the contents discussed here.

In this chapter, a substantial literature on provenance variation in conifers is reviewed by categorizing studies according to the experimental approach used, the implied application of the study, and the species being tested. Specific results of select studies are discussed in detail, and common patterns of geographically mapped genetic variation are summarized. It is beyond the scope of this chapter to cite all the rich and relevant literature on this topic. We have tried to include key

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_8

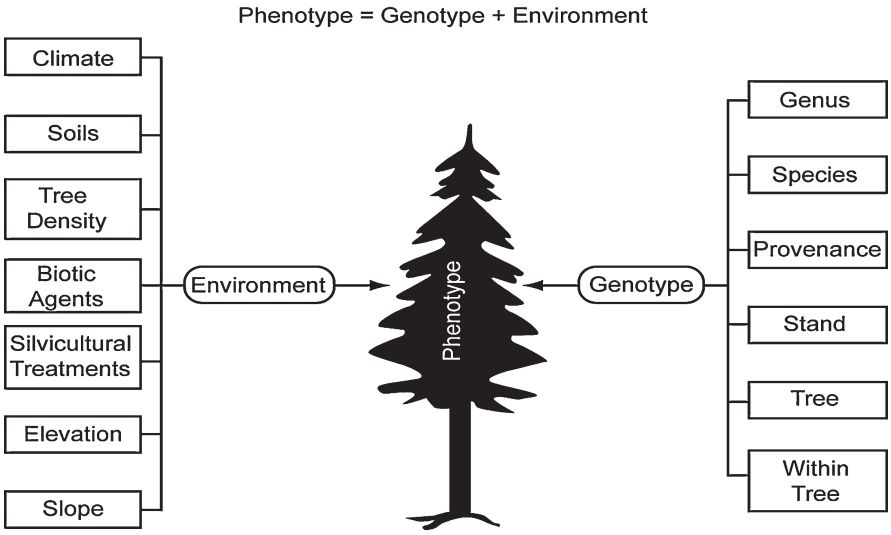


Fig. 8.1 Schematic diagram illustrating potential factors that contribute to a tree's phenotype. Most conifer species exhibit abundant phenotypic variation. (From White et al. 2007)

historical studies and references and a large sampling of more recent publications representing a broad taxonomic and geographic distribution. Detailed discussions of experimental design and analysis of results are not covered here. Readers are encouraged to seek out previous treatments that cite much of the earlier literature and cover the concepts and applications of the study of geographic variation in depth (Burley and Nikles 1973; Dorman 1975; Wright 1976; Zobel and Talbert 1984; Turnbull and Griffin 1986; Ladrach 1998; Ledig 1998; White et al. 2007; Morgenstern 2011).

Definitions

Before proceeding to a review of the topical literature, it is necessary to define some of the terms and concepts related to the study of geographically distributed genetic variation in phenotypic traits as they are used here. Our brief definitions follow more thorough treatment on the subject provided by others (Wright 1976; Zobel and Talbert 1984; White et al. 2007).

Terms for Describing the Identity of Experimental Plant Materials: Provenance, Population, Seed Source, and Accession

Conifer common garden trials are typically comprised of seed or seedlings collected from portions or the entirety of a species' natural range. The term *provenance*,

adopted in France in the late 1700s, is used to denote the geographic location, and its associated climate, from which the seeds were collected in natural stands. The term is not easily defined by any measure of size (area) or physical description and may refer to very different things depending on sampling density, range heterogeneity, etc. An accepted criterion for discerning between two provenance collections is that trees from those collections are not likely to be interbreeding. The term *population* refers specifically to the individuals originating at a provenance location. Short-term trials with many "entries" typically refer to populations. The same non-interbreeding criterion applies here as well, but further geographic delineation is typically lacking. *Seed source* also refers to the location from which the seed was collected but provides a means of identifying samples that come from non-native provenances. For example, seed collected from a Finnish provenance of Scots pine, growing in a common garden trial in Madrid, Spain, would have a seed source of Madrid. An *accession* is a term or number assigned to each seed lot in an experimental planting. It is a label used to track specific genetic materials from collection to planting. Each individually sampled tree in a provenance collection will typically receive its own accession number.

Terms for Describing Patterns of Genetic Variation on the Natural Landscape: Cline, Race, and Ecotype

Mapping genetic variation against geographic or climatic variables typically, though not always, reveals a recognizable pattern or distribution. The term *cline* is used to describe a continuous gradient for each individual trait. Traits that vary in a continuous gradient, typically in concert with an associated environmental gradient, are said to be distributed in a clinal pattern (Huxley 1938). A clinal pattern of variation is common for adapted traits in wide-ranging species with continuous distributions, and different traits may exhibit very different clinal patterns of variation. When provenances/populations vary sufficiently (often defined subjectively) from one another along a genetic/environmental gradient, they are classified as belonging to different *races*. In mountainous areas, elevational races are commonly associated with traits like frost tolerance or bud set timing (Campbell 1979; Rehfeldt 1983, 1988; St. Clair et al. 2005). Given that clinal trends may vary by trait, it is possible that a species can be defined by multiple races (Shutyaev and Giertych 2000). A more collective term is *ecotype* which describes a race made up of genotypes adapted to a particular habitat or ecological niche (Turesson 1922; White et al. 2007). An ecotype is defined by a suite of traits common to individuals within a population or geographic location, as may occur in disjunct or widely separated populations (Wright et al. 1969; La Farge 1975; Cordoba-Rodriguez et al. 2011). An ecotype may represent an early stage of taxonomic subdivision or speciation.

Historical Perspectives

The first reported studies recognizing geographic patterns of variation among populations of conifer species occurred in Europe in the 1700s, well before Mendel described the mechanisms of inheritance and Darwin fully developed the concept of natural selection. The earliest seed source trials, conducted by the inspector general of the French navy, with *Pinus sylvestris*, were established in France in the mid-1700s and repeated ~70 years later by another Frenchman, Pierre Vilmorin, with the goal of finding the best seed sources for plantation establishment in their local regions. These studies, along with similarly dated reports from Germany and Scandinavia and the subsequent expansion of studies of geographic variation in common garden studies, are referenced and described in fascinating detail by Langlet (1971) and Morgenstern (2011) among others. Many of these early efforts resulted in the importation, testing, and selection of provenances of conifer species from western North America, ultimately leading to the development of local land races—populations that adapt, through natural selection, to the exotic environment (White et al. 2007).

In the USA, the discipline of forest genetics essentially began with the establishment of intraspecific racial trials of *Pinus ponderosa* and *Pseudotsuga menziesii* in the early 1910s, conducted by staff of the newly created US Forest Service (Munger and Morris 1936; Weidman 1939). These trials, though unreplicated, set the stage for increasingly sophisticated trials to follow. By the 1930s and 1940s, it was widely recognized that forest trees possessed a great deal of heritable variation, much of it distributed on the landscape in nonrandom patterns, and that useful levels of genetic variation among populations and among trees within populations occurred for most species (Bates 1930; Wright 1944, 1953). For much of the remainder of the twentieth century, provenance trials proliferated for forest trees throughout the USA (Wheeler et al. 2015) and abroad.

Though provenance studies were, and still are, predominantly applied science, driven by the desire to improve productivity of managed forests, the common garden concept was similarly embraced for basic research that contributed significantly to development of the scientific discipline of genecology, the study of the relationship between genetic properties of populations and their environments (Turesson 1922; Huxley 1938; Clausen et al. 1940; Langlet 1971). By the 1980s, genecological studies were becoming widely adopted by forest scientists, at first relating patterns of genetic variation to physical/geographical variables (e.g., elevation, latitude and longitude, soil type) and later directly to climatic variables as improved climatic sampling and statistical methods interpolating climatic data between weather stations became available (Campbell 1979, 1986; Rehfeldt 1983, 1990; St. Clair et al. 2005; St. Clair 2006). Today it is widely recognized and accepted that patterns of genetic variation in wide-ranging species predominantly reflect adaptation to climatic gradients, implying natural selection is a significant evolutionary force acting on populations (Heslop-Harrison 1964; Stern and Roche 1974; Morgenstern, 2011; Rehfeldt et al. 2014a). Such patterns seem nearly universal in temperate and boreal forest species, though exceptions occur in conifers, particularly in species with

highly disjunct populations or those with ranges that do not include highly heterogeneous environments. In some of these cases, other evolutionary forces such as migration, drift, mutation, introgression, and inbreeding have variously been invoked to explain patterns of variation.

Application of Common Garden Experiments

Beyond elucidating the influence of evolutionary forces on patterns of genetic variation in natural populations, common garden trials have long been considered the logical first step in conducting forest genetics/tree improvement research (Wright 1976; Zobel and Talbert 1984; White et al. 2007). First and foremost, they provide estimates of the amount and distribution of genetic variation in a species of interest, illuminating the opportunities, or lack thereof, for capturing improvement through selection among or within populations and more intensive tree breeding approaches. In some instances, common garden trials alone are sufficient to meet tree improvement needs, but in most cases, they provide guidance for an array of important tree improvement decisions (modified from White et al. 2007). Provenance studies may help in:

1. Identifying superior performing provenances for a given planting region or climatic range.
2. Setting parameters on how far (physical or ecological distance) seed/seedlings from a given provenance can be safely moved for planting elsewhere and remain adapted to new conditions. Delineating breeding and deployment zones and design of selection and genetic testing programs follow from these results.

While these applications, or variants thereof, comprise the primary motivation for conducting most common garden trials, forest geneticists have also used them to guide gene conservation strategies (CAMCORE Cooperative 2000; Gallis et al. 2007), resolve taxonomic queries (Conkle and Critchfield 1988; Rafil et al. 1996; Rehfeldt 1997, 1999; Potter et al. 2015), quantify and describe provenance-by-test site interactions (Moura et al. 1998; Karlsson et al. 2001; Correla et al. 2010; Rweyongeza 2011), investigate heterosis in interprovenance crosses (Ilstedt and Eriksson 1986; Harfouche et al. 1999), seek sources of insect and disease resistance (Wheeler et al. 1976; Zhang et al. 1997; Dvorak et al. 2007), and to reconstruct populations that have been lost due to stand disturbance (Chalupka et al. 2008).

With the growing realization that climate change and forest health issues are likely to affect the extent to which native populations remain adapted to their local environments, provenance trials have been used to evaluate and guide recommendations for meeting future environmental challenges (Schmidtling 1994; Rehfeldt et al. 1999; St. Clair and Howe 2007; O'Neill et al. 2008; Thomson et al. 2009; Wang et al. 2006a, 2010a; Bansal et al. 2015a, b). Increasingly, such efforts are included in studies identified by terms such as *assisted migration* and *landscape genomics* (Ukrainetz et al. 2011; Pedlar et al. 2012; Alberto et al. 2013; Wheeler and Neale 2013).

Experimental Approaches and Analytical Methods

The previous section addressed "why" studies of genetic variation of phenotypic traits in natural populations are conducted (the application of results or purpose). Here we speak in general terms of "how" they are conducted by adopting the terms *approaches* and *methods* to guide our sorting of studies. Though the terms have been used interchangeably in some treatments, here the term approach refers, principally, to the physical nature and duration of the common garden experiments, while the term method refers to an admittedly subjective amalgam of experimental facets that include the type of dependent and independent traits studied and the analytical tools used to reveal relationships between the traits and how that knowledge can be used. This treatment reflects the increased complexity and sophistication of research approaches that have occurred over the last several decades in this field of study.

Experimental Approaches

In the most recent text on forest genetics, White and co-authors (White et al. 2007) list and describe three experimental approaches for the study of geographic variation: an indirect (genecological) approach which uses short-term common garden tests in artificial environments, a direct approach which uses long-term common garden trials in field experiments, and the use of genetic markers in trees collected in natural populations. The latter, the use of genetic markers to elucidate patterns of variation in neutral and adaptive traits, will not be discussed here but will be covered later in this volume (Chaps. 9 and 10, respectively).

Short-term genecological studies are typically conducted in growth room, greenhouse, nursery, or other artificial environments for periods, defined here, of a few months to a few years (Fig. 8.2). They are typically composed of many provenances/populations and relatively few parent trees (families) within provenance. Environmental conditions of the trials can be carefully monitored and controlled. Genecological studies seek to characterize the statistical relationships between measured traits and the environmental conditions of the provenance source locations. The key function of genecological studies is to elucidate the selective forces that shape the patterns of variation observed in phenotypic traits. Traits that vary in response to geographic or environmental conditions are likely adaptive and under selection. When used to guide delineation of seed zones and seed transfer guidelines, short-term trials are limited by the assumption that local sources are best adapted for local sites. This assumption may not always be accurate.

The direct approach to provenance testing evaluates adaptation and productivity over the long term in field tests located in environmental conditions representative of the species' native range or in non-native locations where investigators seek to identify potentially useful sources for establishing land races (Fig. 8.3). The primary objectives of the direct approach are to identify superior provenances/populations and guide how they are deployed across the landscape. It does not assume

Fig. 8.2 A short-term, artificial environment common garden trial of *Pseudotsuga menziesii* var. *menziesii* (Douglas-fir) where trees from hundreds of provenances (populations) are grown side by side for 2–3 years and evaluated for growth and phenological traits. The phenotypic results are related to the geographic or climatic variables from which the provenances originated. (Courtesy of Dr. Bradley St. Clair, USDA Forest Service, Corvallis, OR)

Fig. 8.3 An expansive provenance trial of *Pinus patula*, a species native to the highlands of Mexico, being evaluated as an exotic species in South Africa. (Photo taken from: CAMCORE Cooperative 2000, p. 167)

adaptation of local sources to local environments. The direct approach considers the responses of provenances/populations to the range of environments sampled by the test sites and the responses of provenances/populations to the geographic or climatic transfer distance between sources and test sites. These alternative experimental approaches employ different sampling methods, address different objectives, and possess varying attributes and shortcomings (White et al. 2007; Table 8.1).

Table 8.1 Advantages, disadvantages, and applications of alternative approaches used in conifer common garden trials

	Method	
	Short-term artificial environment	Long-term field trial
Advantages	Many provenances/populations may be evaluated Trials produce a large amount of information on adaptive traits (phenological, physiological, morphological) in a short time Experiments are statistically powerful (low experimental error) Experimental environments may be manipulated (imposed drought, freeze, pest challenges)	Provides superior information for making final decisions about seed transfer guidelines Provides superior information for selecting the best provenances for reforestation needs Provides an opportunity to test for genotype by environment interactions Allows for a longer evaluation period of adaptive traits
Drawbacks	Results may not mimic those in natural environments Only a short juvenile period of the tree's life cycle is tested Studies do not reveal best provenances for specific environments or silvicultural regimes likely to be experienced	High costs of establishment, maintenance, and measurement Long period required to obtain information of interest Limited number of provenances that can be tested due to space requirements
Applications	Characterizing patterns of adaptive genetic variation Elucidating the selective forces that mold the patterns of variation Developing preliminary seed transfer guidelines within a region Narrowing the number of promising provenances to be tested in long-term trials	Characterizing patterns of adaptive genetic variation Selecting the best specific provenances for an operational planting program Selecting materials to put into an applied tree improvement program

Analytical Methods

Investigators have relied on an array of analytical methods to extract information from common garden trials and to guide them in deployment decisions. Early trials relied on analysis of variance (ANOVA) to estimate genetic and environmental variances and test for significant differences among provenance means. Later analyses relied increasingly on regression and correlation analyses to investigate the relationships between phenotypes and environmental characteristics. Principal component analysis (PCA) and canonical correlation analysis (CCA) were/are often used to permit estimation of simultaneous effects of multiple independent variables on dependent variables.

The types of relationships explored include (1) genecological functions describing performance in a common garden test as a function of the climates of seed sources, (2) response functions describing how individual provenances/populations respond to being planted in different climatic conditions represented by

multiple test sites, and (3) transfer functions describing how far away, usually climatically, populations may be moved between test sites and sources and remain adapted or productive. As discussed above, determining genecological functions is the primary objective of short-term indirect approaches, whereas determining response functions and transfer functions are the primary objectives of long-term direct approaches. Genecological functions may also be determined in long-term direct tests.

A direct approach allows the ability to calculate both the effect of climatic conditions at common garden planting sites (environmental effects) and the effect of climatic conditions on populations at their provenance location (genetic effects). A direct approach can be used to predict the effect of climate change on a population's performance regardless of whether the change was due to moving the population to a different climate or to global climatic change across all provenances. Response functions must be calculated individually for all populations since they may vary genetically. Transfer functions are calculated as the difference in performance of multiple populations when growing under test site climates and when growing under their provenance climates (i.e., site climate minus provenance climate) (Matyas 1994; Rehfeldt et al. 1999). Wang et al. (2010a) subsequently integrated the environmental effects of response functions and the genetic effects of transfer functions in a single "universal response function (URF)". They used the URF to predict responses of *Pinus contorta* populations as a function of the climate in which they were growing and the source climate in which they evolved. While all methods noted here possess limitations, they provide valuable predictive power for guiding future reforestation efforts in rapidly changing environments, assuming appropriate common garden trials are established in a timely manner.

Dependent and Independent Variables

By far the most commonly studied dependent variables in conifer common garden trials are those related to survival, growth, and growth rhythm (phenology) such as tree height, diameter (caliper), volume, timing of bud flush, bud set, and shoot elongation, as well as stem/crown form traits. Less frequently, variation related to insect and disease tolerance or resistance is studied (Wells and Wakeley 1966; Wheeler et al. 1976; Yanchuk et al. 1988). These are often accompanied by estimates of cold and drought hardiness and to a lesser extent studies of tolerance to salinity and by measures of physiological traits such as water use efficiency (WUE), carbon isotope discrimination (CID), cavitation, and photosynthetic rates. Common independent variables used include geographic and topographic traits such as latitude, longitude, elevation, slope, aspect, and distance from an ocean or mountain range crest. Occasionally, soil type is included. More importantly today, independent variables include a wide array of climatic variables that reflect seasonal temperature and moisture regimes.

Common Garden Testing Literature

The preceding discussion was intended to provide a context and framework within which to introduce literature focused on describing phenotypic variation in conifers as detected in common garden (provenance) trials (Table 8.2). While no doubt incomplete, the listing here broadly covers recent, topical, and seminal studies in the field. In most cases, the objectives or results of the study are briefly summarized, and the intended applications are noted. Studies are organized by experimental approach and species being studied.

Amount, Distribution, and Pattern of Genetic Variation in Phenotypic Traits of Conifers

Conifers in general exhibit a great deal of genetic variation for frequently observed phenotypic traits. This is particularly true for widely dispersed species that inhabit highly heterogeneous environments. As noted in Fig. 8.1, genetic variation exists at many levels, but those of provenance (population) and parent tree (family) are of greatest interest, from the standpoint of both natural and artificial selections. They are also the factors most easily measured in common garden studies and dissected using common statistical approaches such as ANOVA, regression, and correlation. The amount and distribution of genetic variation in conifers is very briefly discussed below before looking, in much greater depth, at how that variation is patterned on the landscape.

Amount and Distribution of Genetic Variation

The amount and distribution of genetic variation of phenotypic traits are highly variable and influenced by factors such as the species and trait being studied, research methods used, test environments, and geographic range of samples being evaluated. The amount of genetic variation detected is also a function of the heritability of a trait, and heritability estimates are, in part, driven by the precision of the test and its ability to distinguish between environmental and genetic effects. Short-term trials that carefully control environmental conditions and study adaptive traits that manifest very early in the life cycle typically have relatively high estimates for observed genetic variation and heritabilities (Campbell 1986; St. Clair et al. 2005). For example, the average percentage (and range) of total variance attributed to populations (provenances) and families within populations in a short-term nursery trial of *Pseudotsuga menziesii* (Douglas-fir) was 20.5% (2.9–48.5%) and 15.3% (6.6–26.1%), respectively, for 12 seedling growth and phenology traits measured at ages 1 and 2 (St. Clair et al. 2005; Table 8.3). When traits were combined in a canonical analysis, the proportion of variation accounted for by genetic factors increased to 55.3% and 37.9% for the first and second canonical variables, respectively. The values noted here are quite high however and reflect, in part, the design of such

Table 8.2 A compendium of conifer common garden studies conducted, with some exceptions, over the last 30 years, presented by experimental approach, study application or purpose, and species of interest. Dependent and independent traits studied as well as a summary of results are noted

Application	Species	Dependent trait	Independent trait	Result	Citation
Experimental approach: short-term trials in artificial environments					
Selection of best sources in native range	*Pinus canariensis*, *P. pinaster*, *P. pinceana*	Drought-related traits: cavitation, hydraulics, root biomass	Climatic	Genetic variation in drought resistance traits low (2–5%); high levels of genetic-by-test site interaction and phenotypic variability (*P. canariensis*) but low levels of phenotypic variation in *P. pinaster*. Ecotypic variation noted in *P. pinceana* attributed to isolated populations	Lopez et al. (2013), Lamy et al. (2014), Cordoba-Rodriguez et al. (2011)
	Pinus pinaster	Biochemical traits affecting stem form	None	Determined variation among provenances in biochemical traits can be used as indicators for early selection of form	Sierra-DeGrado et al. (2008)
	Taxodium distichum	Growth, survival, tolerance to saline conditions	Soil/water salinity	Significant variation among populations in growth and foliar ion content under varying site conditions	Krauss et al. (2000)
	Taxus brevifolia	Chemical traits, growth, growth rhythm	Physiographic	Significant variation in all traits among and within populations; bud flush clinal but all other traits distributed randomly; epigenetic and environmental factors significant	Wheeler et al. 1992 (1995)
Selection of best sources in non-native range	*Abies lasiocarpa*	Growth, growth rhythm, needle color, branch traits	Non-native Christmas trees	Trees from Yukon to Arizona tested in Denmark, Iceland, and Norway; high survival, significant provenance variation, and large provenance-by-test site interactions	Hansen et al. (2004)

(continued)

Table 8.2 (continued)

Application	Species	Dependent trait	Independent trait	Result	Citation
	Pinus ponderosa	Seed and seedling traits	Physiographic	Early results of a long-term study established on many sites outside the native range; variation discontinuous for most traits; genetic variation described as ecotypic and varietal	Wells (1964)
	Pinus ponderosa and *Pseudotsuga menziesii*	Growth, growth rhythm	None	Studied provenance variation in nursery trials to establish origin of existing plantations growing in Argentina and to identify superior sources for afforestation	Rehfeldt and Gallo (2001), see also Wright et al. (1971)
Systematics, conservation, restoration	*Chamaecyparis thyoides*	Growth, winter damage, survival, growth rhythm	Physiographic	Test site variation exceeded provenance variation; growth patterns ecotypic but phenology mildly clinal; southern sources cold intolerant	Mylecraine et al. (2005)
	Cupressus species	Growth, growth rhythm, tree form	Demographics, native ranges	Assessed genetic structure of seven native *Cupressus* species; structure largely reflects uncertain, chance events that are dispersed throughout evolutionary history	Rehfeldt (1997)
	Pinus washoensis	Growth, growth rhythm	Physiographic	Washoe pine and related members of the *Ponderosae* share similar genetic structures, and variation is distributed along similar physiographic (elevation) clines	Rehfeldt (1999)
Define seed/ breeding zones and transfer guidelines	*Picea glauca*	Growth and growth rhythm	Physiographic	Modest levels of variance explained predominantly by longitude, a surrogate for precipitation and temperature clines; sources from warm southern provenances grow best; short-term trials mimic results of several older trials	Lesser and Parker (2004)

	Picea mariana	Bud break	Physiographic	Bud break occurred earlier and more quickly in northern provenances; provenance explained more variation than family within provenance; no elevational effect detected	Rossi and Bousquet (2014)
	Pinus contorta	Growth, growth rhythm, shoot elongation, freezing tolerance, seed germination	Physiographic and climatic	Clinal patterns of variation that generally reflect elevational and geographical gradients of frost-free periods and, to a lesser extent, moisture	Rehfeldt (1983, 1988), Sorensen (1992)
	Pinus halepensis	Water use efficiency (WUE), xylem native embolism (XNE), growing season length (GSL)	Climatic	WUE increased on dry test sites—provenances from dry sites had increased WUE the most; GSL decreased clinally from moist to dry test sites	Klein et al. (2013)
	Pinus monticola	Growth, growth rhythm, cold hardiness	Physiographic and ecotypic	In a range-wide study, variation found within but very little among populations; northern sources (Washington and Oregon) grow well, with modest cold hardiness, while southern sources (California) grow slowly, but are hardier; a large transition zone exists between them	Rehfeldt et al. (1984)
	Pinus patula	Growth	Physiographic	Sources from low elevations performed best over all sites	Saenz-Romero et al. (2011)

(continued)

Table 8.2 (continued)

Application	Species	Dependent trait	Independent trait	Result	Citation
	Pinus ponderosa	Growth, growth rhythm, emergence	Physiographic	Variation among populations typically large and distributed along relatively steep elevational clines and gentler geographic clines, interacting with slope and aspect	Rehfeldt (1990, 1991), Sorensen (1994)
	Pinus pseudostrobus, *P. devoniana*, and *P. leiophylla*	Growth	Physiographic	Clinal variation for *P. pseudostrobus*, but not for other species, with low elevation sources growing best over all sites in an elevational gradient	Castellanos-Acuna et al. (2013), Saenz-Romero et al. (2012)
	Podocarpus totara	Growth, stem form, branch length	Physiographic	Growth traits negatively correlated with latitude, presumably a surrogate for temperature	Bergin and Kimberley (1992)
	Pseudotsuga menziesii	Growth, growth rhythm, cold hardiness, seed emergence, partitioning of growth	Physiographic	Strong clines across elevational gradients within a single watershed; patterns varied by traits	Campbell (1979)
				Strong clinal variation associated with elevation and distance from the ocean across environmentally heterogeneous regions; procedures for using mapped genetic variation to develop seed zones and seed transfer rules	Campbell (1986, 1991), Campbell and Sugano (1993)
				Variance among populations low; distributed into provinces defined by temperature for interior variety	Rehfeldt (1978)

				Variation distributed clinally with distance (west to east) from crest of the Cascade Range and with elevation	Sorensen (1979)
			Physiographic, climatic	Considerable genetic variation among populations for all traits; bud set, emergence, fall cold hardiness, and growth clinally related mostly to elevation and cold season temperatures; bud burst and biomass partitioning mostly related to latitude (summer temperatures) and summer drought	St. Clair et al. (2005), St. Clair (2006)
				For coastal variety, variation in bud burst timing largely an adaptation to summer drought; variance in growth most influenced by minimum winter temperatures; test environment strongly influences phenological traits	Gould et al. (2011, 2012)
	Thuja plicata	Growth, growth rhythm, cold damage	Physiographic	Statistically significant differences among populations, but clines were very gentle; seed zone bands 600 meters and two degrees latitude recommended	Rehfeldt (1994a, b)
Climate change/assisted migration guidelines	*Pseudotsuga menziesii*	Growth, growth rhythm, emergence and biomass partitioning	Physiographic and climatic, climatic models	Models based on known patterns of adaptive variation and climate models estimate likelihood of maladaptation for populations under climate change scenarios, based on numerous previous studies; significant maladaptation anticipated; source movement recommendations given	St. Clair and Howe (2007)
	Pseudotsuga menziesii and *Pinus ponderosa*	Growth, growth rhythm	Physiographic and climatic, climatic models	Models based on known patterns of adaptive variation and climate models estimate likelihood of maladaptation for populations under climate change scenarios, based on numerous previous studies; predicted impacts of climate change greater for inland varieties, more so for ponderosa pine than for Douglas-fir.	Rehfeldt et al. (2014a, b, c)

(continued)

Table 8.2 (continued)

Application	Species	Dependent trait	Independent trait	Result	Citation
Experimental approach: long-term trials in field tests					
Identify best sources in native range	*Agathis* spp.	Growth, survival, silvicultural traits		Describes efforts to use provenance trials to illuminate how to grow trees in nurseries and plantations	Bowen and Whitmore (1980)
	Cupressus nootkatensis (syn. *Callitropsis nootkatensis*)	Growth, survival, and cold hardiness	Geographic	Significant region and provenance (population) within region variation for growth but not adaptability as measured by survival and cold hardiness, at individual test sites, but significant site-by-population interactions; most GxE attributed to marginal populations	Russell and Krakowski (2012)
	Picea abies	Growth, growth rhythm	None	Provenance performance relatively stable across planting sites in Denmark and Sweden but clone within provenance GxE observed	Karlsson et al. (2001)
				Unique application of an older provenance trial to reconstitute the genetic makeup of a superior provenance in the form of a select seed orchard	Chalupka et al. (2008)
	Picea glauca	Growth, growth rhythm	None	Family within provenance and provenance variances generally small (2–4% of variation) for growth and bud set, but curiously lacking for bud burst; height correlated with bud set	Li et al. (1993)
	Picea rubens	Winter photosynthetic rates	None	Provenance variation existed; winter photosynthetic rates approached those seen during the growing season in some populations	Schaberg et al. (1995)
	Pinus brutia	Growth, survival, stem form	Physiographic	Mid-elevation provenances performed best on all test sites in Turkey; provenance and family within provenance each accounted for ~10% of total variance	Isik et al. (1999, 2000)
	Pinus cembra	Growth, branch traits	Physiographic	Considerable provenance and provenance-by-test site interaction variation for all traits measured	Blada and Popescu (2007)
	Pinus kesiya	Growth, stem form, branching traits, foxtailing		Large and statistically significant differences among provenances for all traits; provenance-by-test site interactions mostly nonsignificant	Costa e Silva and Graudal (2008)

	Pinus pinaster	Growth and stem form	Ecotypic/ regions	Study concentrated on genotype by environment interactions in Portugal; GxE noted in all traits as well as significant variation among and within provenances	Correla et al. (2010)
		Growth, stem form, and insect resistance	None	Study used provenance variation to guide a "provenance hybridization" breeding and testing trial looking for heterosis	Harfouche et al. (1999)
	Pinus pinea	Growth, survival, phase change	Ecotypic/ regions	Studies vary in scope and objective; soil type of provenance may influence subsequent performance elsewhere (Provence, France); across the Mediterranean region, little provenance variation was observed for adaptive traits—phenotypic plasticity was observed	Court-Picon et al. (2004), Mutke et al. (2010)
	Pinus ponderosa	Growth, survival, stem form, WUE, Carbon isotope discrimination	Ecotypic/ regions	A wide range of provenance trials established across the mid-West and Lakes state regions of the USA, largely with the goal of identifying best provenances for planting in those regions	Cregg et al. (2000), La Farge (1975), Read (1980), Wright et al. (1969)
	Pinus resinosa	Growth, survival	Physiographic	Longitudinal clinal variation in growth, with western sources doing best, maritime sources worst, at test site in NW Ontario (test age = 48 years)	Rahi et al. (2010)
	Pinus sylvestris	Growth, biomass partitioning, filled seed counts, terpenoid and flavonoid concentrations	Ecotypic, physiographic, age	A sampling of large array of provenance trial publications on Scots pine, here ranging in age from 5 to 90; in Poland, Central European provenances grow best; northern provenances have a greater percentage of filled seed per cone (fewer lethal alleles) than southern sources; age does not affect provenance variation in flavonoid concentration; in Bulgaria, historical biogeographic forces appear to dictate patterns of provenance variation	Karkkainen et al. (1996), Manninen et al. (2002), Naydenov et al. (2005), Oleksyn et al. (1999, 2000), Oleszek et al. (2002), Savva et al. (2002)
	Pinus species of Mexico and Central America	Growth and survival mostly	Geographic	A compendium of chapters on provenance trials covering 11 native pine species in Mexico and Central America	CAMCORE Cooperative (2000)

(continued)

Table 8.2 (continued)

Application	Species	Dependent trait	Independent trait	Result	Citation
	Pseudotsuga menziesii	Growth	Physiographic	Near rotation-age, coastal range-wide trials in British Columbia exhibit low (1–6% of total) provenance variation, and large variation among blocks, sites and residual; high elevation sources consistently worst performers; results support extending seed source movement across a latitudinal range wider than currently accepted	Krakowski and Stoehr (2009)
Identify best sources in non-native range	*Picea abies*	Wood density, volume, ring width	Geographic	Several trials, some exceeding 50 years of age, evaluating Norway spruce in Canada; most variation found within provenances, but minor differences between southwestern and northeastern European provenances	Blouin et al. (1994)
	Pinus elliottii	Volume	Geographic	Evaluates southern US provenances in Brazil and Argentina; low type B correlations between US and South American test sites, but high type B correlations among South American sites	Rockwood et al. (2001)
	Pinus oocarpa	Volume, stem form, branch traits	Geographic	Selection opportunities among and within provenances, but provenance by site and family-by-site interactions are significant when *P. oocarpa* is planted in Brazil	Moura et al. (1998)
	Pinus radiata	Growth	Geographic	Study identified best provenances to establish on earthquake ravaged dry river valley in Sichuan, China	Bi et al. (2013)
		Growth, stem form, branch traits	Geographic	Specific provenances were generally best on all test sites in Australia though high provenance-by-site interactions were observed	Gapare et al. (2012)
	Pinus taeda	Growth and disease resistance	Geographic	A review of efforts to move seed sources to nonlocal sites to capture excellent growth and disease resistance, using improved families and provenances	Lambeth et al. (2005)

Systematics, conservation, restoration	*Cupressus sempervirens*	Terpene levels	None	Chemotypes observed among provenances used to identify sources requiring attention for conservation purposes	Gallis et al. (2007)
	Pinus nigra	Terpene levels	None	Chemotypes used to predict phylogenetic relationships among races or ecotypes	Rafil et al. (1996)
	Pinus ponderosa	Varied	Geographic		Conkle and Critchfield (1988), Potter et al. (2015)
Define seed/ breeding zones and transfer guidelines	*Larix occidentalis*	Growth, growth rhythm, disease resistance	Physiographic and biotic environment	Determined genetic correlations among growth, growth rhythm, and disease incidence, in relation to biotic and abiotic environmental factors	Rehfeldt (1992)
	Picea abies	Cold hardiness	Physiographic	Most variation in cold hardiness due to test site; acclimation played greater role than selection on provenance performance	Gömöry et al. (2010)
	Picea glauca	Methods paper	Physiographic and climatic	Combined a biophysical site index model with a seed source transfer model based on temperature and precipitation to estimate plantation yield in current and future climates	Beaulieu and Rainville (2005)
	Pinus densiflora	Growth	Physiographic and climatic	Study focused on provenance-by-site interactions in Korea; significant interactions noted but driven by relatively few provenances	Kim et al. (2008)
	Pinus halepensis	Cone serotiny	Climatic	Determined the trait is highly heritable and under selection; expression related to temperature regime at provenance origin	Hernandez-Serrano et al. (2014)
	Pinus pinaster	Growth and stem form	Physiographic and climatic	Provenance differences associated with major climatic and geographic features in Spain	De La Mata and Zas (2010)

(continued)

Table 8.2 (continued)

Application	Species	Dependent trait	Independent trait	Result	Citation
	Pinus sylvestris	Growth	Physiographic and climatic	Study assessed genetic response to climate variables and developed population-specific response functions driven by those variables in a large and mature provenance trial; appears a majority of populations occur at locations that are suboptimal for growth for those sources	Rehfeldt et al. (2002)
	Podocarpus totara	Growth	Physiographic and climatic	Large differences among provenances associated with latitude and summer temperatures	Bergin et al. (2008)
	Pseudotsuga menziesii	Cold and drought hardiness	Climatic	Populations varied threefold in cold hardiness; driven primarily by winter minimum temperatures and date of first fall frost; to a lesser extent by summer temperature and aridity of source locations (warmer, drier sites confer some cold hardiness); trees from warm, dry locations and from cool, high elevation locations had greater drought tolerance	Bansal et al. (2015a, b, 2016)
Climate change	*Larix occidentalis*	Growth	Physiographic and climatic	Study concluded that long-term provenance trials can be used to infer forest species growth responses to anticipated climate change	Leites et al. (2012)
		Growth, growth rhythm, disease tolerance	Physiographic and climatic	Study accurately predicted current species range based on climate variable models; models used to predict the future range of the species based on anticipated climate changes	Rehfeldt and Jaquish (2010)

	Picea abies	Growth	Climatic	The study sought to identify best provenances for future climatic conditions in Austria by using climate-response functions from 15-year-old provenance trials	Kapeller et al. (2012)
	Picea mariana	Growth and survival	Climatic	Used population-response functions, site-transfer functions, and universal response functions (Yang) to predict effects of climate change on likely survival and growth of black spruce across Ontario and the Great Lakes region; southernmost sources likely to suffer in warmer conditions	Thomson et al. (2009), Yang et al. (2015)
	Pinus contorta	Growth, growth rhythm, survival	Climatic	A complex, sophisticated analysis of genetic response to climatic factors and anticipated climate change patterns; results suggest provenances have varying climatic optima and often occupy sites that afford suboptimal growth opportunities, driven by both environmental and density dependent selection forces	Rehfeldt et al. (1999)
		Growth	Climatic, geographic	Latter paper introduces the Universal Response Function (URF); analytical approach permits simultaneous estimates of environmental and genetic effects; mean annual temperature (MAT) accounted for 73% of variation in height growth, environmental effects ≫ genetic effects; for provenances from MAT <4.5 °C, a one-degree increase in MAT will improve growth, but for provenances from MAT >4.5 °C, it will decrease growth; applications of URF are enumerated	Wang et al. (2006a, 2010a)
		Radial annual growth rings	Climatic	Study seeks to predict climate change impacts based on measurements of response in radial annual growth rings to seasonal and annual climate variables in a mature, well-replicated provenance trial in British Columbia; provenances from cool locations will likely respond well to warming climates while those currently in warmer environments likely will suffer	McLane et al. (2011)

(continued)

Table 8.2 (continued)

Application	Species	Dependent trait	Independent trait	Result	Citation
	Pinus patula and *P. tecunumanii*	Growth and survival	Climatic	A pair of papers that use existing provenance test results to estimate climate change effects on species ranges and productivity using climate envelope models (CEM); concluded species are currently doing well in CEM predicted to be unfavorable; fragmentation likely to have greater negative impact on these species than climate change	Leibing et al. (2009), van Zonneveld et al. (2009)
	Pinus ponderosa and *Pseudotsuga menziesii*	Growth potential	Climatic	A series of papers that synthesized data from a broad range of provenance trials for multiple varieties of two species; range-wide climate-based clines in growth traits described statistically, thereby converting clines to climatypes; climatypes were broader for ponderosa pine than for Douglas-fir, and broader for interior than coastal varieties of either species; both species will require human intervention to insure future reforestation efforts are done with adapted seed sources	Rehfeldt et al. (2014a, b, c)
	Pinus strobus	Many	Physiographic and physiographic and climatic	Redefined clines in traits in terms of climatic variables rather than physiographic variables (i.e., latitude); climate modeling methods used to identify realized niches and predict future niche congruity; *P. strobus* would likely suffer in predicted future climates	Joyce and Rehfeldt (2013)

		Growth	Climatic	Used URF models to guide future planting decisions. Climate warming expected to enhance *P. strobus* growth for provenances from sites colder than MAT 11 °C.	Yang et al. (2015)
	Pinus taeda and *Picea abies*	Growth	Climatic	Results of mature provenance trials for both species used to predict effects of anticipated climate changes on growth in native ranges; yearly average minimum temperatures at provenance source best predictor of transfer success	Schmidtling (1994)
	Pseudotsuga menziesii	Growth	Climatic	Using a North American CEM compiled from a large array of provenance trials, authors predict best provenances for planting in future European conditions	Isaac-Renton et al. (2014)
	Multiple	Primarily growth and growth rhythm	Climate modeling	Review papers that predict genetic responses of forest trees in temperate and boreal forests of North America and Europe to anticipated climate change conditions based on results from myriad provenance trials; recommendations made for future use of common garden studies	Rehfeldt et al. (2004), Alberto et al. (2013)
Assisted migration	Multiple	Adaptive traits	Climate modeling	Selected citations defining the concept of "forestry"-assisted migration, its applications, and desired reliance of common garden trials and climate envelop modeling for guidance	McLane and Aitken (2012), Pedlar et al. (2012), Ukrainetz et al. (2011), Williams and Dumroese (2014)

Table 8.3 Amount and distribution of genetic variation for seedling traits in a short-term genecology trial of *Pseudotsuga menziesii*

	Percentage of total variance		
Trait	Among provenances	Among families within provenances	Error
Shoot weight	16.8	18.9	64.3
Root weight	14.3	18.5	67.1
1st year height	13.3	26.1	60.5
2nd year height	17.9	13.4	68.7
Height increment	17.4	8.9	73.8
Diameter	17.1	18.6	64.3
Root-shoot ratio	8.9	6.6	84.5
Bud set year 1	36.3	14.6	49.1
Bud set year 2	18.6	7.9	73.5
Bud burst year 2	34.5	21.9	43.6
Rate of seed emergence	48.5	21.1	30.4
Llamas growth	2.9	7.5	89.6
Canonical trait 1	55.3	10.6	34.1
Canonical trait 2	37.9	13.7	48.4

Modified from St. Clair et al. (2005)

studies where emphasis is placed on sampling many populations (often hundreds), covering most of the species' range, and relatively few trees (families) within populations (often only one or two).

More commonly, the proportion of variation observed that can be attributed to provenances and families (1–10%) is significantly less than noted above and often shifts, from one study to another, between which factor contributes more than the other (Isik et al. 1999, 2000; Li et al. 1993; Wheeler et al. 1995a; Krakowski and Stoehr 2009; Lopez et al. 2013). The distribution may be dependent on species and trait studied, age of test, sampling method, and so forth. For instance, growth (height, diameter, volume), measured in older tests, typically exhibits far less variation among provenances than that measured in short-term trials. Long-term trials generally sample fewer provenances and many families within provenances, in part because it increases the opportunity to identify exceptional sources for deployment. In some species, racial variation is low, and virtually all genetic variation detected is among families within provenances across large portions of a species' range (e.g., *Pinus monticola* and *Thuja plicata*, Rehfeldt 1994a, b; *Pinus resinosa* and *Thuja occidentalis*, Wright 1976). Likely the most important factor affecting proportion of variation attributable to population is the range of environment samples.

Patterns of Variation

Patterns of genetic variation for phenotypic traits in natural populations are many and varied but are frequently displayed as clines reflecting adaptation of annual growth and dormancy cycles to environmental gradients, most notably temperature

Table 8.4 Elevational and climatic distances needed to detect genetic (racial) differences in conifer species native to the Northern Rocky Mountains, USA

Species	Elevation (m)	Number of frost-free days	Evolutionary mode
Pseudotsuga menziesii (Douglas-fir)	200	18	Specialist
Pinus contorta (lodgepole pine)	220	20	Specialist
Picea engelmannii (Engelmann spruce)	370	33	Intermediate
Pinus ponderosa (ponderosa pine)	420	38	Intermediate
Larix occidentalis (western larch)	450	40	Intermediate
Thuja plicata(western red cedar)	600	54	Generalist
Pinus monticola (western white pine)	None	90	Generalist

Modified from Rehfeldt (1994a)

and precipitation. More specifically, the environmental gradients most affecting patterns of adaptation are measures of minimum temperature (average monthly, winter, or annual minimum temperature), maximum temperature (average summer temperature, average maximum temperature of the warmest month), number of annual frost-free days, precipitation, and indices of drought. White et al. (2007) point out that seed zones delineated for four southern pine species in the USA, based on long-term provenance studies, mirror the US Department of Agriculture's Plant Hardiness Zones based on average minimum temperatures.

It follows that for most conifer species, especially those growing in highly heterogeneous environments, provenances (populations) are at least moderately adapted to the local climatic conditions, though the patterns and extent of adaptation may vary appreciably. A summary of studies delineating elevational races in conifers of the Northern Rocky Mountains of the USA indicates the extent to which species vary in sensitivity to the number of frost-free days at the location of provenance tested (Rehfeldt 1994a; Table 8.4). The author classifies species as being specialists, intermediates, or generalists with respect to the distance, in meters of elevation, required to detect genetic differences in growth and growth rhythm traits, ranging from as little as 200 m to 600 m or more (Table 8.4). For species considered to be specialists, seed transfer rules would necessarily be restrictive while for generalists, highly permissive. This type of information would be helpful in guiding the design of a tree improvement program or assisting in managing of climate change effects on the health and adaptability of future forests.

For most studies conducted with wide-ranging Northern temperate and boreal species, the manifestations of clinal patterns of variation follow similar trends, especially with respect to temperature gradients. Thus, trees from colder locations (e.g., high elevations, northern latitudes, distance from the ocean), when grown at warmer test sites, tend to grow more slowly, flush earlier in the spring, set buds earlier in the fall, and are more resistant to cold damage from early fall cold events or extreme winter cold events than sources from warmer areas. They flush earlier because they

require fewer heating degree days to force flushing. The corollary is that provenances originating from areas with favorable growing environments (i.e., warmer) tend to grow faster at most test sites, except for those in the most extreme locations. Similarly, trends occur along moisture gradients, with trees from drier provenances growing more slowly, and investing more resources in belowground biomass, than trees from moister origins. Wright (1976) notes that trees from dry sites tend to have larger seeds and bluer foliage than those from moist sites. Foliage color in conifers is often a function of the accumulation of epicuticular waxes on the needles which may be as much a response to increased solar radiation as it is to moisture.

Exceptions to most general patterns occur. For instance, many studies report a lack of significant clinal variation in traits studied (e.g., Wheeler et al. 1992, 1995a; Isik et al. 1999; Oleszek et al. 2002; Costa e Silva and Graudal 2008; Mutke et al. 2010; Castellanos-Acuna et al. 2013). In some cases, clinal variation exists for traits like growth, but different species exhibit opposite trends. For example, in eastern North America, the best growing provenances of *Picea glauca* originate in eastern, maritime provenances and decline in growth as they move west and north, into more continental climes (reviewed in Beaulieu and Rainville 2005). The opposite trend exists for *Pinus resinosa*, with western sources growing best and maritime sources least (Rahi et al. 2010).

Are Local Sources Best?

For most of the period covered by modern forestry practices, recommendations for which seed/seedling should be used for reforestation have strongly favored use of local sources, generally codified by the creation of seed zones specific to each species of conifer in a region. This policy was created to reflect the concept that tree populations are adapted to their local environments and risk of plantation failure was minimized by using local sources. Increasingly, common garden studies have revealed that local sources may not always be best and that some seed source movement may be acceptable or even desirable (Wells and Wakeley 1966; Namkoong 1969; Rehfeldt et al. 2002; Lambeth et al. 2005). Reasons for this are likely manifold, almost certainly differ for each case studied, and are dependent on how the term "best" is defined. A well-documented case in forestry is that described for *Pinus taeda* in the southeastern USA (Wells and Wakeley 1966; Zobel and Talbert 1984). Sources from the western portion of the species range possess resistance to a damaging rust that greatly diminished growth potential in eastern provenances. Transfer of western sources several hundreds of miles east have resulted in significant growth increases. More generally useful however are studies that seek to match optimal growth environments with growth potential. Carefully controlled common garden trials that have characterized population response functions relative to climate variables have revealed, in many cases, that local populations are growing in locations that are suboptimal for their growth (Rehfeldt et al. 1999, 2002, 2004). The results of these tests suggest that "populations inhabit climates colder than their optima, with the disparity between the optimal and inhabited climates becoming

greater as the climate becomes more severe" (Rehfeldt et al. 2002). In such cases, a warming climate may improve growth potential for some populations and species in the near term but would eventually have a detrimental effect on overall species adaptation and growth (McLane et al. 2011). The consensus, however, is that climate change will have significant and negative effects on conifer survival and growth in the coming decades and centuries, requiring increasing human intervention in the form of assisted migration to assure continued adaptation to changing conditions (see citations under climate change in Table 8.2).

Case Studies

Variation among provenances has been observed for the large majority of conifer species evaluated in common garden trials, regardless of the traits studied or experimental approaches employed. As noted previously, many of these studies have been enumerated in a few comprehensive summaries published over the years. The concepts and findings of these studies have, so far, only been discussed in very general terms here. In this section we will discuss, in much greater detail, the results of a select group of studies, with a focus on describing patterns of genetic variation observed and identifying the interplay of environmental factors and evolutionary forces molding the distribution of variation on the landscape. Attention is given to well-studied species and studies that seek to use results from common garden trials to guide reforestation or restoration plans to cope with current and/or anticipated future climates. The studies chosen for review represent both short-term (genecological) and long-term (field tests) provenance trials and showcase the evolution in methodologies developed to analyze and interpret common garden trial results.

Pseudotsuga menziesii (Douglas-fir)

Pseudotsuga menziesii is among the most thoroughly investigated conifer species in the world with respect to studies of geographic and climatic variation. A species native to western North America, *P. menziesii* ranges from British Columbia, Canada to Mexico (~4500 km), from sea level to over 3000 m, and from the Pacific coast to eastern slopes of the Rocky Mountains. Taxonomists recognize two varieties, one occupying the coastal regions west of the Cascade, Sierra Nevada, and Canadian coastal mountain ranges (*P. m.* var. *menziesii*) and an interior or Rocky Mountain variety (*P. m.* var. *glauca*). The species is the subject of several tree improvement programs and is planted extensively, both within its native range and as an exotic in Europe and New Zealand (Howe et al. 2006).

Common garden trials of *P. menziesii* began both in the USA (Munger and Morris 1936) and in Europe in the early 1900s. More recently (1960s to present), long-term replicated trials for one or the other varieties have been established, both in North America (Ching 1965; Wright et al. 1971; White and Ching 1985; Jaquish 1990; Sziklai 1990; Gould et al. 2011, 2012; Bansal et al. 2015a, b, 2016) and

abroad (Sweet, 1965; Breidenstein et al. 1990). These have been supplemented by a striking array of short-term nursery trials, the nature of which has strongly influenced modern approaches to the study of conifer genecology (Hermann and Lavender 1968; Rehfeldt 1977, 1978, 1979, 1989; Campbell and Sorensen 1978; Campbell 1979, 1986, 1991, Campbell and Sugano 1979; Sorensen 1979, 1983; St. Clair et al. 2005; St. Clair 2006). We have drawn from a number of these studies to illustrate genecological relationships, contrasts between short- and long-term trial results, and the use of provenance information to interpret plant responses to climate change.

For seed and seedling traits expressed in short-term nursery, growth chamber, and greenhouse trials, both interior and coastal varieties of Douglas-fir appear to be closely adapted to steep environmental gradients (e.g., elevation, aspect, latitude, distance from the ocean). Hermann and Lavender (1968) collected seed along a single transect (~25 km long), at 500-foot elevational intervals, between 1500 and 5000 feet, and from north- and south-facing slopes on the western flank of the Cascade Mountains in Central Oregon, USA. Seeds were grown in nursery beds at two sites and measured for growth and phenology traits. They found racial differences, associated with both elevation and aspect, for most traits, including timing of bud burst and bud set (Fig. 8.4). Seedlings from lower elevations broke buds earlier and set buds later than those from higher elevations, and seedlings from southern aspects set buds sooner than those on northern slopes. Trees from northern slopes grew larger and had higher root-shoot ratios than those from southern aspects. The responses were interpreted as adaptation to cold (elevation) and drought (aspect) gradients.

A more intensively sampled study conducted with seed collected from a single watershed in the central Cascade Mountains of Oregon, USA, revealed similarly complex patterns of growth and phenology with regressions of traits on elevational gradients confounded by aspect (Campbell 1979). Based on the levels of genetic differentiation observed, the author estimated that a transfer of seedlings from as little as 3.5 km apart, at the same elevation but from different aspects, could lead to ~80% of the seedlings being poorly adapted in the new environment. The results demonstrate that populations can vary greatly as result of the conditions under which they have evolved. These and other studies were conducted in portions of the coastal *P. menziesii* range that were environmentally highly heterogeneous and climatically extreme, and their results led to relatively conservative recommendations for defining breeding zones and seed transfer rules for the regions studied.

Beginning in the 1990s, a series of larger studies based on comprehensive seed collections from throughout the coastal range of *P. menziesii* in Washington and Oregon, USA, were conducted using short-term or genecological (St. Clair et al. 2005; St. Clair 2006) and long-term (Gould et al. 2011, 2012; Bansal et al. 2015a, b, 2016) experimental approaches. For the short-term studies, wind-pollinated seed was collected from over 1300 parent trees in natural stands at 1048 locations (Fig. 8.5). Paired samples from 291 locations permitted estimates of family-within-provenance variance. In the first of these studies (St. Clair et al. 2005), seedlings were grown for 2 years in nursery beds in Corvallis, Oregon, USA, and measured

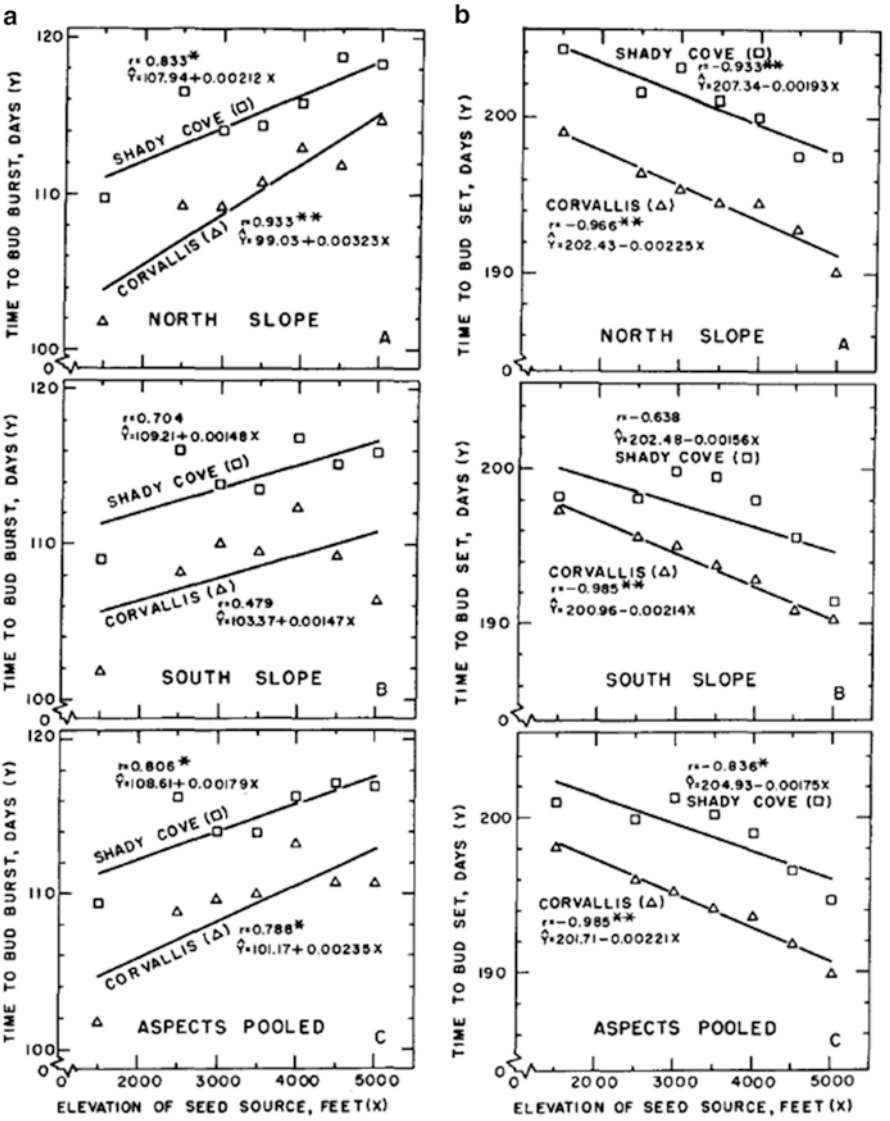


Fig. 8.4 Regressions of days from beginning of year to (**a**) bud burst and (**b**) bud set at two nurseries for 50% of *Pseudotsuga menziesii* seedlings of seven altitudinal origins (A) from north-facing aspects, (B) from south-facing aspects, and (C) for aspects pooled. * denotes statistical differences at the 5% level of probability. (From Hermann and Lavender 1968, pp. 143–151)

for 16 growth and phenology traits. All phenotypic traits were analyzed with ANOVA to estimate components of variance, and with regression and correlation analyses, and CCA to investigate the relationships between phenotypic and environmental traits.

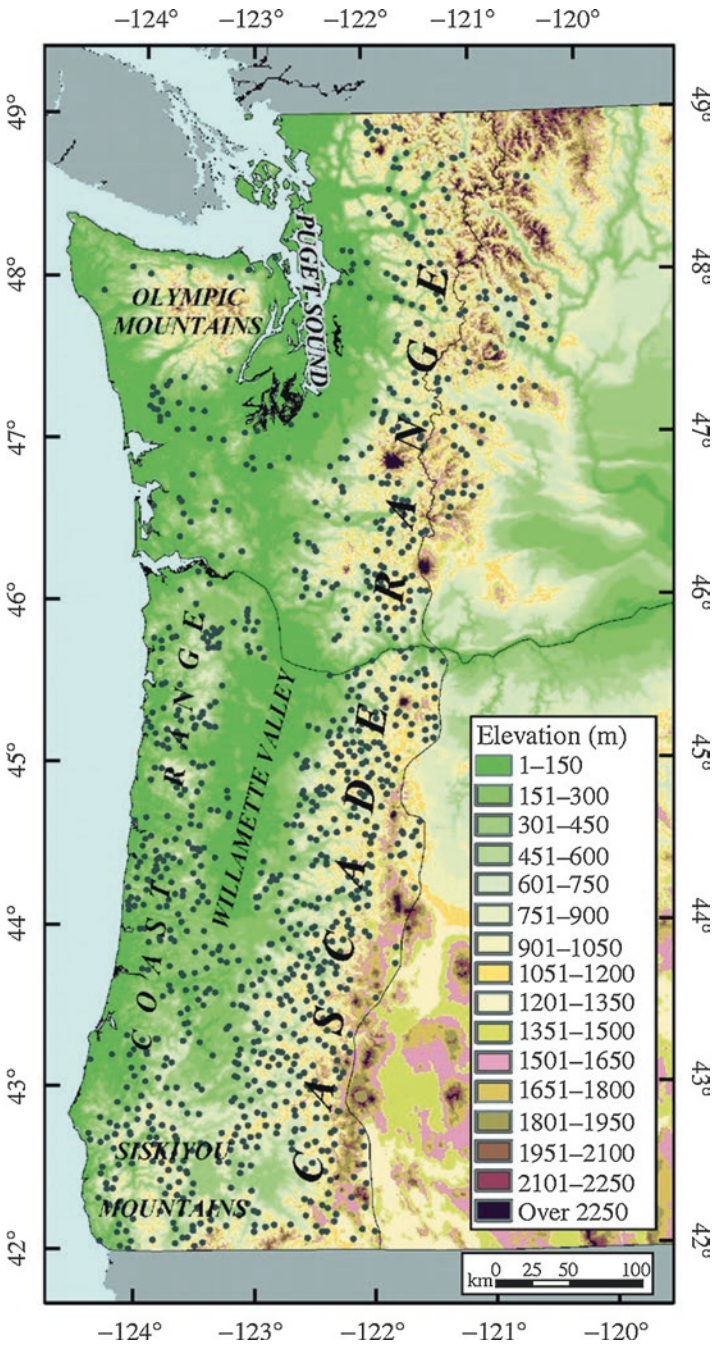


Fig. 8.5 Sample locations for over 1000 *Pseudotsuga menziesii* (Douglas-fir) populations evaluated in short-term nursery trials. (From St. Clair et al. 2005, p. 1201)

CCA revealed most traits measured were strongly related to the environment of seed source (canonical correlations for the first and second canonical variables were 0.82 and 0.70) and combined, the two variables accounted for 27% (20% and 7%, respectively) of the total variation in phenotypic data. The first canonical variable was most strongly related to temperature (late fall and winter minimums) and correlated traits like elevation and dates of first and last frosts, while the second canonical variable was most strongly related to summer precipitation, temperature, and aridity. The variables, when mapped, reveal clear patterns of racial variation (Fig. 8.6). Higher values of the first variable are related to vigor (i.e., later bud set, faster seed germination, larger seedling sizes, and increased shoot to root ratios). Higher values for the second canonical variable were related to earlier bud burst and greater partitioning to second-year diameter versus height. Contour intervals shown in Fig. 8.6 reflect a 30% level of risk of maladaptation from source movement. Overlaying the maps of the two canonical variables revealed additional insights into

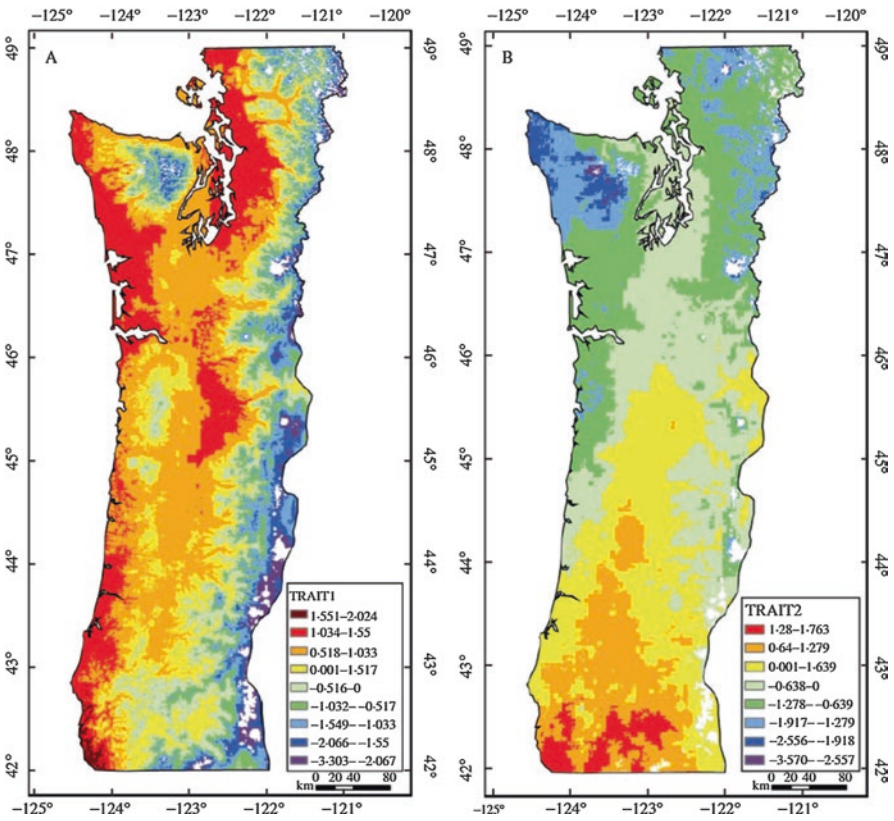


Fig. 8.6 Geographical variation in (**a**) the first and (**b**) second canonical variables. Mean values are shown as the zero contour between yellow and light green. Contour intervals represent a 30% level of risk of maladaptation from source movement in *Pseudotsuga menziesii*. (From St. Clair et al. 2005, p. 1207)

shared genetic types (St. Clair et al. 2005). The results suggest both elevation and latitude should be considered when stratifying the region into areas of similar genetic types. What is clear from these figures is that trees from relatively large areas, often quite distant from one another, may share quite similar genetic makeups with respect to well-defined trait arrays. Operationally speaking, the results shown here suggest that seed transfer guidelines and breeding zone delineation may be much more relaxed than implied by the earlier studies conducted in narrow geographies (Campbell 1979) but require that environments be thoroughly characterized using appropriate climatic and topographic variables.

Unlike earlier nursery-based studies that suggested *P. menziesii* populations were narrowly adapted over steep environmental gradients and were vulnerable to genetic-by-environmental interactions, results garnered from multiple long-term provenance trials of *P. menziesii* (>25-year-old data) and extensive field tests of tree improvement select families find provenance (and/or family)-by-site interactions to be modest and, when present, driven by only one or a few provenances or families (White and Ching 1985; Stonecypher et al. 1996; Krakowski and Stoehr 2009). Similar results have been found in provenance-by-site interaction studies for other species (Karlsson et al. 2001; Costa e Silva and Graudal 2008; Kim et al. 2008). How can such inconsistent findings be explained? It is important to recall the differences between short- and long-term testing (Table 8.1; Howe et al. 2006).

Short-term studies evaluate juvenile plant traits in artificial environments over very short durations. The amount and pattern of genetic variation observed in seedlings may vary substantially as the trees mature (White and Ching 1985). New seedlings are particularly vulnerable to environmental stresses as they tend to grow vigorously and enter dormancy much later in the fall than older seedlings. In natural stands, the clear majority of seedling mortality occurs within a year or two of establishment when trees are vulnerable and selection forces can be extreme. Long-term studies are established with much hardier, older seedlings that are less vulnerable to environmental stresses. These hardier trees may allow for their full genetic potential to be expressed over time, the ultimate manifestation of which is superior growth and survival. In brief, short-term seedling tests provide information on seedling tolerances and performance, and only indirectly inform decisions on seed-transfer guidelines and breeding zones, while long-term trials provide direct evidence for provenance performance in diverse, natural habitats.

To conclude this overview of common garden trials in *P. menziesii*, we shift to studies addressing climate change concerns. As noted in the previous studies reviewed here, coastal *P. menziesii* exhibits substantial adaptation to temperature and precipitation gradients. Similar clinal trends are demonstrated in extensive studies conducted on the Rocky Mountain variety of *P. menziesii* (Rehfeldt 1977, 1978, 1979, 1989). Climate change scenarios for most regions of the world, including the range of *P. menziesii*, are expected to include co-occurring changes in temperature, precipitation, and the extent and frequency of extreme weather events. To address concerns about the long-term adaptation of *P. menziesii* to anticipated climate change, a set of common garden trials, referred to as the Douglas-fir Seed-Source Movement Trial, were established using a reciprocal transplanting design.

Open-pollinated seed from multiple trees in each of the five populations sampled in each of the 12 regions was established, as 2-year-old seedlings, in 9 separate field trials spanning the extremes of warm/dry to cool/wet conditions in the range of coastal *P. menziesii*, from Washington to California, USA. Several investigations have used these trials to study cold and drought hardiness (Bansal et al. 2015a, b) and growth phenology (Gould et al. 2011, 2012).

To better understand the natural covariation in stress tolerance to multiple traits, the data from the cold and drought studies were combined in a separate analysis using PCA (Bansal et al. 2016). Trees from two families in each of 35 populations from across the range of the species were measured for drought hardiness at a warm, dry test site and for cold hardiness at a cool, moist test site. Drought hardiness was estimated based on measures of transpiration, specific leaf area, and water saturation deficit, while cold hardiness was estimated using freeze tests of needles, twigs, and buds. Using PCA, the data were combined into two generalized stress hardiness traits, HARDINESS 1 and HARDINESS 2.

Hardiness trait values were strongly related to seed source climate variables and explained 47% (HARDINESS 1) and 30% (HARDINESS 2) of the variation among populations in drought and cold hardiness traits. The mapped hardiness traits reveal a complex pattern of covariation that suggest multiple genetic mechanisms are likely at play (Fig. 8.7).

The authors found that drought and cold hardiness converged among populations along winter temperature gradients (elevation) and diverged along summer precipitation gradients (latitude). Notably, populations from regions with cold winters had relatively high tolerance to both drought and cold stressors, while populations from regions with warm and dry summers had increased drought tolerance but reduced cold hardiness. Co-tolerance to both stressors in the cooler, mountainous regions was suspected to result from common adaptive mechanisms (dealing with frost and winter desiccation, for instance), while the divergence in warmer, drier areas seemed to reflect fundamentally different physiological processes and trade-offs.

The results of these studies guide recommendations on source movement to mitigate deleterious effects of climate change (assisted migration). The presumptive recommendation prior to the revelation of covariance in drought and cold hardiness traits was that trees originating in current warm regions should be moved to current cold regions, in the future. While this may still be the case, it implies we should avoid moving from warm, wet conditions to cold, dry conditions.

While assisted migration may be viewed as a means to move sources within a species' natural range, or possibly to extend the natural range, to insure adaptation to future climates, the use of a species as an exotic must also be considered in this context. *P. menziesii* has been planted throughout Europe for over 100 years and has benefitted from extensive provenance testing during that time (reviewed in Isaac-Renton et al. 2014). Scientists developing bioclimate envelope models to guide assisted migration recommendations for Douglas-fir in its native range examined whether these models could retrospectively predict the success of provenance transfers in Europe (Isaac-Renton et al. 2014). The study's meta-analysis was based on long-term growth data of 2800 provenances established on 120 European test sites,

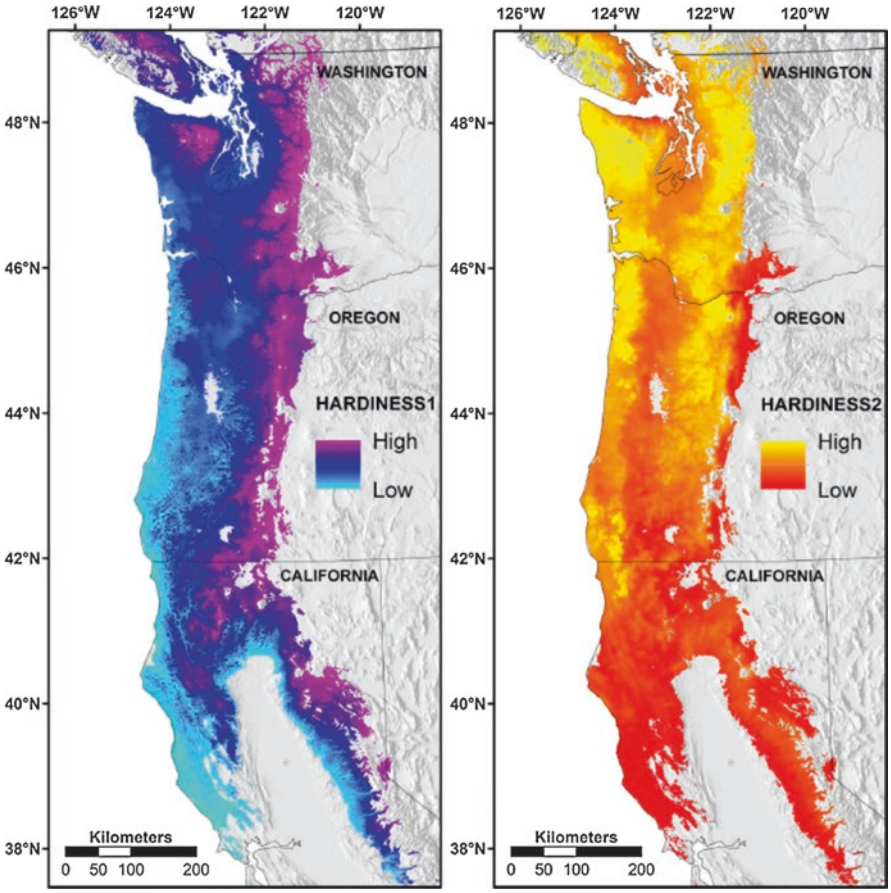


Fig. 8.7 Geographic variation in the first (HARDINESS) principal components of Douglas-fir (*Pseudotsuga menziesii*) in the Pacific Northwest, USA. The principal components were derived by combining drought and cold hardiness trait data from two common gardens. The components were then modeled using seed source climate variables. HARDINESS 1 values exhibited a strong longitudinal gradient associated with elevation and cool winter temperatures; HARDINESS 2 values exhibited a latitudinal gradient associated with summer precipitation. Higher values of HARDINESS 1 corresponded with greater drought and cold hardiness. Higher values of HARDINESS 2 corresponded with greater cold hardiness but reduced drought hardiness. (From Bansal et al. 2016)

the results of which were used to validate the climate envelope model projections for Europe. The correlation between observed provenance performance and climate model predictions were generally good for Western Europe, including France, Belgium, and Southern Germany, but the models were less productive, when considering only growth, for Central and Eastern European sites. The inclusion of cold and drought tolerance to the model improved results considerably. The effectiveness of the model predictions changed with the observed climate variable period used, noting the general warming trend observed over the last 30 years (see Fig. 8.8). The

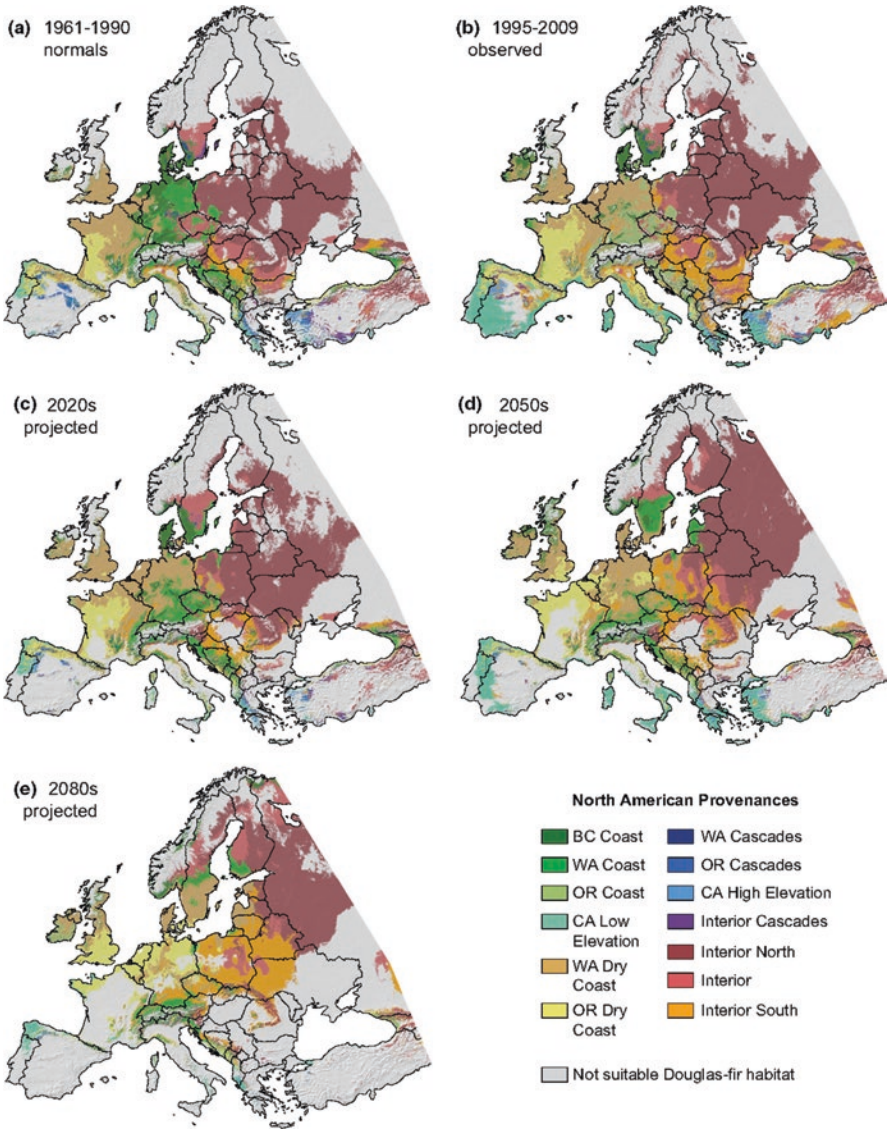


Fig. 8.8 Random forest predictions of suitable North American Douglas-fir provenances under 1961 to 1990 climate normal conditions, a recent 15-year climate average from 1995 to 2009, and ensemble projections for the 2020s, 2050s, and 2080s under the A2 emission scenario. (From Isaac-Renton et al. 2014)

study went on to recommend appropriate provenance selections for projected future climatic conditions (Fig. 8.8). Sources from wet, coastal areas of North America are gradually replaced by sources from more southern, dry coastal areas in Western Europe, while Rocky Mountain varietal sources are recommended for Eastern and Northern European locations.

Pinus

Studies of phenotypic variation in the genus *Pinus* are well represented by common garden and provenance trials (see Table 8.2) both within species' native ranges and as exotics in non-native habitats. Few conifer species have been more intensively and extensively studied than *P. contorta* (lodgepole pine) (Illingworth 1978; Sorensen 1992; Rehfeldt et al. 1999; Wang et al. 2010a) and *P. sylvestris* (Scots pine) (Giertych 1979; Oleksyn et al. 1999; Oleszek et al. 2002; Shutyaev and Giertych 2000; Rehfeldt et al. 2002), two of the most widely dispersed tree species in the world.

The natural range of *P. contorta* exceeds 33° of latitude and 3900 m of elevation, from the Pacific coast of North America to the Rocky Mountains and from Canada's Yukon Territory (64° N) to Central California in the Sierra Nevada Mountains. Four varieties or subspecies of *P. contorta* are recognized (Critchfield 1957), each occupying a distinct biogeoclimatic region (Wheeler and Critchfield 1985). Within the province of British Columbia, Canada, alone, the species (including subspecies *contorta* and *latifolia*) occupies a niche that "spans climates differing by >35 °C in mean temperature in the coldest month, 3 m in mean annual precipitation, and 180 days in the length of the frost-free season" (Rehfeldt et al. 1999). Both short- and long-term common garden and provenance trials have demonstrated that variation among populations within subspecies of *P. contorta* is large, arranged along steep altitudinal and latitudinal clines for adaptive traits, and is generally characterized by a strong negative relationship between growth potential and environmental harshness of the parent population provenance (Ying et al. 1985; Rehfeldt 1988; Xie and Ying 1995; and others).

Much of what has been learned comes from a landmark study established by the British Columbia Ministry of Forest's Research Division in the mid-1970s (Illingworth 1978). The study tested 140 natural populations (provenances), representing *P. contorta* subspecies *latifolia, contorta*, and *murryana*, planted on 60 replicated test sites distributed throughout the province. The results of comprehensive investigations of growth and survival at age 20 in this set of trials (Rehfeldt et al. 1999; Wang et al. 2010a) is reviewed here because of the innovative analytical approaches, data interpretations, and applications of these seminal studies.

In the first of these investigations, the study methods involved (1) the use of polynomial regression models using physiographic descriptors and known climatic conditions at 513 weather stations to predict 7 climatic variables for the remainder of the species range, including sampled populations; (2) developing population response functions describing the height or survival of each population using predicted climate variables for each test site; (3) developing general transfer functions; and (4) interpreting results in terms of niche breadth, effects of climate change on adaptedness of populations, and reforestation in a changing environment.

Most of the 118 populations studied by Rehfeldt et al. (1999) generated statistically significant response functions for 20-year height (89%), survival (82%), or both (70%). Climatic predictors most influencing height were mean annual temperature, mean temperature of the coldest month, and temperature differential

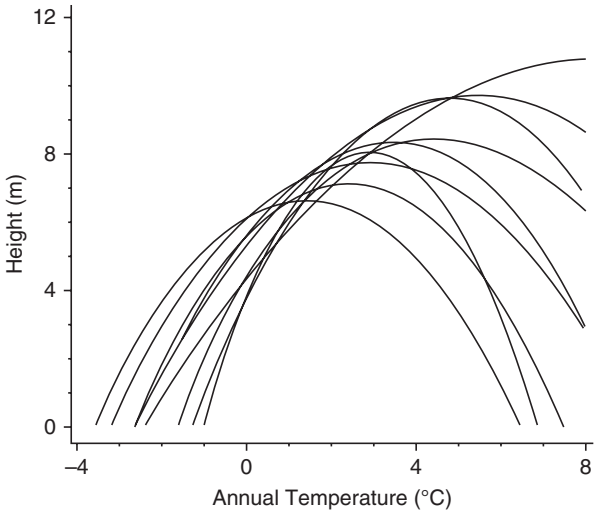


Fig. 8.9 Response functions using mean annual temperature as a predictor of height for nine populations that represent a variety of responses for *Pinus contorta* ssp. *latifolia*. (From Rehfeldt et al. 1999)

between coldest and warmest months. Survival was most influenced by the number of days without freezing temperatures and the summer heat: moisture index (an indicator of drought). A selected array of population response functions using mean annual temperature (MAT) as a predictor of height reveals much about the genetic variation, among and within populations, and niche characteristics of *P. contorta* (Fig. 8.9). For instance, the optimum climate for growth is identified as the apex of the response function, and the breadth of the function denotes the range in climates within which a population is predicted to survive and grow. Population response functions shown here (Fig. 8.9) illustrate large differences in growth potential and tolerance to cold and reflect clearly the negative relationship between growth potential and cold hardiness.

The breadth of the functions suggests that the fundamental niche of most populations (climates within which they can survive and grow) spans nearly 10 °C in MAT, reflecting both broad phenotypic plasticity and large within-population genetic variability though it must be noted that production may vary greatly across that environmental spectrum. Variation among populations is further illustrated by population response functions generated for pairs of populations for both height and survival, predicted using three different climatic factors (see Fig. 8.10).

Perhaps the most revealing finding of this study was that most of the populations studied here occupied suboptimal environments. That is, the climates of the locations from which the populations were sampled were colder (by 2 °C on average) than their response functions indicate would be optimal for growth. The discrepancy between occupied habitats and ecologically optimal habitats is accentuated as the climate becomes more severe (colder, drier). The authors propose that this discrepancy results from a combination of environmental selection, which produces broadly adapted populations, and density-dependent selection, which produces a relatively narrow realized niche within which most populations are relegated to suboptimal

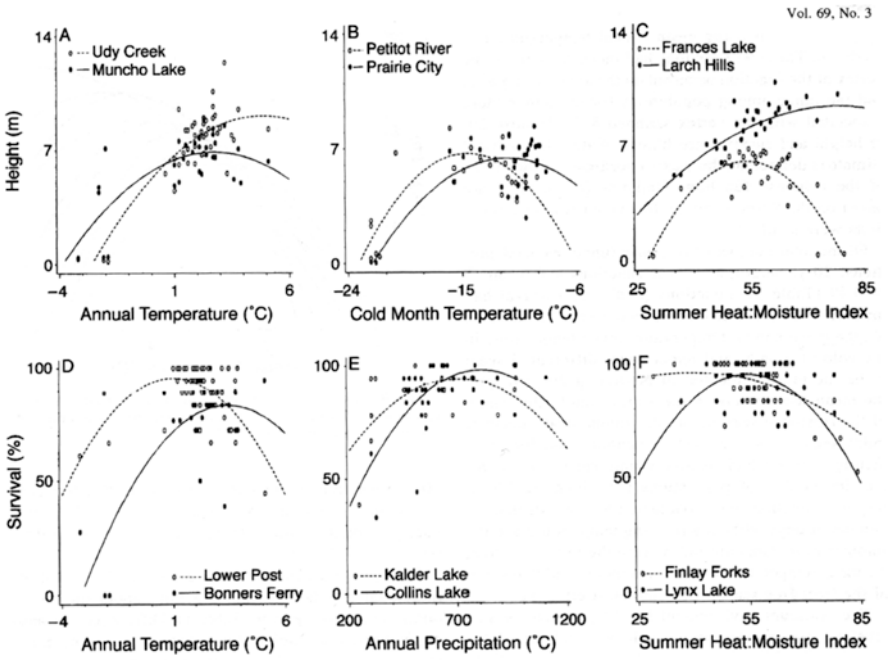


Fig. 8.10 Response functions for *Pinus contorta* ssp. *latifolia* populations plotted in relation to the observed data illustrating differential climatic adaptation between pairs of populations. (**a**–**c**) relate 20-year height to climate; (**d**–**f**) relate survival to climate. (From Rehfeldt et al. 1999)

environments. Both selective forces are regulated, suggests the authors, by asymmetric gene flow from the center of the species distribution toward the peripheries of the species range.

A comparable study (Rehfeldt et al. 2002) conducted with 110 populations of *P. sylvestris* growing on 47 planting sites in both Eurasia and North America revealed remarkably similar results as shown by the *P. contorta* study. Large differences in growth potential and climatic optima were demonstrated among populations, with populations originating in warmer, milder locations showing superior growth potential and climatic optima reflective of their origins. Moreover, this study also found populations were typically growing in suboptimal environments. Both studies suggest that relatively small changes in climate will greatly affect growth and survival of forest tree populations and that maintaining current forest productivity levels will require large-scale redistribution of genotypes across the landscape (assisted migration).

In a subsequent study using the same data sets as Rehfeldt et al. (1999), Wang et al. (2010a) introduced the URF analytical method. The URF integrates both environmental and genetic effects of climate, on dependent variables, like growth, by describing phenotype as a function of the climate of the test site and the provenance origin, simultaneously. In the study reported, Wang et al. (2010a) modeled the effect of mean annual temperature (MAT) on height and plotted the model surface to

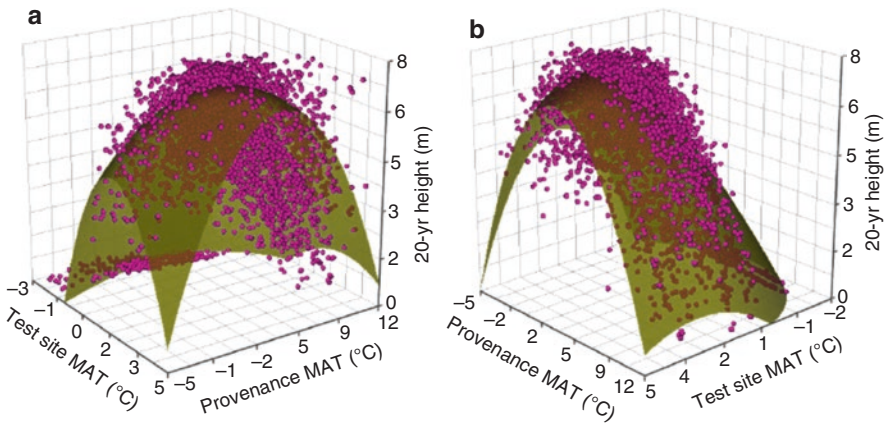


Fig. 8.11 Observed heights of populations in *Pinus contorta* provenance tests and modeled universal response function in which height is predicted as a joint function of site and provenance mean annual temperature (MAT). The fits of the data points to the model surface are shown in two rotations (**a** and **b**, where **b** is rotated 90° from **a**). The provenance test analyzed in this study consists of a range-wide collection of 140 populations planted at 60 test sites in British Columbia, Canada. (From Wang et al. 2010a)

observed performance from the long-term trials (Fig. 8.11) demonstrating the predictive power of a relatively simple model. An enhanced model, using additional independent variables, was then used to generate the predicted effect of a 1 °C MAT increase on both environmental and genetic components of height growth (Fig. 8.12). This exercise revealed that (1) both environmental and genetic effects are linearly associated with MAT, (2) environmental effects are substantially stronger than genetic effects, and (3) there is a predicted positive effect on mean tree height with increasing temperatures for provenances that originate from sites with MAT lower than 3 °C and 4.5 °C for genetic and environmental effects, respectively. For populations from MAT above those values, the effect of warming was negative. The authors further illustrated that selective planting of sources to their predicted optimal climatic conditions, as anticipated by climate modeling, will result in substantially improved productivity. Yang et al. (2015) found very similar results for another subboreal species, *Picea mariana*, using the URF approach both with respect to the relative strength of genetic and environmental effects and the tipping point for provenance MAT.

The development of the URF, as described here, likely heralds the adoption of this analytical approach for future, well-designed common garden studies. Indeed, the study authors (Wang et al. 2010a) note many positive attributes of the URF, including (1) improved prediction of climate change impacts on phenotypes, (2) reduced size and cost of future provenance trials, (3) quantifying and comparing environmental and genetic effects of climate on population performance, and (4) prediction of performance of any population growing in any climate. Still, much is to be learned, and scientists and resource managers must use caution in their

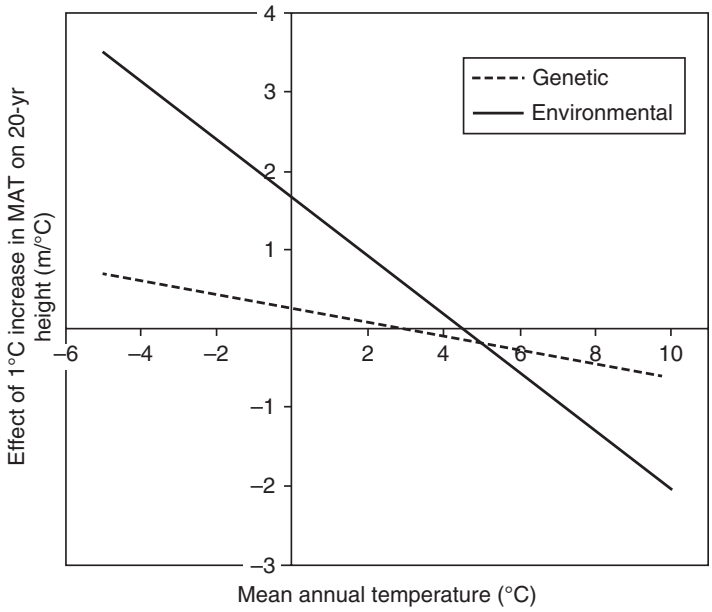


Fig. 8.12 Environmental vs. genetic effects of mean annual temperature (MAT) on *Pinus contorta* 20-year height. Effects are expressed by the rate of change in 20-year height with an increase of MAT by 1 °C at a given MAT (*x*-axis). (From Wang et al. 2010a)

interpretation of the results. In both studies noted here (Wang et al. 2010a; Yang et al. 2015), results suggest climate warming may be beneficial to forest productivity, particularly so when assisted migration is imposed. But these results are predicated on studies looking at temperature only which may not reflect the totality of climate change. Shifts in moisture regimes (seasonal distribution, amount, and form of precipitation) as well as insect and disease incidence will surely add complexity to anticipated future forest conditions. What we can say is that the technique offers another powerful tool for managing forests of the future.

Summary

Virtually everything we learned about genetic (genotypic) variation in forest trees before the development of allozymes, genetic markers, and genome sequences derives from the study of phenotypic variation among and within populations of trees. The predominant research approach for such studies has been the common garden trial which provides an objective means for dissecting observable phenotypic variation into genotypic and environmental effects. In forestry, the common garden test is frequently known as a provenance trial, when conducted with accessions from natural populations of a single species, or a genetic test, when conducted with pedigreed accessions from domesticated populations. Here we are concerned

only with the study of genetic variation in natural populations. Though common garden trials may serve many purposes, the most common applications are to select the best sources (i.e., most productive, best adapted) for planting in native or exotic locations and to identify the best deployment strategy for those sources across the biogeoclimatic landscape.

Provenance testing in conifers has been evolving for over 200 years. Today there are two main experimental approaches to provenance testing: short-term testing in artificial environments (genecological trials) and long-term testing in field trials. Short-term trials identify traits and environmental variables important for adaptation. Long-term trials provide direct evidence of survival, growth, and adaptation in multiple environments. Today it is widely recognized that patterns of genetic variation in wide-ranging species predominantly reflect adaptation to climatic gradients. Integrative measures of temperature and moisture such as mean annual temperature (MAT) or drought indices are typically used to model provenance performance. Due in large part to significant improvements in climatic data bases and climate modeling, new analytical approaches have been developed for interpreting provenance test results and guiding selection and deployment recommendations. The approaches produce mathematical models that reflect our understanding of climatic effects on environmental (population transfer functions) and genetic (population response functions) components of variation. The universal response function (URF) combines both effects in a single model. Increasingly such models are being used to improve predictions of climate change impacts on phenotypes.

Neutral Genetic Variation

9

Introduction and Background

Neutral genetic variation is described as that which is unaffected by natural selection. The neutral theory of molecular evolution, proposed in the late 1960s (Kimura 1968; King and Jukes 1969), holds that most genetic variation at the molecular level is evolutionarily neutral, the product of mutation, migration, drift, and mating systems rather than selection. For some time, disagreement over the extent or amount of molecular variation that is "neutral" was featured in the evolutionary literature, labeled as the neutralist/selectionist debate. Ohta (2002) suggested that slightly deleterious mutations can lead to nearly neutral variation, and today's literature generally uses the descriptor "neutral or nearly neutral" when describing most types of molecular variation. The neutral theory of molecular evolution was formalized in the late 1960s, in part due to recognition of developing technologies that provided scientists with the ability to observe variation at the level of individual genes. The neutral theory continues to be refined (Kimura 1983; Nei 2005, 2013; Nei et al. 2010).

The theory has its underpinnings in our understanding of genome organization and the central dogma of molecular biology. For most organisms, especially conifers, a very large majority of the genome (>90%) is non-coding (Chap. 4), and single nucleotide mutations or indels in these non-coding regions have little to no biological effect. In coding regions, the degenerate nature of the genetic code in which the same amino acid may be encoded by different nucleotide triplets (synonymous substitutions) insures many new mutations are of little biological consequence. The nature and relevance of non-synonymous substitutions in structural gene-coding regions and variation in regulatory elements is the subject of Chap. 10, Adaptive Variation, and outside the purview of this discussion.

This chapter reviews a substantial conifer literature based on selectively neutral molecular markers. That is not to say the descriptive results presented here are of little evolutionary consequence. Quite to the contrary, these studies have greatly informed our knowledge of other evolutionary forces and biological processes (i.e.,

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_9

drift, migration, mutation, mating systems, introgression, and hybridization) in conifers. The following text briefly reviews markers used to identify neutral genetic variation in conifers, statistical measures of genetic diversity and differentiation, and the objectives and purposes of studies using neutral genetic markers. General results are summarized with additional attention given to allozymes, the historical workhorse of diversity studies in conifers. Results based on biparentally inherited (nuclear genome) markers and uniparentally inherited (chloroplast and mitochondrial genomes) markers are compared. The contribution of the evolutionary forces of the mating system, migration (gene flow), and drift to measures of diversity and divergence are discussed. To conclude we provide summaries of case studies to illustrate the types of data obtained and how they are interpreted.

Molecular Markers Used in the Study of Neutral Variation

A wide range of molecular markers, developed over the last four decades, have found use in studies of neutral variation (White et al. 2007; Duran et al. 2009; Wheeler et al. 2011). Comprehensive tables listing marker types, characteristics, and applications have been published in classroom texts (e.g., Hillis et al. 1996, p. 517) and topical reviews (e.g., Bagnoli et al. 2011, Tables 4.1 and 4.2, p. 145, 148, respectively) and are only briefly discussed here. The most commonly used markers in population genetic studies of conifers prior to the development of the polymerase chain reaction (PCR) and high-throughput sequencing technologies were *allozymes*. Allozymes afforded an inexpensive and tractable means of characterizing the amount and distribution of genetic variation in large numbers of individuals (single trees) and populations but comparison of results from different studies may be confounded by the number and polymorphic nature of loci used. Though still occasionally used, their utility today is limited. Much of the historically summarized data on diversity and differentiation in conifers are based on allozyme studies.

Dominant (bi-allelic only) markers such as *RAPDs* (random amplified polymorphic DNAs) and *AFLPs* (amplified fragment length polymorphisms) enjoyed a fashionable but brief application in conifer population studies. Very large numbers of these markers could be generated relatively quickly and inexpensively but comparisons among species and organisms using these markers were difficult to interpret since marker loci are seldom orthologous. *RFLPs* (restriction fragment length polymorphisms) have found limited utility in population genetic studies primarily due to their relatively high costs and low throughput of genotyping efforts.

As recently as 2009, Duran et al. noted that *SSRs* (simple sequence repeats or microsatellites) and *SNPs* (single nucleotide polymorphisms) were the markers of choice in most modern genetic analyses. SSRs are powerful, co-dominant, and multi-allelic markers based on DNA fragment length, a function of variable number of tandem repeats of simple base-pair motifs. They are relatively expensive to develop and genotype. SNPs represent single nucleotide mutations in DNA sequences, are plentiful throughout the genome, and, today, are relatively easy to identify and genotype due to decreased sequencing costs. SNPs have been

frequently used in applied tree improvement and studies seeking to identify genetic associations with phenotypic and environmental traits. Increasingly, however, DNA sequencing of specific PCR products or genomic DNA is finding use in genetic analyses and will likely continue to grow in popularity as genetic sequencing's costs decline.

Three Conifer Genomes

As discussed in Chap. 2, conifers, like all plants, have three distinct genomes: nuclear (*nDNA*), chloroplast (*cpDNA*), and mitochondrial (*mtDNA*). The conifer nuclear genome is very large, composed mostly of non-coding, repetitive DNA elements, and is biparentally inherited. In contrast, the organellular genomes are tiny, are comprised mostly of gene-coding DNA, and are uniparentally inherited. The conifer chloroplast genome (~ 120–160 kb) is inherited paternally and is dispersed through wind-borne pollen (Neale et al. 1986; Neale and Sederoff 1989). The mitochondrial genome is generally inherited maternally, dispersed primarily through seed, though exceptions do occur. Maternal inheritance appears to be universal in the Pinaceae, but evidence of paternal inheritance occurs in several genera from the Cupressaceae, Araucariaceae, and Podocarpaceae (Neale et al. 1989, 1991; Whittle and Johnston 2002). The organellular genomes are highly conserved (low mutation rates) and haploid. They have effective population sizes roughly half that of the organism's diploid nuclear genome. Given the highly variable characteristics of these genomes, it is possible to find markers that will address virtually all scientific queries, except for those seeking genetic cause and effect, such as genotype/phenotype association.

Most of the types of markers noted above, including DNA sequencing approaches, have been used in studies of all three genomes noted here. Choice of markers for such studies is, however, heavily dependent on study objectives. For instance, Bagnoli et al. (2011) consider sequencing of cpDNA and mDNA to be a superior approach for identifying species, studying phylogeographic events, and estimating population differentiation within species. Nuclear DNA markers (RFLP, SNP) or sequences are preferred for estimates of within-population genetic diversity and gene flow. For many studies, use of two or all three genomes may be desirable to fully characterize the biological questions of concern.

Purpose and Applications of Neutral Genetic Variation Studies

The development of neutral genetic markers opened the investigative door to a wide range of basic and applied studies in conifers and other organisms. Most notably they made it possible to empirically address population genetic concepts previously relegated largely to theoretical treatments. It is also important to note that the introduction of molecular technologies and markers brought about profound and fundamental changes in the nature of genetic research of forest trees, including the focus

of new faculty hires, the types of students recruited to graduate studies, and the scope and scale of funding for genetics research (Wheeler et al. 2015).

The primary purpose of most molecular genetic studies using neutral genetic variation has been to characterize genetic diversity within and divergence among populations of interest, but the applications of the data (study objectives, analytical approaches) are remarkably varied. Basic research has been dominated by population genetic studies that seek to describe the relationship between the amount and distribution of genetic diversity and various life history traits and evolutionary forces. A common theme of many of these studies is to use diversity and divergence results to infer phylogeographic patterns, a topic covered well in a review of neutral genetic variation in conifers by Bagnoli et al. (2011). Phylogeography studies seek to identify the historical stochastic processes that shaped the observed patterns of neutral genetic variation on the landscape. The main objectives of phylogeographic studies are the identification of glacial refugia and the characterization of the processes that resulted in the patterns observed (Avise 2009; Bagnoli et al. 2011). The term "phylogeography," introduced by Avise et al. (1987), appears to have superseded the historically used term "biogeography," based largely on the former's integration of phylogenetic methods including gene genealogies. Diversity studies are used to inform the nature of mating systems, gene flow, drift (founder effects, bottlenecks), and hybridization/introgression events. They are used extensively in species conservation investigations (Chap. 13), and they have enlightened and advanced the field of conifer phylogenetics. Then again, some studies seek merely to contrast similar measures of diversity among taxa or to compare diversity measures obtained using different marker types.

Neutral genetic markers have been similarly embraced in applied research. Early studies sought to characterize the effects of forest management (silvicultural treatments) and tree improvement activities (domestication) on the levels and patterns of genetic diversity in select species. Markers provide a convenient means of genetically fingerprinting clones, seedlots, or species and insure genetic integrity in breeding programs. The ability to track parental contribution in controlled and uncontrolled crosses or crosses with multiple paternal sources has been remarkably useful in both applied and basic research. Applications of this type include quantifying pollen contamination in seed orchards, supplemental mass pollination success, and differential parental success.

Several statistical measures have been devised to measure the amount and distribution of genetic diversity *both within and among species and populations*, the most common of which are described here (Table 9.1). It is important to note the basis for specific diversity measures. Measures of diversity within species are not confounded by how diversity is partitioned within or among populations and are used to characterize species-level diversity. Within-population diversity is a measure of the observed or expected proportion of loci that are heterozygous, usually averaged over all populations studied.

In short, the study of neutral genetic variation has significantly and fundamentally contributed to our understanding of the levels and distribution of genetic diversity in conifer species and the evolutionary forces that drive those patterns. In the following section we draw from selected studies to illustrate how measures of

Table 9.1 Summary of statistical measures devised to measure the amount and distribution of genetic diversity

Metric and derivation	Definition
P	The proportion of polymorphic loci (loci with two or more alleles). Generally calculated on a population basis and averaged over all populations. Most studies consider a locus polymorphic when the most common allele occurs less than or equal to 95% of the time, though occasionally the bar is set at 99%
A **A_e** Calculated by $1/\Sigma p_i^2$, where p_i is the frequency of the i^{th} allele	Number of alleles per locus. Calculated by dividing the total number of alleles by the number of loci studied. It is a function of the proportion of polymorphic loci and the number of alleles at those loci. Occasionally studies report the number of alleles for polymorphic loci alone (A_p), or the *effective* number of alleles (A_e), which includes the frequency of each allele in the calculation. The latter is generally substantially less than **A** (30–50%)
H_e	Expected heterozygosity (also referred to as *genetic diversity*) is the expected proportion of heterozygous loci (two alleles) per individual. It is calculated based on Hardy-Weinberg proportions from allele frequencies. It is a function of allele frequencies, the number of alleles per polymorphic locus, and the proportion of polymorphic loci. It is the most frequently used composite measure of genetic diversity
H_o Calculated by $1-\Sigma p_i^2$	Observed heterozygosity or the proportion of loci observed to be heterozygous per individual based on actual *genotype* frequencies at each locus averaged across all loci. This measure requires gathering of genotype data (characterizing both alleles at a locus) as opposed to population gene (allele) frequencies, as is often done in conifer population studies
F_{ST}, G_{ST}, R_{ST}, Q_{ST} $F_{ST} = \sigma^2_p/\bar{p}_i(1-\bar{p}_i)$, where σ^2_p is the variance in allele frequencies and $\bar{p}_i$ is the mean frequency of the i^{th} allele	Measures of the proportion of total genetic diversity that occurs among populations. The measures noted here are used interchangeably though they are not precisely equivalent. **F_{ST}** is calculated on individual loci, and averaged over all loci studied, while **G_{ST}** is calculated directly on a multi-locus basis. The latter two measures are calculated based on haplotype frequencies rather than allele frequencies and are typically used with **SSR** and biallelic markers
D	Nei's genetic distance is defined as the accumulated number of detectable gene substitutions per locus between two populations (Nei 1972). It is calculated on a per-locus basis and averaged over all loci studied. Genetic distance provides another means to gauge the genetic differences between two populations. Populations fixed for different alleles at a given locus would have a genetic difference of 1.0 for that locus. Identical allele frequencies in two populations would result in a **D** = 0
F_{IS}	A measure of the level of inbreeding in a population, calculated on a single-locus basis and averaged over all loci studied
N_m	A measure of the average number of individuals migrating into a population each generation and a reflection of the level of gene flow among populations. It is generally high (> ~5.0) when diversity among populations is low and vice versa

diversity have informed specific research applications. These case studies are preceded by a general review of the literature, facilitated by a lengthy and information-dense table (Table 9.2). Regrettably, due to the extensive literature available on this topic, many worthy citations have surely been omitted and for this we apologize to those authors. We have concentrated on citing studies published over the last 20–25 years and on providing findings of previous reviews that have summarized results of literally hundreds of studies from the previous 20+ years (Hamrick et al. 1992; Ledig 1998; Petit et al. 2005).

General Diversity Results: Allozymes

The proliferation of electrophoretic studies of genetic diversity in plants induced a series of early reviews that sought to (1) characterize the amount and distribution of variation in allozymes and (2) correlate levels and patterns of diversity with ecological and life history characteristics (Hamrick et al. 1979, 1992; Nevo et al. 1984; Loveless and Hamrick 1984; Hamrick and Godt 1989). The last of these reviews compiled information from over 968 studies of which approximately 200 provided genetic parameter estimates for woody plants. Of these, 121 studies of conifers were summarized. The exceptional efforts required to compile and analyze these data have not been repeated in recent times, and they remain common citations when comparing diversity estimates in new studies. For many of the associations between life history traits and measures of the amount and distribution of diversity, results are confounded by the fact that data sets include both conifer and angiosperm tree species, but for some traits, the trends observed seem not to deviate appreciably from data on conifers alone. These results are generally summarized below (values noted are averages over variable sample sizes). It should be noted here that the authors being cited (Hamrick et al. 1992) have used the term "endemic" in a way we have interpreted to mean a species has a highly restricted natural range native to a specific geography.

Variation Within Species

- Long-lived woody species have a higher proportion of polymorphic loci (P), more alleles per locus (A), and more genetic diversity (H_e) than other life forms such as annuals or herbaceous perennials.
- Conifers have significantly higher P (71.1% vs. 59.5%) and A (2.38 vs. 2.10) but similar levels of H_e (0.169 vs. 0.183) compared to angiosperm trees.
- Species with endemic (and presumably restricted) geographic ranges have significantly lower H_e (0.087) than species from narrow (0.165), regional (0.169), or widespread distributions (0.257). Geographic range accounted for 44.5% of the variation in genetic diversity explained by ecological and life history traits or 15.1% of total variation in genetic diversity observed within species. Less important factors (breeding system, successional status, and seed dispersal mechanisms) were more relevant in separating conifers from angiosperms.

Table 9.2 Compiled estimates of genetic diversity within populations and divergence among populations of 88 conifer species (and subspecies)

Family/species	M[1]	Ob[2]	Pops	Loci	*P*	*A*	H_o	H_e	F_{IS}	$F_{ST}/G_{ST}/R_{ST}/Q_{ST}$	N_m	Citation
Conifers (w/i species; *N* = 89)	**A**		**8.9 (1.0)**	**17.3 (0.9)**	**71.1 (2.6)**	**2.38 (0.09)**		**0.169 (0.008)**				Hamrick et al. (1992)
Conifers (w/i populations; *N* = 102)	**A**		**8.9 (1.0)**	**17.3 (0.9)**	**53.4 (2.4)**	**1.83 (0.58)**		**0.151 (0.008)**		**0.073 (0.010)**		Hamrick et al. (1992)
Conifers (w/i populations; *N* = 27)	**A**		**19.1**	**17.5**				**0.175 (0.003)**	**0.100**	**0.100 (0.003)**		Petit et al. (2005)
Conifers (w/i populations; *N* = 34)	**C**							**0.51 (0.007)**		**0.165 (0.036)**		Petit et al. (2005)
Conifers (w/i populations; *N* = 21)	**M**							**0.667 (0.010)**		**0.764 (0.008)**		Petit et al. (2005)
Araucariaceae												
Agathis robusta	AF	D, Cm						0.470				Peakall et al. (2003)
Agathis robusta	A	D, Cm						0.190	0.080			Peakall et al. (2003)
Araucaria angustifolia	A, C, N	Cm, C										Stefenon et al. (2009)
Araucaria angustifolia	AF	P	9	673	46–89		0.270					Souza et al. (2009)
Araucaria angustifolia	R	TI, C	3		59.7		0.220				5.400	Medri et al. (2003)
Araucaria cunninghamii	AF	D						0.140				Peakall et al. (2003)

(continued)

Table 9.2 (continued)

Family/species	M[1]	Ob[2]	Pops	Loci	*P*	*A*	H_o	H_e	F_{IS}	$F_{ST}/G_{ST}/R_{ST}/Q_{ST}$	N_m	Citation
Araucaria cunninghamii	N	D						0.610	0.370			Peakall et al. (2003)
Araucaria cunninghamii	R	Cm	8	7	74.0		0.410		0.042			Pye et al. (2009)
Araucaria cunninghamii	I	Cm		6	81.0		0.390					Pye et al. (2009)
Wollemia nobilis	A	D	1	13	0.0	0.00	0.000			0.000		Peakall et al. (2003)
Wollemia nobilis	AF	D		800	0.0	0.00	0.000			0.000		Peakall et al. (2003)
Wollemia nobilis	N	D		20	0.0	0.00	0.000			0.000		Peakall et al. (2003)
Cupressaceae												
Austrocedrus chilensis	A	D	2	12	41.0	1.80	0.071					Ferreyra et al. (1996)
Austrocedrus chilensis	A	S	27	12		1.50	0.135	0.143		0.060		Pastorino and Gallo (2009)
Chamaecyparis obtusa	A	TI	14	11	56.7	1.80	0.187					Uchida et al. (1993)
Cupressus chengiana	N	C	6	6	100.0	4.50	0.510	0.640		0.07–0.11		Lu et al. (2014a)
Cupressus duclouxiana	N	C	10	6	100.0	4.90	0.540	0.620		0.07–0.12		Lu et al. (2014a)
Cupressus funebris	N	C	10	6	100.0	4.10	0.780	0.590		0.120		Lu et al. (2014a)
Cupressus gigantea	N	C	9	6	100.0	6.10	0.530	0.690		0.03 Amova		Lu et al. (2014a)

Thuja occidentalis	N	S	4	6		8.3 vs. 10.8[3]	0.54 vs. 0.64 [3]	0.56 vs. 0.65[3]	0.047 vs. 0.0[3]			Pandey and Rajora (2012)
Widdringtonia cedarbergensis	A	LH, M, C	7	19	27.7	1.30	0.300	0.390	0.170	0.176		Thomas and Bond (1997)
Widdringtonia nodiflora	A	LH, M, C	7	19	31.9	1.50	0.250	0.400	0.241	0.313		Thomas and Bond (1997)
Widdringtonia schwarzii	A	LH, M, C	3	19	23.5	1.20	0.390	0.400	0.025	0.050		Thomas and Bond (1997)
Pinaceae												
Abies spp. ($N = 7$)	**A**	**Many**	**5.4**	**13.6 (3.2)**				**0.130 (0.007)**		**0.063 (0.019)**		Hamrick et al. (1992)
Abies alba	C	Cm, S	17			90 haplotypes				0.133		Vendramin et al. (1999)
Abies balsamea	A	B, M, S	4	22	6.9 vs. 34.1[3]	1.07 vs. 1.34[3]	0.005 vs. 0.025[3]	0.01 vs. 0.025[3]	0.50 vs. 0.01[3]	0.037	6.5	Shea and Furnier (2002)
Abies bracteata	A	C, M	6	30	10.0	1.20	0.017 vs. 0.034[4]	0.028 vs. 0.036[4]	0.39 vs. 0.05[4]	0.02 vs. 0.080[4]	9.8 vs. 3.5[4]	Ledig et al. (2006a)
Abies cephalonica, etc.	A	Pg		23	65.9	1.80	0.184	0.239	0.228	0.048		Fady and Conkle (1993)
Abies flinckii	A	P	6		30.2	1.60	0.100	0.110	0.110	0.271		Aguirre-Planter et al. (2000)
Abies guatemalensis	A	P	10		20.0	1.40	0.050	0.070	0.160	0.120		Aguirre-Planter et al. (2000)

(continued)

Table 9.2 (continued)

Family/species	M[1]	Ob[2]	Pops	Loci	*P*	*A*	H_o	H_e	F_{IS}	$F_{ST}/G_{ST}/R_{ST}/Q_{ST}$	N_m	Citation
Abies hickeli	A	P	6		26.0	1.40	0.060	0.080	0.100	0.070		Aguirre-Planter et al. (2000)
Abies nebrodensis	A	C	1	11	36.4	1.80	0.107	0.162	0.340			Ducci et al. (1999)
Abies religiosa	A	P	5		31.8	1.50	0.080	0.110	0.150	0.250		Aguirre-Planter et al. (2000)
Abies siyuanensis	AF	Cm, C	7	389	35.7			0.136		0.390		Tang et al. (2008)
Abies siyuanensis	N			8	62.5	1.97		0.337		0.210		Tang et al. (2008)
Cathaya argyrophylla	M	Pg, C	15	2		3.0 [5]	0.000	0.000		1.000		Wang and Ge (2006)
Cathaya argyrophylla	N2	Pg, C	16	8		4.00				0.220		Wang and Ge (2006)
Cathaya argyrophylla	N	C	2	50	94.0	2.80	0.370	0.410				Wang et al. (2010b)
Cedrus atlantica	C	P	6	25 haplotypes			0.950			0.107		Terrab et al. (2006)
Larix decidua	A	M		9					0.651–0.675[6]			Lewandowski and Burczyk (2000)
Larix decidua	A	Pg	2	16	50.0	2.00	0.125	0.126				Semerikov and Lascoux (1999)
Larix gmelinii	A	Pg	6	16	68.8	2.50	0.154	0.155	0.024	0.020		Semerikov and Lascoux (1999)

Larix kaempferi	A	Pg	1	16	56.3	1.80	0.122	0.136				Semerikov and Lascoux (1999)
Larix laricina	A	Pg	3	16	40.0	1.60	0.075	0.113	0.003	0.026		Semerikov and Lascoux (1999)
Larix lyallii	N	B	19	7	84.9 [7]	3.20	0.389	0.418		0.066–0.15 [8]		Khasa et al. (2006)
Larix lyallii	A	Pg	1	16	37.5	1.40	0.082	0.094				Semerikov and Lascoux (1999)
Larix occidentalis	A	P	19	23	30.4	1.50	0.061	0.082		0.086		Fins and Seeb (1986)
Larix occidentalis	N	B	9	7	100.0	5.50	0.521	0.580		0.037–0.052 [8]		Khasa et al. (2006)
Larix occidentalis	A	Pg	4	16	43.8	1.60	0.124	0.150	0.178	0.100		Semerikov and Lascoux (1999)
Larix olgensis	A	Pg	3	16	37.5	1.40	0.080	0.082				Semerikov and Lascoux (1999)
Larix sibirica	A	Pg	15	16	56.3	1.80	0.117	0.130	0.059	0.080		Semerikov and Lascoux (1999)
Picea spp. ($N = 28$)	**A**	**Many**	**9.4**	**12.2 (1.6)**				**0.218 (0.020)**		**0.055 (0.011)**		Hamrick et al. (1992)
Picea abies	A	P	11	12	86.1			0.115		0.009		Radu et al. (2014)

(continued)

Table 9.2 (continued)

Family/species	M[1]	Ob[2]	Pops	Loci	*P*	*A*	H_o	H_e	F_{IS}	$F_{ST}/G_{ST}/R_{ST}/Q_{ST}$	N_m	Citation
Picea abies	A	P	9	21	45.5	1.83		0.165		0.042		Giannini et al. (1991)
Picea abies	S	B	8	7			0.580	0.410		0.118		Scotti et al. (2000)
Picea abies	N	P					0.465	0.640	0.251	0.029		Tollefsrud et al. (2009)
Picea chihuahuana	A	M, B	10	24	23.0	1.37	0.073	0.093	0.185	0.248		Ledig et al. (1997)
Picea chihuahuana	M	P, B, Cm	16	16			0.000			1.000		Jaramillo-Correa et al. (2006)
Picea chihuahuana	C	P, B, Cm	17	6			0.415			0.23–0.362 [9]		Jaramillo-Correa et al. (2006)
Picea engelmannii	A	P	16	24	80.0	2.40		0.255		0.147		Ledig et al. (2006b)
Picea glauca	A, Q	C, Cm	22	6	100.0	2.20	0.306	0.290	−0.051	0.038–0.540 [10]		Furnier et al. (1991)
Picea glauca	A, Q	C, Cm	26	14	83.7	2.38	0.348	0.341	0.012	0.014		Jaramillo-Correa et al. (2001)
Picea glauca	E	Cm	26	11	84.4	2.43	0.332	0.313	0.052	0.020–0.246 [10]		Jaramillo-Correa et al. (2001)
Picea glauca	N	B	22	6	100.0		0.737	0.770	0.192–0.241	0.057–0.036 [8]		Anderson et al. (2011)
Picea glauca	N	TI	12	4		11.30	0.514	0.647	0.172	0.032	7.5–12.25	Fageria and Rajora (2013)

Picea glauca x engelmanni	A	TI	9	17	63.0	2.16	0.189	0.210		0.040	7	Stoehr and El-Kassaby (1997)
Picea glehnii	A	P	10	12	75.0	1.98	0.080	0.088		0.022		Wang and Nagasaka (1997)
Picea mariana	A	P	21	23	38.0	1.44	0.120	0.107	Negative	0.059		Yeh et al. (1986)
Picea mariana	A	TI	16	32	67.0	2.52	0.222	0.308				Rajora and Pluhar (2003)
Picea omorika	A	C	13	16	20.9		0.073	0.067	0.125	>0.26		Ballian et al. (2006)
Picea omorika	C	P	6	5		4 haplo			−0.026	0.227 [7]		Nasri et al. (2008)
Picea pungens	A	P	4		42.7	1.60		0.138		0.086		Ledig et al. (2006b)
Picea rubens	A	P, M	10	37	29.1	1.60	0.097	0.100	~ 0	0.047		Rajora et al. (2000a)
Picea sitchensis	A	P	10	22	0.5		0.150			0.079		Yeh and El-Kassaby (1980)
Picea sitchensis	A	TI	11	13	67.0	1.82		0.183		0.082		Chaisurisri and El-Kassaby (1994)
Pinus spp. (n = 112)	**A**	**Many**		**20.5 (0.10)**	**53.5 (0.35)**			**0.187 (0.008)**		**0.083 (0.013)**		Ledig (1998)
Pinus spp. (n = 93)	**A**	**Many**	**7.8**	**19.9 (1.3)**				**0.136 (0.009)**		**0.065 (0.008)**		Hamrick et al. (1992)

(continued)

Table 9.2 (continued)

Family/species	M[1]	Ob[2]	Pops	Loci	*P*	*A*	H_o	H_e	F_{IS}	$F_{ST}/G_{ST}/R_{ST}/Q_{ST}$	N_m	Citation
Pinus spp. (n = 41)	**A**	**Many**						**0.198 (0.013)**		**0.129 (0.025)**		Delgado et al. (2002)
Pinus spp. (n = 9)	**C**	**Many**						**0.582 (0.096)**		**0.245 (0.080)**		Delgado et al. (2002)
Pinus albicaulis	A	P	30	20	23.0	2.2 [7]	0.088	0.096	0.083	0.034	4	Jorgensen and Hamrick (1997)
Pinus attenuata	A	PG	17	32	55.0	1.70		0.131		0.120		Millar et al. (1988)
Pinus attenuata	A	P, Pg, Cm	4	32–36	40.3	1.50		0.110				Wu et al. (1999)
Pinus attenuata	R	P, Pg, Cm	4	>90	48.2	1.50		0.150		0.360		Wu et al. (1999)
Pinus attenuata	M1	P	4	13			0.210			0.790		Wu et al. (1998)
Pinus attenuata	C1	P, Cm	4	25						~ 0		Hong et al. (1993)
Pinus balfouriana	A	P	16	11	27.3	1.27	0.075		0.203–0.443	0.075–0.242		Oline et al. (2000)
Pinus banksiana	A	S	82	14	78 [7]	3.20	0.200	0.215	0.048	0.022	11	Saenz-Romero et al. (2001)
Pinus brutia	A	B	4	7	57.1	1.78	0.205	0.216		0.021		Panetsos et al. (1998)
Pinus canariensis	N2, Q	Cm, P, B,	24	8						0.077–0.225 [12]		Heredia et al. (2014)
Pinus caribaea var. bahamensis	C	P, B, C	10	6	68.0	10.00	0.555			0.023		Sanchez et al. (2014)

Pinus caribaea var. bahamensis	N	P, B, C	10	5		22.00	0.074			0.020		Sanchez et al. (2014)
Pinus caribaea var. bahamensis	C	B, Pg	2	9 [5]				0.865 [11]				Jardón-Barbolla et al. (2011)
Pinus caribaea var. caribaea	C	B, Pg	4	32 [5]				0.974 [11]				Jardón-Barbolla et al. (2011)
Pinus caribaea var. hondurensis	N	C	2	6		2.8–4.0	0.422	0.465 [11]	0.053	0.030	13.2	Delgado et al. (2011)
Pinus caribaea var. hondurensis	C	B, Pg	12	52 [5]				0.951 [11]				Jardón-Barbolla et al. (2011)
Pinus cembra	N	C	25	10	100.0	5.80	0.480	0.537	0.063			Dzialuk et al. (2014)
Pinus cembra	N	B, Cm	11	6			0.48–0.61		0.093	0.055		Lendvay et al. (2014)
Pinus cembra	C	B, Cm	11	4			0.73–0.95			0.095		Lendvay et al. (2014)
Pinus contorta	A	PG	17	25	28 /38[13]	1.65 /1.8 [13]	0.096 /0.119 [13]			0.035–0.065		Aitken and Libby (1994)
Pinus contorta	A, Q	P, Cm	34	42	58–73	1.62–1.94	0.109–0.126			0.060		Wheeler and Guries (1982a) or (1982b)?
Pinus contorta	M1									0.356		Godbout et al. (2008)

(continued)

Table 9.2 (continued)

Family/species	M[1]	Ob[2]	Pops	Loci	*P*	*A*	H_o	H_e	F_{IS}	$F_{ST}/G_{ST}/R_{ST}/Q_{ST}$	N_m	Citation
Pinus cubensis	C	B, Pg	2	95				0.752 [11]				Jardón-Barbolla et al. (2011)
Pinus flexilis	A	M	5	28	50.0	1.70	0.159	0.162		0.031	6.5–7.8	Schuster and Mitton (2000)
Pinus flexilis	M	B	40	1						0.800	0.12	Mitton et al. (2000)
Pinus halepensis	A	B	20		32.6	1.53		0.110				Korol et al. (2002)
Pinus halepensis	A	S	15	5			0.252	0.304	0.172			Agúndez et al. (1999)
Pinus halepensis	N	C, TI	3	9			0.289		0.138	0.319		Steinitz et al. (2012)
Pinus kesiya	A	B	9	12	45.6	1.60	0.153			0.121		Myburg and Harris (1997)
Pinus kesiya	A	Cm, Pg	12	16	56.3		0.130	0.148		0.023		Szmidt et al. (1996a)
Pinus koraiensis	A	Cm, B	12	18	46.8	1.95	0.158	0.169	0.045	0.069	3.4	Kim et al. (2005)
Pinus koraiensis	R	Cm, B	12	38						~0.10		Kim et al. (2005)
Pinus kwangtungensis	C1	C	17	1						0.540		Tian et al. (2008)
Pinus leucodermis	A	Cm, B	7	23	39.7	1.50				0.054		Boscherini et al. (1994)
Pinus leucodermis	C	Cm, B	8	1						0.000		Boscherini et al. (1994)

Pinus leucodermis	A	C, Cm	7	23						0.016–0.04		Bucci et al. (1997)
Pinus leucodermis	R	C, Cm	7	16				0.231		0.04–0.098		Bucci et al. (1997)
Pinus longaeva	A	Cm	3	36	38.9	1.92	0.122	0.134	0.078	0.011		Lee et al. (2002)
Pinus longaeva	R	Cm	4		34.1	1.40		0.130		0.039		Lee et al. (2002)
Pinus maestrensis	C	B, PG	2	7[J]				0.889 [11]				Jardón-Barbolla et al. (2011)
Pinus maximartinezii	A	M, B	1	33	30.3			0.122	0.081			Ledig et al. (1999)
Pinus merkusii	A	Cm, PG	5	16	56.3		0.043	0.073		0.337		Szmidt et al. (1996a)
Pinus muricata	A	PG	17	32	47.0	1.60		0.118		0.220		Millar et al. (1988)
Pinus muricata	A	P, Pg, Cm	6	32–36	40.6	1.50		0.110				Wu et al. (1999)
Pinus muricata	R	P, Pg, Cm	6	>90	40.8	1.39		0.130		0.450		Wu et al. (1999)
Pinus muricata	M1	P	6	13			0.230 [11]			0.750		Wu et al. (1998)
Pinus muricata	C1	P, Cm	10	25			0.009 [11]			>0.80		Hong et al. (1993)
Pinus occidentalis	C	B, PG	2	11[J]				0.874 [11]				Jardón-Barbolla et al. (2011)

(continued)

Table 9.2 (continued)

Family/species	M[1]	Ob[2]	Pops	Loci	*P*	*A*	H_o	H_e	F_{IS}	$F_{ST}/G_{ST}/R_{ST}/Q_{ST}$	N_m	Citation
Pinus oocarpa	A	C	5	12	25.0	1.90	0.115	0.102	−0.114	0.001	227	Sáenz-Romero and Tapia-Olivares (2003)
Pinus palustris	A	B	23	22	39.5	1.9 [7]	0.103	0.105	−0.002	0.041		Schmidtling and Hipkins (1998)
Pinus parviflora	A	Cm, P, B	16	11	61.3	2.10	0.231	0.259	0.113	0.044	5.4	Tani et al. (2003)
Pinus parviflora	M	Cm, P, B	16	1						0.870	0.075	Tani et al. (2003)
Pinus pinaster	A	P, C	11	19	28.8	1.56	0.310	0.346	0.104	0.104		Wahid et al. (2004)
Pinus pinaster	A, Q	Cm								0.019–> 0.73 [10]		González-Martínez et al. (2002)
Pinus pinaster	N	P, B				2.00	0.403			0.170–0.201 [8]		Naydenov et al. (2014)
Pinus pinaster	N	Cm	23	3		23.00	0.630	0.752	0.144	0.111		Mariette et al. (2001)
Pinus pinaster	AF	Cm	23	122				0.167		0.102		Mariette et al. (2001)
Pinus pinceana	A	M, C, P	8	26	56.3	1.80	0.144	0.174	0.116	0.152	4.12 northern, 1–1.6 total	Ledig et al. (2001)
Pinus ponderosa	N	P	4				0.447	0.677	0.27–0.44	0.01–0.1		Lesser et al. (2013)

Pinus radiata	A	PG	17	32	58.0	1.90		0.141		0.130		Millar et al. (1988)
Pinus radiata	A	P, Pg, Cm	3	32–36	47.9	1.76		0.170				Wu et al. (1999)
Pinus radiata	R	P, Pg, Cm	3	>90	49.8	1.50		0.170		0.260		Wu et al. (1999)
Pinus radiata	M1	P	3	13			0.210 [11]			0.790		Wu et al. (1998)
Pinus radiata	C1	P, Cm	5	25			0.028 [11]			~ 0		Hong et al. (1993)
Pinus resinosa	C	B	10	10				0.138		0.560		Walter and Epperson (2001)
Pinus resinosa	C	B, S	29	10	40.0	1.60	0.152 [11]					Walter and Epperson (2005)
Pinus resinosa	N	B, M	17	5	100.0	9.00	0.185	0.508	0.504	0.280	<1.0	Boys et al. (2005)
Pinus strobus	A	C	9	20	47.8	1.75	0.215	0.195		0.061		Rajora et al. (1998)
Pinus strobus	A	B, S	1	11 poly		2.80	0.371	0.370	−0.013			Jones et al. (2006)
Pinus strobus	A	S, P	6	21	23.8	1.36 a	0.074–0.068 [3]	0.08–0.078 [3]	0.12–0.07 [3]	0.033		Myers et al. (2007)
Pinus strobus	C	S, P	6					0.78–0.84 [3]				Myers et al. (2007)
Pinus strobus	C	S	2			7.70			0.01–0.05	0.005		Marquardt and Epperson (2004)

(continued)

Table 9.2 (continued)

Family/species	M[1]	Ob[2]	Pops	Loci	*P*	*A*	H_o	H_e	F_{IS}	$F_{ST}/G_{ST}/R_{ST}/Q_{ST}$	N_m	Citation
Pinus sylvestris	A	Cm	2	20	70.0		0.283	0.269	−0.065	0.022		Szmidt et al. (1996b)
Pinus sylvestris	A	D	7	11		2.80	0.310	0.325	−0.001	0.040		Prus-Glowacki and Stephan (1994)
Pinus sylvestris	A	C	6	17	62.7	2.58		0.294		0.036	6.32	Bilgen and Kaya (2007)
Pinus sylvestris	A	P	13	8		2.60	0.357	0.356	−0.001	0.035–0.070		Prus-Glowacki and Bernard (1994)
Pinus sylvestris	A	TI	1	18			0.253	0.262	0.038			Wang et al. (1991)
Pinus sylvestris	C	P, Cm	12	6	10.0	3.60	0.410			4.9–11.0		Naydenov et al. (2005)
Pinus sylvestris	M1	B	20							0.370		Sinclair et al. (1998)
Pinus sylvestris	R	Cm	2	22	45–88				−0.262	0.023–0.062		Szmidt et al. (1996b)
Pinus sylvestris	A	TI						0.280				Savolainen and Kärkkäinen (1992)
Pinus taeda	A	TI	3	18	76.0	2.19		0.171				Schmidtling et al. (1999)
Pinus taeda	C	TI										Lambeth et al. (2001)

Pinus thunbergii	A	P	13	27	55.3	2.20	0.214	0.212		0.044		Kim et al. (1997)
Pinus torreyana	A	B	2	59	3.4	1.00	0.000					Ledig and Conkle (1983)
Pseudotsuga spp. ($N = 11$)	**A**	**Many**	**15.5**	**15.5 (3.0)**				**0.163 (0.012)**		**0.074 (0.029)**		Hamrick et al. (1992)
Pseudotsuga menziesii	M2	Cm, Pg	29			3.00	0.760					Aagaard et al. (1995, 1998a)
Pseudotsuga menziesii	R	Cm, Pg	29	20		1.38	0.490			0.080		Aagaard et al. (1995, 1998a)
Pseudotsuga menziesii	A	Cm, Pg	29	36		2.75	0.210			0.050		Aagaard et al. (1995, 1998a)
Pseudotsuga menziesii	A	P	49	20	85.0	2.14		0.163	0.008	0.045	5	El-Kassaby and Ritland (1996a)
Pseudotsuga menziesii	A	P, C	11	18	28.3 P, 83.3 S [14]	1.5 P, 2.9 S [14]	0.079	0.077		0.298	0.59	Cruz-Nicolás et al. (2011)
Pseudotsuga menziesii	A	S, Cm, Q	4	27	0.5	1.79		0.171		0.068 among elevations		El-Kassaby and Sziklai (1982)
Pseudotsuga menziesii	A	TI		17								Adams et al. (1998)
Pseudotsuga menziesii	A	TI	>300	20	53.0	2.14		0.171		0.042		El-Kassaby and Ritland (1996b)
Pseudotsuga menziesii	A	TI	28	22		2.8–3.5		0.234–0 0.282				Wheeler et al. (1995b)

(continued)

Table 9.2 (continued)

Family/species	M[1]	Ob[2]	Pops	Loci	*P*	*A*	H_o	H_e	F_{IS}	$F_{ST}/G_{ST}/R_{ST}/Q_{ST}$	N_m	Citation
Pseudotsuga menziesii	N	TI		7–9	90.0	39.00				0.002		Slavov et al. (2005)
Tsuga canadensis	N	C, M, B	60	13	97.2	4.90	0.526	0.566	0.073	0.048		Potter et al. (2012)
Podocarpaceae												
Lagarostrobos franklinii	A	C	32	6	46.0	1.60	0.044	0.169	0.552	0.095		Shapcott (1997)
Podocarpus sellowii	N	B, C, D	3	3	100.0	2.33	0.393		−0.323	$G_{st} = 0.266$, $R_{st} = 0.298$, $F_{st} = 0.357$		Dantas et al. (2015)
Taxaceae												
Taxus baccata	A	C	7	9	76.0	2.70	0.238	0.274	0.130	0.062	3.8	Klumpp and Dhar (2011)
Taxus baccata	A	C, M, P	1	18	61.1	1.40		0.279	0.049			Lewandowski et al. (1995)
Taxus baccata	N	C, M, P	6	5	100.0	14.00	0.526	0.804	0.067	0.297		Chybicki et al. (2012)
Taxus baccata	R	C	14	?						0.10–0.15		Hilfiker et al. (2004)
Taxus brevifolia	A, Q	P	15	22	33.5	1.50	0.122	0.124		0.104		Wheeler et al. (1995a)
Taxus brevifolia	A	S	9	21	42.3	1.70		0.166	0.472	0.080	3	El-Kassaby and Yanchuk (1994)
Taxus cuspidata	A	C	5	14	45.7	1.70	0.172	0.168		0.067	3.5	Lee et al. (2000)

Taxus cuspidata	A	P	6	23	44.9	1.78	0.154	0.192	0.299	0.056	4.2	Chung et al. (1999)
Taxus wallichiana	N	C	3	10		2.6–4.4	0.227–0.527	0.299–0.578				Gajurel et al. (2013)

Data are drawn from 128 citations, the majority of which were published after the mid-1990s. Summary results of previous reviews are included in bold. The data capture estimates of diversity as measured using several types of genetic markers (M). Study objectives (Ob) have been ascribed to each citation. While characterizing genetic diversity is the basis for most papers, the primary purpose of the papers may be more inclusive. A key to marker and objective labels is provided in the footnotes, along with intra-table points of clarification

Abbreviations: *Pops* number of populations sampled (for summary entries, parenthetical values are standard errors), *Loci* number of loci used in estimates, *P* proportion of polymorphic loci, *A* mean number of alleles per locus, H_o observed heterozygosity, H_e expected heterozygosity, F_{IS} inbreeding, $F_{ST}/G_{ST}/R_{ST}/Q_{ST}$ measures of differentiation or estimates of the proportion of total diversity that occurs among populations, and N_m number of individuals migrating among populations per generation

[1] *A* allozyme, *AF* AFLP, *C* chloroplast SSR, *C1* chloroplast RFLP, *E* EST untranslated region, *M* mitochondrial SSR, *M1* mitochondrial minisatellite *M2* mitochondrial RAPD, *N* nuclear SSR, *N1* nuclear RFLP, *N2* nuclear SNP, *I* internal SSR, *R* RAPD, *S* SCAR, *Q* quantitative trait

[2] B biogeography/demographics, *C* conservation, *Cm* compare markers, *D* diversity, *IH* inheritance, *LH* life history, *M* mating systems, *P* population structure *Pg* phylogenetics, *S* spatial ptterns of variation, *TI* tree improvement, *Q* quantitative trait comparison

[3] denotes values comparing a measure of diversity for perpheral/marginal populations with central populations across a species' range

[4] denotes values measured in embryos vs mature trees

[5] denotes the number of observed haplotypes

[6] denotes value based on single locus estimate vs multilocus estimates

[7] denotes values based on polymorphic loci only

[8] denotes estimates of the proportion of total diversity attributed to among populations based on Rst followed by Fst methods

[9] denotes estimates of the proportion of total diversity attributed to among populations based on Rst followed by Gst methods

[10] denotes estimates of the proportion of total diversity attributed to among populations based on Gst followed by ANOVA of tree height in provenance trial

[11] denotes estimates based on haplotype diversity

[12] denotes estimates of the proportion of total diversity attributed to among populations based on Qst followed by ANOVA for height

[13] denotes diversity values for subspecies Bolanderi followed by values for subspecies contorta

[14] denotes values calculated on a population basis (P) and species basis (S) separately

Variation Within Populations

- Conifers retain more genetic variation within populations than angiosperms (*P*, 53.4% vs. 45.1%; *A*, 1.83 vs. 1.68; and H_e, 0.151 vs. 0.141).
- All parameters of genetic variation are significantly influenced by the geographic range of a species. Average values for endemic, narrow, regional, and widespread species are, respectively, *P*, 26.3%, 44.3%, 69.2%, and 74.3%; *A*, 1.48, 1.61, 2.31, and 2.56; and H_e, 0.056, 0.143, 0.194, and 0.228).
- Similarly, all genetic parameters varied significantly as a function of regional distribution with measures of diversity being highest for boreal/temperate species and declining, in order, for temperate species, temperate/tropical species, and tropical species. Genetic diversity in populations of species from lower latitudes ranged from 61% to 71% of the genetic diversity in populations of the boreal/temperate trees.

Distribution of Variation Among Populations (Based on Polymorphic Loci only)

- Long-lived woody species exhibit a much lower proportion of total diversity distributed among populations (G_{ST} = 0.084% or 8.4%) than do short-lived woody species (0.155), long-lived herbaceous perennials (0.278), or annuals (0.355).
- Variation among populations (G_{ST}) is significantly less for conifers (0.073) than for angiosperms (0.102).
- The distribution of variation among populations is significantly affected by the geographic range and regional distribution of the species. G_{ST} for endemic, narrow, regional, and widespread ranges averaged 0.141, 0.124, 0.065, and 0.033, respectively. For boreal/temperate, temperate, temperate/tropical, and tropical distributed species, G_{ST} was 0.038, 0.092, 0.109, and 0.119, respectively.

Differences in Measures of Diversity Among Conifer Genera and Families

Hamrick et al. (1992) compiled summary statistics for four conifer genera (*Abies*, *Picea*, *Pinus*, and *Pseudotsuga*) and found that they all exhibited relatively limited amounts of variation distributed among populations (F_{ST} range = 0.055 to 0.074) but differed significantly in genetic diversity: H_e = 0.13, 0.218, 0.136, and 0.163, respectively. Ledig (1998) summarized data for 112 species of *Pinus* and reported substantially higher H_e (0.187) and slightly more variation distributed among populations (0.083 vs. 0.065) than did Hamrick et al. (1992). A subsequent review of allozyme studies in 41 pine species, based on taxonomic groupings, gave higher estimates of diversity (H_e = 0.198) and differentiation (F_{ST} = 0.129) than either of previous

reviews (Delgado et al. 2002). The relatively modest number of citations for allozyme variation in remaining families and genera make meaningful comparisons suspect (see Table 9.2).

Allozyme Summary

The summarized results from the myriad allozyme studies noted in previous reviews, and those listed in Table 9.2, show that conifers typically retain relatively high levels of genetic diversity within populations but modest differentiation among populations, though significant variation in these measures occurs among species. In rare instances, a complete lack of heterozygosity at all or most loci studied has been observed (*Pinus torreyana*, Ledig and Conkle 1983; *Wollemia nobilis*, Peakall et al. 2003; *Pinus resinosa*, many studies). Similarly, genetic diversity estimates exceeding 0.30 are relatively uncommon (Prus-Glowacki and Bernard 1994; Thomas and Bond 1997; Agúndez et al. 1999; Rajora and Pluhar 2003; Jones et al. 2006). Except for *Pinus resinosa* and other species demonstrated to have gone through severe bottleneck events, very low levels of diversity are associated almost exclusively with rare and endemic species with very restricted ranges. The generally high levels of genetic diversity observed for conifers and other woody plants are attributed to the combination of ecological and life history characteristics that ensure the retention of diversity, such as long life spans, large individual size, predominantly outcrossing mating systems, large and relatively regular pollen and seed crops routinely distributed over great distances (primarily wind dispersed), and large, continuous populations.

Variation in the proportion of variation distributed among populations, while significant numerically and statistically, is generally modest, with most citations listing values below 10% for conifers. Exceptional reports exceeding 20% occur in cases where populations are highly fragmented and occur at great distances from other populations (Szmidt et al. 1996a; Ledig et al. 1997; Thomas and Bond 1997; Aguirre-Planter et al. 2000; Oline et al. 2000; Cruz-Nicolás et al. 2011). Such cases define the incipient stages of speciation events that will only be altered by renewed gene flow among populations.

General Diversity Results: Molecular Markers

Though we report results for 15 different marker types in Table 9.2, very few are represented by more than two to three studies and do not warrant further discussion. The remainder are discussed below, treated separately by genome of origin. By far the most common molecular markers represented in recent works were SSRs or microsatellites. These multi-allelic and highly polymorphic markers typically yield diversity estimates substantially higher than those derived from nuclear allozyme markers.

Organelle Markers

Petit et al. (2005; supplemental data) report average genetic diversity (H_e) for chloroplast and mitochondrial SSR markers in conifers to be 0.51 (N = 34) and 0.667 (N = 21), respectively. The average genetic diversity (both observed (H_o) and/or expected (H_e)) derived from 13 studies reported here (Table 9.2) for *cpSSR* is 0.725. Given the relatively limited number of comparisons between marker types for the same populations and species, it is difficult to derive even general conclusions about the relationships between them. What is clear is that considerable diversity is retained in neutral SSR markers, obtained virtually exclusively from repetitive DNA sequences outside of gene-coding regions. Such markers have great utility for studies of phylogeography.

Nuclear Markers

A summary of nuclear SSR results for 19 studies (Table 9.2) gives a mean H_{ex} intermediate to the organelle marker estimates (0.593) with a range of 0.337–0.804 and only a loose relationship between allozyme and nuclear SSR measures of diversity. Likewise, there seems to be little correlation between estimates for nuclear allozyme diversity and organelle SSR-based diversity measurements. Diversity measures based on modest sample sizes for AFLPs (0.197, N = 6) and RAPDs (0.241, N = 5) are slightly higher than conifer allozyme measures, but not appreciably so.

Population Differentiation

Perhaps more relevant than measures of diversity are estimates of how diversity is distributed among populations given the diverse modes of inheritance for the different markers. Mode of inheritance has a major effect on G_{ST} and provides insight into patterns of gene flow via pollen and seed (Petit et al. 2005). Mean estimates of G_{ST} for chloroplast, mitochondrial, and nuclear SSRs reported by Petit et al. (2005) were 0.165, 0.764, and 0.119, respectively. Bagnoli et al. (2011) report G_{ST} values where estimates have been made for the same set of species (N = 16 or 17). Their estimates are remarkably like those of Petit et al. (2005): 0.205, 0.763, and 0.127 for chloroplast, mitochondrial, and nuclear SSRs, respectively. The median value of 0.07 for the estimate of nSSR G_{ST} reported in Table 9.2 is slightly lower than the previous estimates and essentially equivalent to levels of differentiation observed for allozymes. The proportion of total diversity attributed to among populations for nuclear AFLPs and RAPDs summarized from Table 9.2 was like that observed for cpDNA (0.246 and 0.161, respectively).

Based on the data reported above, maternally inherited mitochondrial genomes exhibit considerably more subdivision than either paternally inherited (chloroplast) or biparentally inherited (nuclear) genomes, evidence that gene flow via seed is substantially reduced compared to that of pollen dispersal.

Factors Affecting Amount and Distribution of Genetic Variation

Mating Systems

Mating systems have a direct influence on both the amount and distribution of genetic diversity. Outcrossing tends to maintain diversity and reduce population differentiation. Conversely, inbreeding reduces heterozygosity and the rate of recombination and promotes genetic organization within populations leading to differentiation (Mitton 1992). Conifers exhibit a mixed-mating system characterized by both outcrossing and inbreeding, the level of either influenced by an array of factors such as age, stand density, population fragmentation, patterns of floral phenology, within and among tree pollen and seed abundance, and genetic load. Inbreeding, which includes both selfing and consanguineous mating, is commonly measured using the F_{IS} statistic. Other frequently used measures are t, the outcrossing rate, and s, the selfing rate, defined as $(1-t)$, where t can be calculated for individual loci and averaged or using multi-locus methods. The latter better reflects effects of all manner of inbreeding and is therefore more likely to produce estimates consistent with F_{IS} (e.g., $(1-F_{IS})$).

Neutral marker studies have shown that most conifers are predominantly outcrossing most of the time, but that their mating systems are dynamic and often variable among and within species, populations, and even individual trees (reviewed in Mitton 1992; Ledig 1998; Williams 2009). Average rates of inbreeding (calculated as F_{IS} or as $(1-t)$) compiled for multiple conifer studies, all based on allozyme results, consistently estimate levels to be near 10% or less (Schemske and Lande 1985; Ledig 1998; Petit et al. 2005; Table 9.2) but these averages disguise the considerable variation observed. In the current compilation (Table 9.2) F_{IS} ranged from zero (negative values considered to be zero for computation), observed for many species, to 0.552 for *Lagarostrobos franklinii* (Shapcott 1997), a Podocarp, and 0.675 for *Larix decidua*, a larch in the Pinaceae (Lewandowski and Burczyk 2000). While selfing rates may be quite high at time of pollination, allozyme studies have demonstrated that conifers in general exhibit strong selection against inbred embryos and seedlings (Plessas and Strauss 1986; Ledig 1998; Ledig et al. 2006a). Inbred progeny and surviving trees exhibit lifelong inbreeding depression (Sorensen and Miles 1974; Fowler and Park 1983; Wilcox 1983). Conifers have been shown to harbor high genetic loads, defined here as the number of lethal alleles or lethal-allele equivalents per individual, relative to other plants and animals (Sorensen 1969, 1971; Koski 1971; Wheeler 1989). A lethal allele, when homozygous, results in the death of a seed or plant. Most deleterious alleles are expressed at the time of embryo development and seed set. Consequently, relatively few selfed, filled seeds mature. In a controlled mating study of *Pseudotsuga menziesii* using 40 maternal trees, the average proportion of selfed, filled seeds was 6%, with a range of 0.1–17.9% (Wheeler 1989). This equated to a range in the number of lethal equivalents of 7–22, like that reported for another study conducted in *Pseudotsuga menziesii* by Sorensen (1969).

Inbreeding levels are typically higher in marginal populations than in populations near the center of a species range (Ledig et al. 2006a; Myers et al. 2007; Pandey and Rajora 2012). Similarly, though not universally, elevated inbreeding levels co-occur in populations that are highly differentiated, fragmented, or both (Thomas and Bond 1997; Aguirre-Planter et al. 2000; Ballian et al. 2006). Inbreeding rates in conifers typically vary inversely with levels of genetic diversity but notable exceptions occur here as well such as with *Widdringtonia cedarbergensis* and *Widdringtonia nodiflora* (Thomas and Bond 1997).

Gene Flow

Migration or gene flow is a powerful evolutionary force and critical determinant of population structure. In conifers, high levels of pollen gene flow are generally credited with producing low levels of population differentiation and relatively uniform or homogenized allele frequencies observed among and within populations. In contrast, maternally inherited mitochondrial markers have revealed that gene flow via seed is substantially less than that through pollen, leading to high levels of population differentiation for that genome. Gene flow has been studied in conifers for decades. Early efforts focused on physical measures of pollen and seed dispersal (reviewed in Lanner 1966; Koski 1971; Burczyk et al. 2004a; White et al. 2007) and later, with the introduction of allozymes, on studies seeking to directly estimate migration rates by determining paternity of seeds in controlled and surrounding populations (Wheeler and Jech 1992; Adams and Burczyk 2000). Subsequent population genetic studies, including many of those reported in Table 9.2, relied on an indirect measure of migration, N_m, to estimate gene flow. N_m was defined by Wright (1951) as the effective number of individuals migrating into a population every generation and is calculated using the statistic F_{st}, a standardized measure of genetic variance among populations: $F_{ST} = 1/(4\ N_m + 1)$. This estimator of gene flow has subsequently been shown to be flawed, as natural populations under study seldom meet the conditional biological assumptions of the infinite island model used by Wright (Whitlock and McCauley 1999). Regardless of the method's shortcomings, the story told by estimates of N_m based on Wright's approach likely approximates actual phenomena relatively well.

In brief, migration rates in conifers, based on nuclear and chloroplast markers, vary widely, from near zero (Tani et al. 2003; Boys et al. 2005; Cruz-Nicolás et al. 2011) to rates exceeding 10 (Saenz-Romero et al. 2001; Delgado et al. 2011; Fageria and Rajora 2013), but in general reveal gene flow via pollen is quite high. Measures of gene flow may vary, even within species, depending on the genome under study. For instance, allozyme studies in *Pinus flexilis* estimate N_m to be from 6.5 to 7.8 (Schuster and Mitton 2000), while studies based on maternally inherited mitochondrial markers estimate N_m to be near zero (Mitton et al. 2000). These disparate results suggest gene flow mediated by pollen movement is high, even between widely separated populations, but infrequent when mediated by seed movement even though seed in this species is typically distributed by birds and mammals.

Migration rates in three species of *Taxus* (yews), the seed of which is also distributed by birds, have consistently exceeded 3.0, even though populations of these species tend to be quite widely dispersed and separate (El-Kassaby and Yanchuk 1994; Chung et al. 1999; Lee et al. 2000; Klumpp and Dhar 2011).

Perhaps more appealing, intuitively, are studies that attempt to illustrate distances genes flow from point or stand sources, based on direct measures of paternity and parentage. A review of eight studies conducted in natural stands of conifers shows average pollen dispersal distances that vary from less than 10 m in *Picea abies* and *Pinus densiflorus* to over 160 m in *Araucaria angustifolia* (Bagnoli et al. 2011). These same studies estimated the proportion of seeds or seedlings sampled that resulted from long-distance pollen paternity to vary from 1% to over 50%. Studies conducted in carefully controlled clonal seed orchards of conifers using paternity exclusion analyses reveal a classic leptokurtic curve for effective pollen dispersal, with the majority of pollen from individual male parents falling within a few 100 m or less (Erickson and Adams 1989; Williams 2009). As in the natural stands noted previously, however, rates of pollen contamination from sources far from the orchards often exceeded 30% (range = 4–89%; reviewed in Wheeler and Jech 1992; Adams and Burczyk 2000). Anecdotal evidence of long-distance seed and pollen dispersal, noted in Chap. 1 of this volume, clearly illustrate the magnitude of gene flow possible in conifers. Though gene flow may vary substantially among species, populations, seasons, and even individual plants, at physical isolation distances of hundreds to thousands of meters, gene flow levels are sufficient to counteract genetic drift and moderate levels of directional selection (Wheeler and Jech 1992).

Genetic Drift

Genetic drift refers to the random change in population allele frequencies due to sampling error. The term "random drift," first used by Wright (1929), is generally discussed in cases related to two additional population phenomena, bottlenecks and founder effects. A bottleneck event occurs when a taxon's population number is drastically reduced, as may occur due to environmental catastrophe (fire, volcanism, glaciation) or human-induced habitat loss. In time, such small populations are susceptible to loss of low-frequency alleles due to increased inbreeding and random mating opportunities. Species such as *Pinus torreyana, P. resinosa, P. pinea,* and *Wollemia nobilis* are often cited as likely having gone through extreme bottlenecks that have reduced their genetic diversity to, or nearly to, zero (Ledig and Conkle 1983; Simon et al. 1986; Peakall et al. 2003). In the case of *Wollemia*, it appears that the entire extant population of individual trees descended from a single surviving ancestor (Peakall et al. 2003).

Founder effects refer to the random sampling of individuals from a large or ancestral population that contributes to the establishment of new populations as may occur when the range of a species expands following glaciation or alternatively when a range shrinks as might occur due to climate change, leaving scattered and

isolated populations in ecologically desirable refugia (Schuster and Mitton 2000; Godbout et al. 2008). The literature on genetic diversity measured with neutral genetic markers is replete with references to founder effects (Table 9.2, Study objective = Biogeography/Demographics).

Case Studies

Summary statistics, such as those noted in the previous sections, allow geneticists to make relatively broad and sweeping statements about the nature of genetic diversity, as measured using neutral genetic markers, for conifers, collectively. Thus, conifers appear to maintain large stores of genetic variation, most of which is found within populations. For widespread species, with large populations, gene flow is significant and genetic differentiation between populations is modest (<10%). Geologic and climatic events (e.g., mountain building, glaciation) that result in range fragmentation lead to increased population differentiation due primarily to genetic drift. Migration, particularly following large-scale disturbance, leaves genetic clues about such things as location of refugial populations, migratory routes, and current and ancient introgressive events. We have also learned that conifers are primarily outcrossing and tolerate inbreeding poorly, though exceptions occur.

In the following section we strive to present case studies that elucidate, in greater detail, results of experiments that address a broad spectrum of research objectives. Early studies concentrated on characterizing diversity and population structure and soon expanded to investigations of mating systems and species biogeography. We have chosen a single species, *Pinus contorta*, to illustrate in some detail many of the facets discussed earlier in this chapter and conclude with examples of how neutral genetic markers have provided insight on how forestry and applied tree improvement have influenced genetic diversity in domesticated and/or managed populations of commercial conifer species.

Diversity, Population Structure, and Biogeography

Pinus contorta (lodgepole pine) is one of the most widespread conifers in the world and has been an important component of the temperate and sub-boreal conifer forests of western North America for several million years (Critchfield 1984a). Discussed at some length in Chap. 8 with respect to geographic variation, the species has enjoyed widespread attention by geneticists, including several studies of neutral genetic variation (Yeh and Layton 1979; Wheeler and Guries 1982a, b, 1987; Wheeler et al. 1983; Dong and Wagner 1993; Yang and Yeh 1995; Fazekas and Yeh 2006; Godbout et al. 2008). Interest in the population genetics of the species is elicited by several factors: much of the species range currently exists in areas glaciated as recently as 10,000–14,000 years ago (biogeography), its range is sympatric with the range of a closely related species (*Pinus banksiana*) in Central Alberta, Canada (hybridization and introgression), and it is often recognized as consisting of four largely allopatric subspecies or varieties (phylogenetics), based on

morphological traits (Critchfield 1957). We draw from these works to illustrate several applications of the study of neutral genetic variation in lodgepole pine beginning with a comprehensive investigation focused on population structure and diversity.

Forty-two allozyme loci were assayed in 1895 trees (35–60 trees per population), representing 33 populations from across much of the range of *Pinus contorta,* including all four currently recognized subspecies (Fig. 9.1, Wheeler and Guries 1982b). Standard measures of genetic diversity within and differentiation among populations and subspecies were calculated for both allozymes and an array of cone and seed traits measured on the same trees.

Of the 42 loci studied, 35 were polymorphic in at least one population, and the remaining 7 were essentially monomorphic throughout the species. This is a relatively common finding in conifers when many loci are studied. For most polymorphic loci (32 of 35), the same allele was most common in all populations, though actual allele frequencies often varied. For three loci, two or more alleles occurred in nearly equal frequencies, a condition sometimes associated with balancing selection. Overall, measures of genetic diversity in lodgepole pine reported in this study (H_e = 0.116; A = 1.85; P = 68%) were low relative to other widespread woody species, including angiosperms (H_e = 0.228; A = 2.56; P = 74.3%), though not too dissimilar from averages for other conifers (H_e = 0.151, A = 1.83; P = 53.4%). As is common in such studies, diversity measures varied considerably among populations (Table 9.3). The authors concluded that contrasts with diversity measures in other studies likely reflect, in part, differences in sampling and locus selection as much as real genetic differences.

Despite the dramatic amplitude in ecological conditions under which the populations in this study grew, the distribution of genetic variation remained predominantly within populations (90.7%) with only 3.2% found among populations and 6.1% among subspecies. In contrast, the distribution of variation for 12 seed and cone morphometric traits was 43.9%, 18.6%, and 37.6%, for within and among populations, and among subspecies, respectively. Average genetic distance (D) estimates between populations within subspecies (range, 0.23–0.66%) and between subspecies (range, 0.82–1.86%) were uniformly low and less than that typically seen for widespread conifer species. Much larger estimates of D have been observed between populations of *Taxus brevifolia* (Wheeler et al. 1995a, D = 5.3%), *Pseudotsuga menziesii* (Cruz-Nicolás et al. 2011, D = 12.5%), and *Picea mariana* (Yeh et al. 1986, D = 1.6%). A separate study, based on RAPD markers, found similarly modest differences both within and between subspecies of lodgepole pine (Fazekas and Yeh 2006), prompting the authors to suggest subspecies status was unwarranted in this species. Based on the substantially more diagnostic differences exhibited by morphometric traits in the Wheeler and Guries study (Wheeler and Guries 1982b), the authors concluded the taxonomic treatment of Critchfield (1957) was legitimate. The obvious question raised by this difference of opinion relates to what level and type of molecular variation constitute taxonomic status designation. In this case, allozymes likely do not accurately reflect the extent of speciation that has occurred.

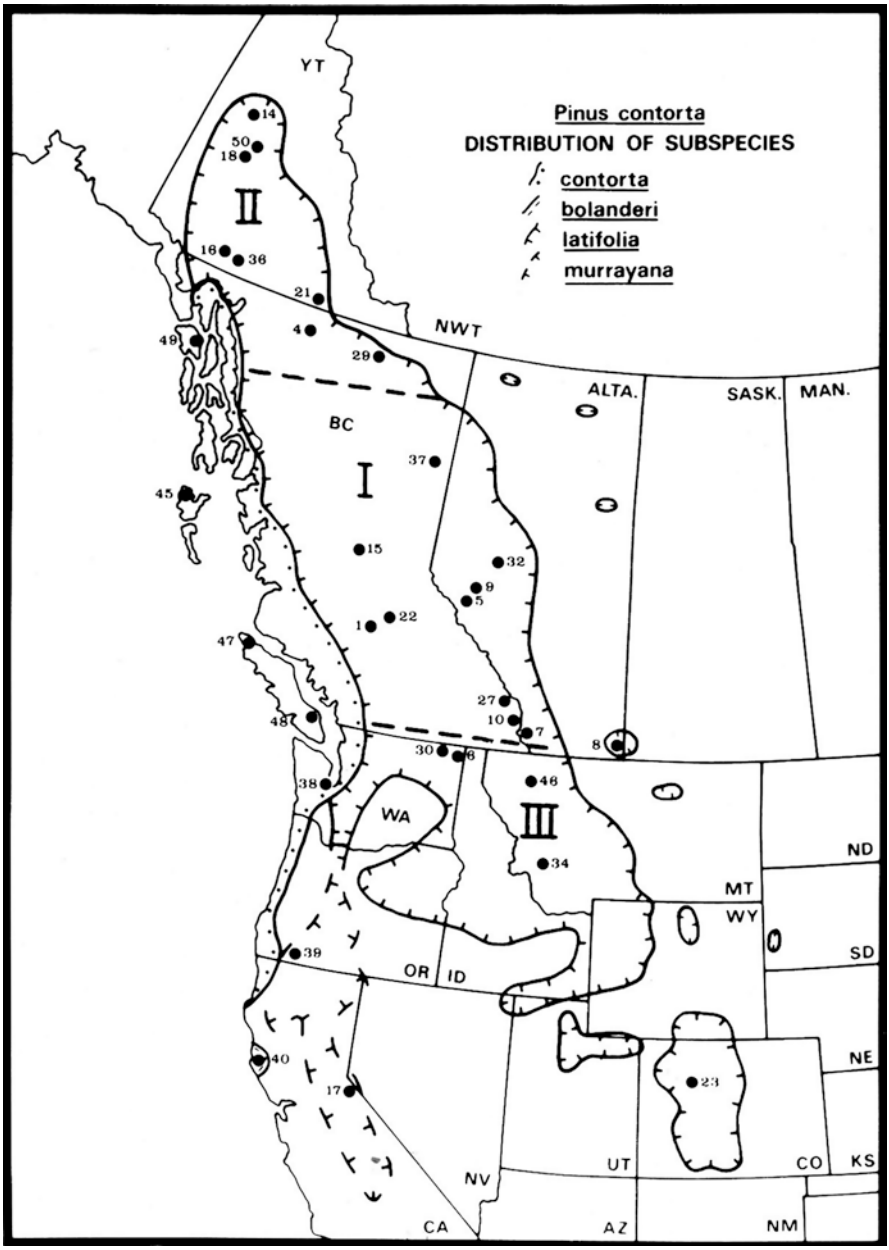


Fig. 9.1 The natural ranges for *Pinus contorta* subspecies and distribution of sampled populations (•). Groups define biogeographic regions (I, *central* Southern British Columbia and Alberta; II, Yukon and Northern British Columbia; III, United States, mostly south of the continental ice sheets of the last glaciation event). (From Wheeler and Guries 1982a)

Table 9.3 Number of trees sampled (*N*), mean expected heterozygosity ($\mathbf{H_e}$), mean proportions of polymorphic loci (**P**), and mean number of alleles per locus (**A**), presented by population and averaged over subspecies for a range-wide collection of *Pinus contorta*

Population	N	H_e[a]	P[b]	*A*
Pinus contorta				
Ssp. *latifolia*				
1	60	0.114 ± 0.026	0.70	1.86
4	60	0.125 ± 0.029	0.61	1.69
5	60	0.103 ± 0.025	0.70	1.81
6	60	0.128 ± 0.028	0.73	1.93
7	60	0.144 ± 0.028	0.82	2.24
8	60	0.102 ± 0.025	0.73	1.88
9	60	0.120 ± 0.027	0.70	1.95
10	60	0.115 ± 0.026	0.70	2.04
14	45	0.108 ± 0.025	0.64	1.71
15	60	0.114 ± 0.028	0.70	1.83
16	60	0.129 ± 0.029	0.70	1.76
18	60	0.109 ± 0.028	0.64	1.76
21	60	0.113 ± 0.028	0.79	1.83
22	60	0.113 ± 0.027	0.67	1.83
23	50	0.119 ± 0.028	0.64	1.81
27	60	0.111 ± 0.025	0.76	1.97
29	60	0.119 ± 0.028	0.73	1.86
30	60	0.114 ± 0.027	0.61	1.76
32	60	0.105 ± 0.025	0.64	1.81
34	60	0.116 ± 0.026	0.73	2.09
36	50	0.130 ± 0.031	0.61	1.76
37	60	0.128 ± 0.029	0.70	1.83
46[c]	60	0.097 ± 0.026	0.58	1.67
50	60	0.144 ± 0.030	0.67	1.88
Mean	58	0.118 ± 0.027	0.69 ± 0.06	1.86 ± 0.12
Ssp. *contorta*				
45[c]	55	0.130 ± 0.030	0.61	1.88
47[c]	60	0.123 ± 0.027	0.70	1.88
48[c]	60	0.125 ± 0.029	0.64	1.18
49	35	0.114 ± 0.029	0.55	1.57
38	60	0.139 ± 0.029	0.73	1.93
Mean	54	0.126 ± 0.029	0.65 ± 0.07	1.81 ± 0.14
Ssp. *murrayana*				
17	60	0.120 ± 0.026	0.70	1.86
39	60	0.128 ± 0.028	0.76	2.02
Mean	60	0.124 ± 0.027	0.73 ± 0.03	1.94 ± 0.11
Ssp. *bolanderi*				
40	40	0.109 ± 0.028	0.58	1.62

[a]Calculated on the basis of allele frequencies for 42 loci
[b]Frequency of most common allele <0.99
[c]*N* for these bulk samples represents the number of seeds assayed per enzyme system
From Wheeler and Guries (1982b)

Two of the four subspecies of *P. contorta* (ssp. *bolanderi* and ssp. *murrayana*) currently occur exclusively south of the Wisconsin glacial maximum, which terminated near the Canada–United States border. Most of the present-day range of the remaining subspecies (ssp. *contorta* and ssp. *latifolia*) was glaciated until rather recently (10,000–14,000 years ago). For *P. contorta*, the obvious biogeographical quandary concerns the number and location of glacial refugia and migratory routes the species used in recolonizing this extensive range (approximately 2400 km north/south and 800 km east/west). The traditional view held that the species survived in refugia south of the Cordilleran and Laurentide ice sheets and rapidly moved northwest, at rates as high as 220–1000 m/yr, following ice melt (Heusser 1965; Strong and Hills 2013) though the possibility of refugia existing in Alberta and the Yukon Territory in Canada and coastal areas of Alaska and Canada were proposed as early as 1937 (reviewed in Wheeler and Guries 1982a). A small ice-free area in west central Yukon and ocean levels as much as 130 m lower than present offered opportunities for vegetation to survive during the last glacier maximum. Palynological studies have consistently been cited as evidence that the species survived south of the ice sheet alone (Cwynar and MacDonald 1987; Strong and Hills 2013; early works reviewed in Wheeler and Guries 1982a) but several lines of genetic evidence suggest otherwise.

Based on levels of genetic diversity, the occurrence and distribution of private or rare alleles, and the degree and pattern of between-population genetic distance measures of allozymes, Wheeler and Guries (1982a) concluded that populations of ssp. *contorta* and ssp. *latifolia* survived in multiple northern, coastal, and possibly nunatak refugia. Predictions for genetic parameters based solely on rapid migration north from southern refugia would include: (1) declining levels of genetic diversity from south to north, (2) isolation by distance or a positive correlation between genetic distance and geographic distance, and (3) a decline in the number of private alleles from south to north. Virtually none of these conditions were met.

Relative to other *P. contorta* subspecies, current coastal populations of ssp. *contorta* occupying areas north of the maximum ice sheet terminus are, on average, the most variable in the species (mean $H_e = 0.126$), exhibit only modest genetic affinity for one another (mean D within subspecies = 0.0066), and retain 44% of the private alleles found in the species. The average genetic distance between pairs of coastal populations is two to four times the average genetic distance observed between pairs of ssp. *latifolia* populations and nearly comparable to the D between the two subspecies (0.007; Fig. 9.2). Of the 11 rare alleles observed in the coastal populations, four are unique to the subspecies, seven are shared with the northern-most populations of ssp. *latifolia* (group II, Fig. 9.1), six are shared with the southern-most populations of ssp. *latifolia* (group III, Fig. 9.1), and only three are shared with centrally located populations (group I, Fig. 9.1). The average number of private alleles observed per population (4.6) for these coastal populations was the highest observed for all lodgepole groupings. These findings imply the species survived in one or more coastal refugia on coastal islands of British Columbia, Canada, and Southeastern Alaska. This should no longer be considered an unusual finding. Studies of the phylogeographic structure of *Pinus banksiana* and *Picea mariana*

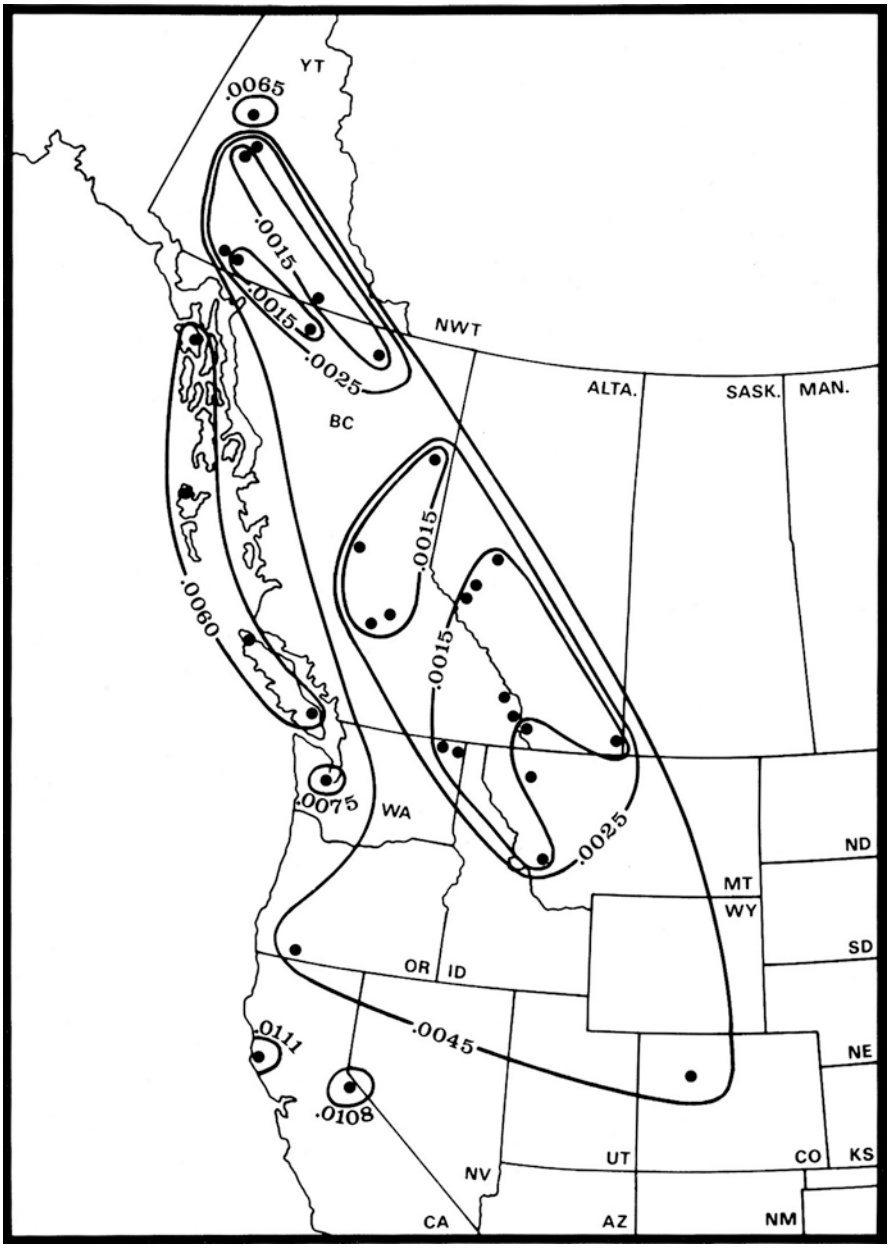


Fig. 9.2 Population clustering based on UPGMA analysis of genetic distance values. (From Wheeler and Guries 1982a)

support the speculation that coastal glacial refugia existed for these species along the now-flooded Atlantic seaboard during the last glaciation (Yeh et al. 1986; Godbout et al. 2010).

Equally compelling are genetic data supporting northern refugia for ssp. *latifolia* (Fig. 9.2). The average genetic diversity of the northern populations (H_{ex} = 0.122, Group II, Fig. 9.1) exceeded that of the population groups from central British Columbia (H_{ex} = 0.115, Group I) and south of the ice limit (H_e = 0.115, Group III). Population clustering based on analysis of genetic distance values (Fig. 9.2) clearly shows northern-most populations to be distinct from those populations in the south. The most distinct population was also the most northerly known population of lodgepole pine in Yukon (#14, Figs. 9.1 and 9.2). Interestingly, it is also located very close to or within the area known to be ice-free during the last glacial maximum.

Rare allele number and distribution paint a similar picture. The average number of rare alleles for southern (4.4), central (2.6), and northern (3.6) populations are inconsistent with loss of diversity during rapid migrations. Southern and central populations share six private alleles not found in the northern populations. Only 4 of the 12 private alleles found in northern populations are found in southern populations of ssp. *latifolia*.

More recent and comprehensive studies support the biogeographical interpretations noted above, based on RFLP markers (Fazekas and Yeh 2006) and a mitochondrial minisatellite marker (Godbout et al. 2008). The latter study, which included sampling of 91 populations, presents compelling results consistent with species re-establishment from multiple genetically distinct and widely separate refugia (Fig. 9.3). Godbout et al. (2008) studied mitochondrial haplotype diversity which exhibited far greater population differentiation (G_{ST} = 0.365 and R_{ST} = 0.568) than observed for allozymes. Population structure based on a Bayesian analysis (BAPS) revealed eight separate groupings, with distinct coastal and northern groups evident as well as multiple sources for northern expansion from areas south of the ice. It is interesting to note that, relatively recently, a new variety of *P. contorta* was described (*Pinus contorta* Douglas ex Loudon var. *yukonensis* W.L. Strong; Strong 2010). The author notes this newly described variety may be a taxon representing a possible Beringian survivor. This would be consistent with the concept of incipient speciation spurred by glacial vicariance events.

Conservation and Mating Systems

Neutral genetic markers are frequently used to characterize genetic variation and guide conservation efforts for conifer species considered to be threatened in some manner (Table 9.2), and an entire chapter in this volume is dedicated to the topic of conservation and restoration (Chap. 13). Still, the thoroughly interesting case of a rare Mexican pinyon pine is worthy of a brief visit here.

There is a rich literature focusing on conservation of Central American conifers, especially the pines and spruces. Mexico and neighboring Central American countries are home to as many as 50 pine species, many of which occur in fragmented

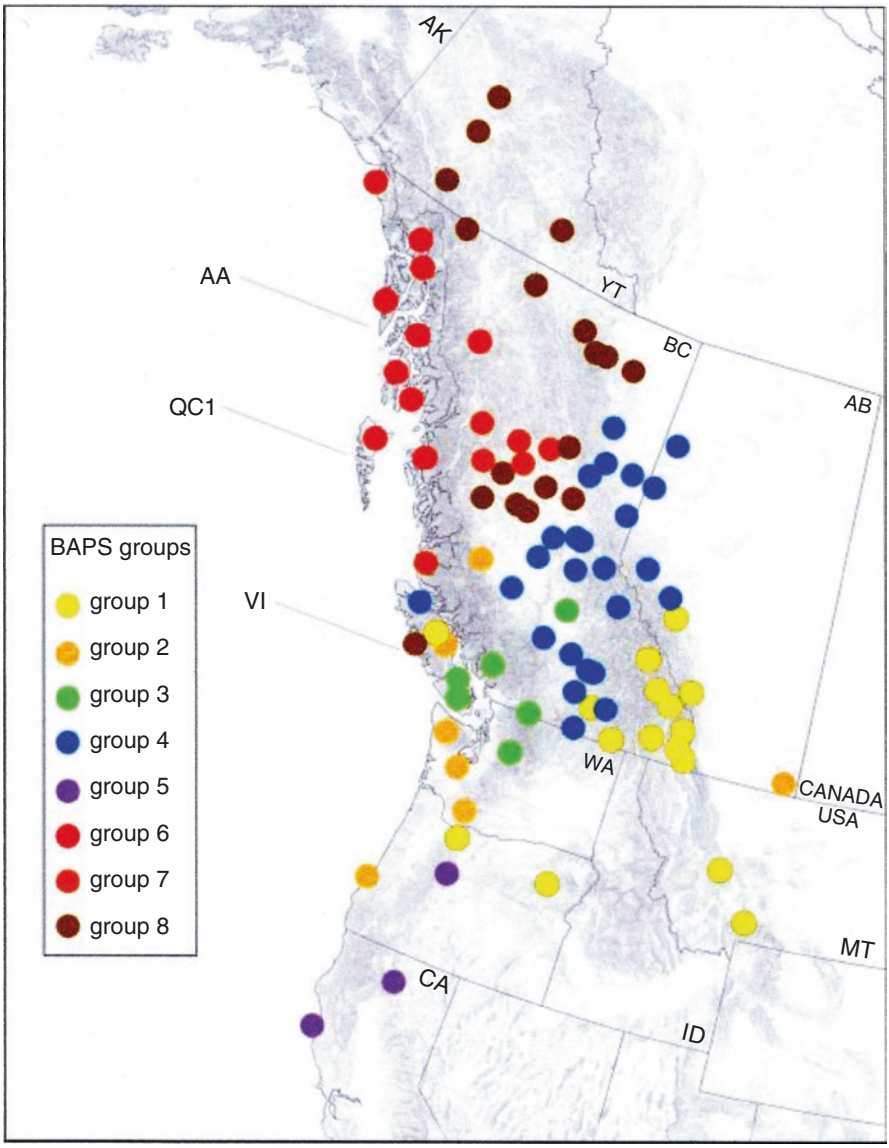


Fig. 9.3 Geographical distribution of 91 lodgepole pine populations following their membership to each of the eight groups representing the optimal solution delineated by Bayesian analysis (BAPS). State/province abbreviations: AB Alberta, AK Alaska, BC British Columbia, CA California; ID Idaho, MT Montana, WA Washington, YT Yukon Territory. Island abbreviations: AA Alexander Archipelago, QCI Queen Charlotte Islands, VI Vancouver Island. (From Godbout et al. 2008)

and widely dispersed populations in mountainous terrain and are listed as threatened or endangered. Among the rarest of these is the maxipinyon or *Pinus maximartinezii*, a unique pinyon pine represented by a single population of <2500 individuals located on ~400 ha in southern Zacatecas, Mexico. It is distinguished among the pinyons by having very large, woody cones and large seeds (~900 seeds per kilogram). The seeds are highly valued, annually collected, and sold in local markets. Natural reproduction of the species, not surprisingly, is rare to non-existent today. An allozyme study of the species provided fascinating insight into the genetic diversity, mating structure, and evolutionary history of this species (Ledig et al. 1999).

The authors found that the proportion of polymorphic loci in *P. maximartinezii* was quite low (P = 30.3%), but genetic diversity was higher than expected (H_e = 0.122). Additionally, estimates of outcrossing rates were higher than expected given the population size (t_m = 0.816), though there was significant variation in estimates for individual trees, ranging from 0.42 to 1.00, with a distinct bimodal distribution among sampled trees. The average inbreeding coefficient (F) was 0.081. The most remarkable facet of the study related to the nature of the 10 polymorphic loci (of the 33 loci studied). In all cases, only two alleles were observed and with a single exception, allele frequencies were intermediate (Table 9.4), patterns rarely if ever seen in other conifers. The typical pattern for conifers includes multiple alleles per polymorphic locus with one common allele and the rest occurring at low frequency (< 0.10). In addition, significant linkage disequilibrium was found in both pollen and maternal gametes, and in most cases the disequilibrium was in the same direction. None of the polymorphic loci studied here are known to be genetically linked in conifers, however.

Ledig et al. (1999) speculated two possible evolutionary scenarios that could have led to the observed patterns of diversity, allozyme frequencies, numbers of alleles per locus, and measures of disequilibrium. The first is that the species

Table 9.4 Allele frequencies (f) and fixation indices (F) for ten polymorphic loci in *P. maximartinezii* in CGF-1995

	f			
	Allele 1	Allele 2	n^a	F
Ald2	0.615	0.385	96	0.033
Got3	0.240	0.760	96	0.085
Lap2	0.490	0.510	96	0.125
Mnr1	0.589	0.411	96	0.158
Mpi	0.761	0.239	71	−0.005
Pgi2	0.552	0.448	96	0.130
Pgm	0.995	0.005	96	0.225
6Pg1	0.568	0.432	96	0.032
Skd1	0.786	0.214	96	−0.083
Tpi1	0.561	0.439	82	0.109
Mean				0.081

[a]Number of diploid individuals sampled
From Ledig et al. (1999)

survived a severe bottleneck, perhaps even as a single seed, and subsequent rapid expansion in the relatively recent past. Based on measures of disequilibrium, the bottleneck would have had to occur within the last four to five generations or less than 1000 years ago. The second alternative is that the species could have originated as a single mutant individual or as a unique, monotypic hybrid between two other pine species. A tantalizing suggestion by the authors is the possibility of human intervention in selecting for a large-seeded variant. In either scenario, the origin of this species from a single seed seems to be a distinct possibility

Effects of Forest Management and Tree Improvement on Genetic Diversity

As the ability to detect and characterize genetic diversity developed, forest geneticists found many applications of neutral molecular markers for quantifying the effects of forest management and tree improvement activities (Savolainen and Kärkkäinen 1992; Wheeler et al. 1995b; El-Kassaby and Ritland 1996b; Chap. 13). Research of this nature was driven by concerns that forestry practices would erode genetic diversity in subsequent generations as a function of domestication activities and population size reductions. In general, most studies in conifers have shown genetic erosion due to forestry and tree improvement activities to be minimal to nil. A few of these studies are reviewed here. Markers were also readily adopted for monitoring and quantifying a wide range of applied tree improvement and seed orchard activities such as supplemental mass pollination, pollen contamination studies, and paternal contributions from polymix crosses or wind pollination (Adams 1983; Wheeler and Jech 1992; Adams et al. 1997; Stoehr et al. 1998; Lambeth et al. 2001; El-Kassaby and Lstibůrek 2009).

Savolainen and Kärkkäinen (1992) reviewed a handful of papers that contrasted genetic diversity and outcrossing rates between natural stands and managed populations (seed tree or shelterwood stands or stands established with seeds from seed orchards based on selected parents). Overall, managed stands possessed as much or more diversity and higher outcrossing rates than natural stands. One of these studies (Neale 1985) compared several measures of genic diversity (P, H_e, A) for four life-cycle stages in *Pseudotsuga menziesii* var. *menziesii* (natural stands, adjacent shelterwood stand leave trees, seed crops from the leave trees, and 3- to 5-year-old regeneration in the shelterwood stand). No significant differences were noted for any of the measures of diversity. In a more recent and comprehensive study, conducted with *Picea glauca* (white spruce) in Northern Alberta, the effects of harvesting at increasing intensities on several measures of genetic diversity and population structure were estimated using both genomic and *EST*-based (expressed sequenced tags) nSSRs (Fageria and Rajora 2013). The study compared unharvested control stands (CON) and naturally regenerated seedlings from post-harvest stands with 75% natural stand stem retention (75R), 50% retention (50R), 20% retention (20R), 10% retention (10R), and clearcuts (CCT) with 2% stand retention. The replicated study was done in both conifer-dominated and mixed-wood forest types. Contrary

to expectations, the authors detected no significant effect of harvest intensity on genetic diversity, inbreeding levels, or population structure in this study (see Table 9.5). The potential effects of genetic drift and inbreeding resulting from harvesting-induced bottlenecks were apparently counterbalanced by the predominantly outcrossing mating system and high gene flow from the residual and surrounding stands. Similarly, measures of genetic diversity were effectively unchanged by impacts of forest fires, forest harvesting levels, and alternative reforestation practices in *Picea mariana* (Rajora and Pluhar 2003).

Tree improvement activities represent an incrementally greater threat to reducing genetic diversity than forest silvicultural practices. The goal of such activities is to select, typically for improved productivity or wood quality traits, while maintaining overall adaptability of deployed populations. This is done by iterative stages of breeding, testing, and selection known as recurrent selection (see Fig. 9.4). The process ultimately results in reduced numbers of genotypes and effective population sizes that contribute to future forests. Improved materials are typically deployed as seedlings produced from seed orchards containing selected parent trees. Intensive tree improvement of conifer species began in the mid-1900s (Wheeler et al. 2015) in most parts of the world, and the most advanced programs have gone through three

Table 9.5 Mean genetic diversity parameters and fixation index (F) for unharvested control and post-harvest natural regeneration of *Picea glauca* in the conifer-dominated (CD) and mixed-wood (MX) forest based on ten microsatellite loci. Treatment abbreviations are defined in the text

HT	A_T	A	A_e	A_R	A_P	H_o	H_e	F
CD								
CON	112.0	11.20	6.50	10.66	0.5	0.492	0.639	0.175
75R	114.5	11.45	7.15	10.38	0.5	0.529	0.656	0.191
50R	108.5	10.85	6.87	9.90	0.5	0.572	0.675	0.152
20R	108.5	10.85	6.14	9.84	0.5	0.523	0.639	0.161
10R	109.5	10.95	7.12	10.28	1.0	0.557	0.659	0.162
CCT	112.5	11.25	6.49	10.05	1.5	0.526	0.666	0197
MW								
CON	114.0	11.40	6.21	10.24	1.0	0.537	0.656	0.171
75R	115.5	11.55	7.28	10.66	0.5	0.536	0.669	0.208
50R	114.5	11.45	6.90	10.40	0.5	0.545	0.651	0.140
20R	111.0	11.10	6.74	10.44	0.0	0.523	0.648	0.202
10R	109.0	10.90	6.82	9.71	1.0	0.440	0.645	0.253
CCT	116.5	11.65	7.03	9.84	0.0	0.553	0.658	0.127
Mean	110.9	11.09	6.71	10.20	0.7	0.533	0.655	0.173
CD								
Mean	113.8	11.38	6.83	10.30	0.5	0.529	0.655	0.177
MW								

Details of harvesting treatments (HT) are provided in Table 1 of Fageria and Rajora (2013). A_T total number of alleles, A mean number of alleles per locus, A_e effective number of alleles per locus, A_R allelic richness, A_P private alleles, H_o mean observed heterozygosity, H_e mean expected heterozygosity, ANOVA did not show any significant differences among harvesting treatments for all eight parameters, as indicated in Table S4 in Fageria and Rajora (2013)
From Fageria and Rajora (2013)

to four cycles of improvement. Most aggressive tree improvement programs seek to retain as much genetic diversity as possible while still meeting their productivity goals. Here we report on two comprehensive studies, both done with Douglas-fir (*Pseudotsuga menziesii*), that used allozymes to monitor diversity in successive cycles of breeding, testing, and selection (Wheeler et al. 1995b; El-Kassaby and Ritland 1996b).

Both studies characterized expected heterozygosity (H_e), proportion of polymorphic loci (P_{95}), and the number of alleles per locus ($\boldsymbol{A}$) using 20–22 loci for populations representing four to five successive stages of domestication (natural stands,

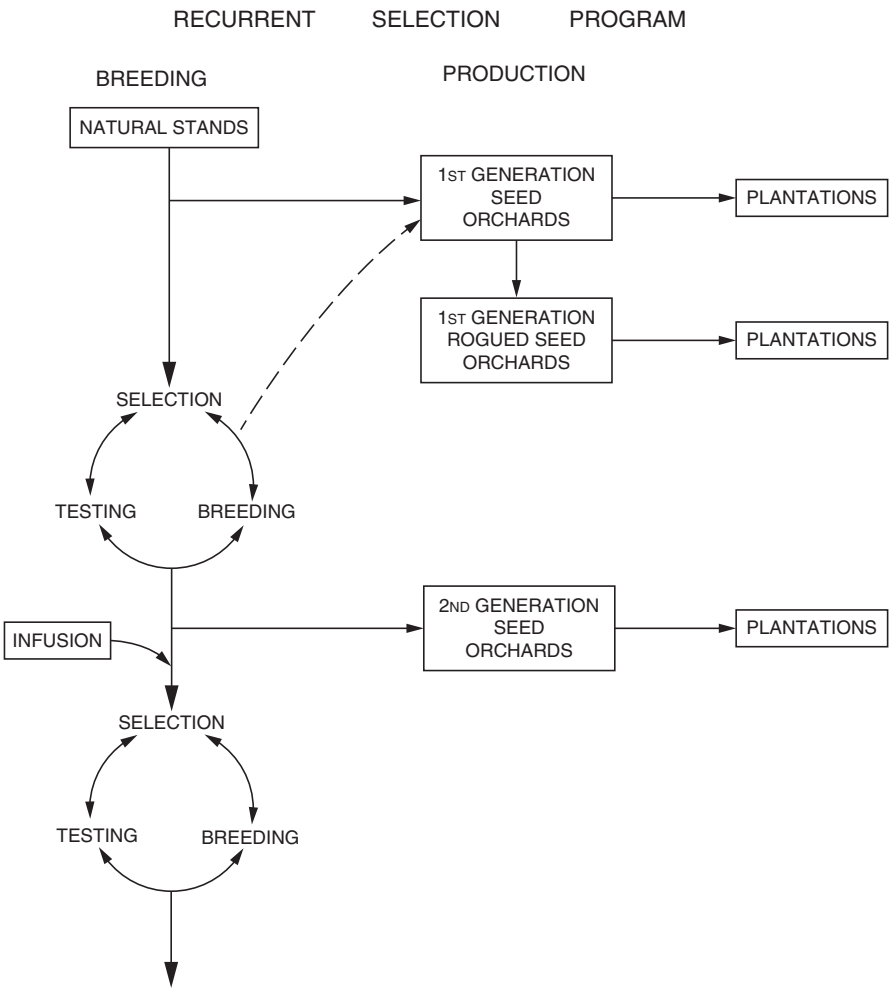


Fig. 9.4 Diagram depicting a *Pseudotsuga menziesii* tree breeding program showing the relationship between natural populations and first- and second-generation seed orchards as well as an infusion population. (Modified from El-Kassaby and Ritland 1996b)

Table 9.6 Measures of genetic diversity for natural stands and progressively more select (fewer numbers of individuals per population selected for growth performance) populations of *Pseudotsuga menziesii* (Douglas-fir) in two studies from the Pacific Northwest United States. Study 1, El-Kassaby and Ritland (1996b); Study 2, Wheeler et al. (1995b)

Population	N		P_{95} (%)		A		H_e	
	Study 1	Study 2	Study 1	Study 2	Study 1	Study 2	Study 1	Study 2
Natural stands	49	24	52.6	77.0	2.1	3.5	0.171	0.285
1st gen production Orch.	12	6	62.5	68.0	2.3	3.5	0.172	0.277
1.5 gen production Orch.	12	6	60.4	68.0	2.2	3.3	0.173	0.265
2nd gen production Orch.	4	6	56.3	77.0	2.3	3.0	0.163	0.240
Rooted cutting parents		1		73.0		2.8		0.266

Table 9.7 Effect of increasing levels of selection for growth traits and reduced population sizes (the number of parental contributors) on the number and frequency of rare alleles in tree improvement populations of *Pseudotsuga menziesii* (Douglas-fir). RC – rooted cutting population

Total number of alleles occurring with a frequency of		Proportion of alleles with frequencies that increased in subsequent generations			Proportion of alleles that were lost in subsequent generations		
		1st	2nd	RC	1st	2nd	RC
≤0.10	42	50.0	33.3	40.4	2.4	23.8	38.1
0.05–0.10	10	50.0	10.0	60.0	0	10.0	10.0
0.01–0.049	15	40.0	46.7	40.0	0	6.7	20.0
0.001–0.009	17	58.8	35.3	29.4	5.9	47.1	70.6

Modified from Wheeler et al. (1995b)

first-generation or cycle-production seed-orchard parents, rogued or 1.5-generation orchards, and second-generation seed orchards; see Table 9.6). Wheeler et al. (1995b) included a fifth level that incorporated a subset of second-generation parents used in a rooted cutting deployment trial. Sample sizes were large, except for the rooted cutting populations, and included replication for all levels. Average population sizes were > 250, 100, 45, and 30 for natural populations, first-generation orchards, rogued orchards, and second-generation orchards in the Wheeler et al. (1995b) study and > 100, 78, 61, and 55 for the El-Kassaby and Ritland (1996b) study. The rooted cutting populations ranged from 9 to 37 per year, over a 5-year period of study.

Measures of diversity were unaffected or declined slightly in the intensively managed populations (Wheeler et al. 1995b) but generally increased slightly in the less intensively managed program (El-Kassaby and Ritland 1996b). In no instances were differences statistically significant except for loss of alleles in the rooted cutting population. The increases in diversity measures noted in the latter study were attributed to the mixing of select trees from different stands (populations) into one artificial population, likely devoid of any inbreeding effects. A similar result was noted in *Pinus taeda* (Schmidtling et al. 1999) and *Picea sitchensis* (Chaisurisri and

El-Kassaby 1994). The differences in absolute diversity measures between the studies can be attributed primarily to the loci studied and possibly to a lesser extent the interpretation of variation observed on electrophoretic gels.

The greatest influence on genetic diversity of the bottlenecking practice of tree improvement is to reduce the number of rare alleles observed in successive generations of domestication where rare is defined as occurring with a frequency of less than or equal to 10%. For the 22 loci analyzed by Wheeler et al. (1995b), 77 alleles were observed, of which 42 (54.5%) were considered rare as defined above (see Table 9.7). Of these 42 alleles, 17 occurred with a frequency of less than 1.0%. Loss of these very-low-frequency alleles was significant, from 1 in the first generation to 8 and 12 in the second generation and rooted cutting crops (Table 9.7). Loss of alleles with frequencies >1.0% was slight (0, 2, and 4, respectively). Ironically, the frequency of most surviving rare alleles increased in the selected populations, contributing to enhanced measures of heterozygosity in the rooted cutting crop.

It is generally accepted that rare alleles are rare for good reason – they are typically deleterious, and their loss is not of concern, though certainly a small proportion of such alleles may be of adaptive significance (Wheeler and Jech 1992). In a study of controlled biparental crosses in *Pseudotsuga menziesii*, Bongarten et al. (1985) found a statistically significant decline in family performance as the number of rare alleles in the cross increased, thus supporting the notion that rare alleles are deleterious in general.

While the results of the studies noted here suggest most early-stage efforts at tree domestication have had relatively little effect on levels of genetic diversity in managed populations, it is likely that successive cycles of breeding, testing, and selection will ultimately lead to increased inbreeding and loss of diversity unless aggressive steps are taken to prevent it. Furthermore, using neutral markers to track genetic diversity, while convenient in the past, is now rather obsolete. In the future, markers, such as SNPs, located in expressed genes, would be preferable for tracking the effects of domestication on genes under selection, particularly for adaptive evolutionary potential.

Summary

The neutral theory of molecular evolution holds that most genetic variation observed at the molecular level is evolutionarily neutral, the product of mutation, migration, genetic drift, and mating systems rather than selection. The development of an array of neutral genetic markers, beginning with allozymes in the 1970s and culminating, recently, with targeted genetic sequences, has provided tools to address a broad range of previously recalcitrant scientific inquiries, including population structure, phylogeography, phylogenetics, genetic conservation, and the evolutionary forces of migration, drift, and mating systems. In addition to studies in natural conifer populations, markers have been used extensively in managed and domesticated populations where knowledge of paternity and parentage is relevant.

The amount and distribution of genetic variation in conifers varies considerably by species, range, geographic distribution, and genome (nuclear vs. organelle), though in general, conifer species retain a great deal of genetic diversity within populations and small to modest levels of differentiation among populations (nuclear and paternally inherited chloroplast markers), indicative of a predominantly outcrossing mating system and extensive gene flow via pollen. Maternally inherited mitochondrial markers exhibit significantly lower levels of diversity within and greater differentiation among populations, reflective of the much lower gene flow associated with seed dispersal. Extreme examples of species exhibiting little-to-no genetic diversity or high levels of population differentiation are explained by genetic bottlenecks and founder effects.

Adaptive Genetic Variation

10

A Short History of Adaptive Genetic Variation in Conifers

Forest geneticists have a strong interest in understanding the distribution and patterning of genetic variation across forest landscapes, the nature of which is shaped by demographic (mating system, gene flow, and genetic drift) and adaptive (natural selection) processes. In Chap. 8, we discussed how these combined forces have resulted in patterns as revealed by common-garden studies. Such studies have been extremely important in guiding seed transfer and reforestation activities but do not explicitly disambiguate between the relative roles of demographic versus adaptive processes. One reason for this lack of clarity is that in common gardens highly polygenic traits are generally evaluated where the patterning of some individual genes reflects demographic processes and for others it is reflective of adaptive processes. To begin to separate the relative roles of demography versus adaptation, we discussed results obtained from neutral-marker studies in Chap. 9. In Chap. 11, we discuss genomic methods to dissect polygenic adaptive traits into their individual gene components. In this chapter, we turn our attention to studies that determine which genes may be nonneutral and thus underlie adaptive patterns of variation based on information in DNA sequences alone.

The ability to distinguish between genes, or more specifically the nucleotide polymorphism within genes, that may potentially be neutral versus nonneutral was enabled by the development of population genetic theory to analyze data resulting from DNA sequencing. We refer the reader to excellent texts and review papers on this topic (Nei 1987). In this chapter, we summarize the conifer genomic literature where these approaches have been used to identify genes, and genetic markers for these genes, that are thought to be nonneutral.

Two general approaches have been used to distinguish between neutral versus nonneutral genes. In the first, two measures of nucleotide diversity (π and θ) are obtained for a gene based on DNA sequencing of the gene in population samples. The technical term used for DNA sequencing in population samples is

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_10

resequencing, which is taken from the human genetic literature where it was first used. Estimates of π and θ have been reported for many genes from many conifer species (Table 10.1). Estimates of π and θ are often calculated based on *All* single-nucleotide polymorphisms (SNPs), just *Silent* SNPs (the combination of noncoding and synonymous sites), or just *Nonsynonymous* SNPs. In this chapter, we are most interested in the latter category as only nonsynonymous SNPs can lead to a change in the primary protein structure which could have selective (adaptive) differences among alleles. Estimates of π and θ based on nonsynonymous sites are given in Table 10.1, except where noted otherwise. Once estimates of π and θ have been made, they can be used to estimate a statistic for departure from neutrality and detect signatures of selection. Many such statistics have been proposed (see Nielsen 2005); however, the one most commonly used, and the one given in Table 10.1, is the Tajima D (Tajima 1989). Values of D significantly different from zero (either positive or negative) indicate a departure from neutrality or the effect of a demographic process. Positive values suggest balancing section or population contraction, whereas negative values suggest positive selection or population expansion. Thus, significant estimates of D can occasionally result from demographic forces and not selective forces. In this chapter, we will limit our discussion of evidence for nonneutral genes to the average estimate of D reported in the conifer literature and the percentages of genes having positive or negative estimates (Table 10.1). We refer the reader to the primary literature for a much more thorough treatment and discussion of neutrality tests applied to conifer DNA sequence data.

A second and widely used approach to identify genes potentially under selection are *outlier tests.* Here again, there is a rich theory and many different types of outlier tests. The original and most widely used test is the F_{ST} outlier test (Beaumont and Nichols 1996). The F_{ST} values are estimated from gene frequency data (SNP genotypic frequencies) among subpopulations. Those loci falling at the extreme end of the distribution are deemed outliers and are potentially genes under selection.

Neutrality tests and outlier tests have been performed in nearly 100 studies to date, although the number of conifer species studied remains rather limited. Neutrality tests have generally been performed on sequence data from a small number of candidate genes in a relatively small sample. In only a small number of studies have both types of tests been performed. In the next section, we summarize general trends that have been observed in conifers from this body of work, and in subsequent sections, we review what specifically has been learned in a number of conifer species.

General Trends in Patterns of Adaptive Genetic Diversity in Conifers Observed from Neutrality and F_{ST} Outlier Tests

Despite very large differences in experimental design and methods of analysis in the nearly 100 studies reported (Table 10.1), it is possible to observe some general trends regarding adaptive genetic diversity as inferred from DNA sequence data in individual genes. First, regardless of statistic used, only a small portion of genes

Table 10.1 Estimates of π and θ, average estimate of Tajima's D, percentage of genes with significant D value, percentage of genes having positive or negative estimates of D (both significant and nonsignificant), and the percentage of F_{ST} outliers for gene sequencing studies in conifers

Species	Research topic	Samples (1 N or 2 N)	Genes	SNPs	π	θ	Tajima D	Percent loci sig D	D +,−	F_{ST} outliers (%)	Reference
Pinus taeda	Wood formation	32	19	35	0.0011	0.0011	0.29	0	8, 13		Brown et al. (2004)
Pinus taeda	Drought stress	32	18	196	0.00166	0.0023	0.26	11	11, 7		González-Martínez et al. (2006)
Pinus taeda	Disease resistance, drought tolerance	32	41	307				10	5, 5		Ersoz et al. (2010)
Pinus taeda	Drought adaptation	907	3059	7216						0.78	Eckert et al. (2010b)
Pinus taeda	Diverse	18	5773	10,591	0.0013	0.0015	−0.47				Eckert et al. (2013b)
Pinus taeda	Growth and adaptation	375	37,718	2,822,609	0.0119					0.0005	Lu et al. (2016)
Pinus sylvestris	Allozyme vs. DNA	31–35	9		0.009		−0.73	0	1, 7		Pyhäjärvi et al. (2011)
Pinus sylvestris	Phytochrome genes	52	2		0.0003	0.0002					Garcia-Gil et al. (2003)
Pinus sylvestris	PAL gene	20	1	12	0.0003	0.0006					Dvornyk et al. (2002)
Pinus sylvestris	Cold tolerance	40	14	74	0.0039		−0.72	33	2, 13		Wachowiak et al. (2009)
Pinus sylvestris	Ionizing radiation	10	2	47	0.0054						Vornam et al. (2012)

(continued)

Table 10.1 (continued)

Species	Research topic	Samples (1 *N* or 2 *N*)	Genes	SNPs	π	θ	Tajima *D*	Percent loci sig *D*	*D* +,−	F_{ST} outliers (%)	Reference
Pinus sylvestris	Demography	40	16				−0.34		0, 4		Pyhäjärvi et al. (2007)
Pinus sylvestris	Bud set, cold tolerance	89	10	149	2.769 (gene)	2.257 (gene)	−0.85	0			Kujala and Savolainen (2012)
Pinus sylvestris	Diversity		17	275	0.0022		−1.17				Wachowiak et al. (2011a)
Pinus sylvestris	Diversity	42	12	48	0.0031		0.12	25	8, 4		Wachowiak et al. (2011b)
Pinus mugo	Adaptation	18	12	308	0.0369	0.0169	−0.68	0			Wachowiak et al. (2013)
Pinus mugo	Adaptation	12	190	900	0.0081	0.0082				1.6	Mosca et al. (2012)
Pinus mugo	Diversity		17	176	0.0022		−0.61				Wachowiak et al. (2011a)
Pinus mugo	Adaptation		383	663						2.9	Mosca et al. (2016)
Pinus uncinata	Adaptation	10	12	205	0.0054	0.0134	−0.12	0			Wachowiak et al. (2013)
Pinus uliginosa	Diversity		17	157	0.0018		−0.24				Wachowiak et al. (2011a)
Pinus uliginosa	Adaptation	3	12	169	0.0051	0.0126	−0.27	0			Wachowiak et al. (2013)
Pinus pinaster	Drought stress	53–197	11	62	0.0028					45	Eveno et al. (2008)

Pinus pinaster	Wood formation	24	8	3	0.0003	0.0005	<0.01	0	3, 0		Pot et al. (2005)
Pinus pinaster	Adaptation	122	6	131	0.0052	0.0045			17, 5		Grivet et al. (2010)
Pinus halepensis	Drought tolerance		10				0.06				Grivet et al. (2010)
Pinus halepensis	Adaptation	93	6	65	0.0031	0.0026			11, 5		Grivet et al. (2010)
Pinus radiata	Wood formation	23	8	2	0.0001	0.0002	0.002	0	0, 1		Pot et al. (2005)
Pinus radiata	Adaptation	447	38	149						6.7	Dillon et al. (2013)
Pinus contorta	Adaptation	92	24	124	0.0025					8.3	Eckert et al. (2012)
Pinus massoniana	Adaptation	60	25	321		0.0045	−0.07	16	10, 15		Zhou et al. (2014)
Pinus hwangshanensis	Adaptation	44	25	480		0.0069	−0.28	8	8, 17		Zhou et al. (2014)
Pinus lambertiana	Disease resistance	24	1	0	0	0					Jermstad et al. (2006)
				180	0.0013	0.0015	−0.4				Eckert et al. (2013b)
Pinus cembra	Diversity	9	171	284	0.0024	0.0025	−0.02			1.8	Mosca et al. (2012)
Pinus cembra	Adaptation	Many	265	455						4.9	Mosca et al. (2016)
Pinus section *Strobus* species	Adaptive	8 to 19	163								Eckert et al. (2013a)

(continued)

Table 10.1 (continued)

Species	Research topic	Samples (1 *N* or 2 *N*)	Genes	SNPs	π	θ	Tajima *D*	Percent loci sig *D*	*D* +,−	F_{ST} outliers (%)	Reference
Pinus strobus				282	0.0018	0.0022	−0.58				Eckert et al. (2013a)
Pinus longaeva				196	0.0014	0.0017	−0.58				Eckert et al. (2013a)
Pinus albicaulis				235	0.002	0.0018	−0.51				Eckert et al. (2013a)
Pinus aristata				83	0.0007	0.0008	−0.27				Eckert et al. (2013a)
Pinus ayacahuite				247	0.0017	0.0019	−0.45				Eckert et al. (2013a)
Pinus balfouriana				174	0.0011	0.0014	−0.78				Eckert et al. (2013a)
Pinus flexilis				278	0.002	0.0021	−0.58				Eckert et al. (2013a)
Pinus monticola				319	0.002	0.0024	−0.69				Eckert et al. (2013a)
Pinus strobiformis				338	0.0021	0.0027	−0.65				Eckert et al. (2013a)
Pinus monophylla				311	0.0022	0.0025	−0.56				Eckert et al. (2013a)
Pseudotsuga menziesii	Wood formation, cold hardiness	24	18	400	0.0021	0.0026	−0.25	11	5, 13		Krutovsky and Neale (2005)
Pseudotsuga menziesii	Cold hardiness	24	121	384	0.0028	0.002				6.6	Eckert et al. (2009)
Larix sibirica	Wood formation		2								Khatab et al. (2008)

Larix sibirica	Cold hardiness		8		0.0054	0.0067	−0.61	26	2, 6		Semerikov et al. (2013)
Larix occidentalis	Cold hardiness		8		0.0043	0.0046					Semerikov et al. (2013)
Larix decidua	Diversity	5	61	219	0.0078	0.0077	1.02			0	Mosca et al. (2012)
Abies alba	Diversity	8	70	197	0.0059	0.0059	−0.13			0	Mosca et al. (2012)
Picea abies	Demography	47	22	38	0.0088	0.013	−0.92	23	2, 19		Heuertz et al. (2006)
Picea abies	Transcription regulation	23	5	17	0.0019	0.0033	−1.51		0, 5		Namroud et al. (2010)
Picea abies	Adaptation	546	290	384						3.8	Scalfi et al. (2014)
Picea abies	Adaptation	860	285	288						1.7	Di Pierro et al. (2016)
Picea abies	Adaptation	303	232	445						6.51	Chen et al. (2012b)
Picea obovata	Phenology		9	357	0.002		−0.5	11.1		7.5	Chen et al. (2014)
Picea glauca	Adaptation	158	345	534							Namroud et al. (2008)
Picea glauca	Diversity	48	105	1443	0.0043	0.0051	−0.4		36, 69	14	Pavy et al. (2012a)
Picea glauca	Climate adaptation	198	7819	11,085						0.41	Hornoy et al. (2015)
Picea glauca	Diversity	142	709	1134						0	Namroud et al. (2012)

(continued)

Table 10.1 (continued)

Species	Research topic	Samples (1 *N* or 2 *N*)	Genes	SNPs	π	θ	Tajima *D*	Percent loci sig *D*	*D* +,−	F_{ST} outliers (%)	Reference
Picea glauca	Transcription regulation	96	5	34	0.0002	0.001	−0.88		0, 5		Namroud et al. (2010)
Picea mariana	Climate adaptation	156	313	583						4	Prunier et al. (2011)
Picea mariana	Climate adaptation	593	47	48						47.92	Prunier et al. (2012)
Picea mariana	Transcription regulation	26	5	19	0.003	0.00075	−0.9		0, 5		Namroud et al. (2010)
Picea rubens	Air pollution	48	36	61						11.5	Bashalkhanov et al. (2013)
Picea sitchensis	Adaptation	24	153				−0.56				Holliday et al. (2010)
Picea likiangensis	Demography		16	256							Li et al. (2009b)
Picea wilsoni	Demography		16								Li et al. (2009b)
Picea purpurea	Demography		16								Li et al. (2009b)
Picea schrenkiana	Demography		16	49							Li et al. (2009b)
Picea morrisonicola	Demography	15	15	37	0.0015	0.0015	0.28	6.6	7, 2		Bodare et al. (2013)
Cryptomeria japonica	Adaptation	186	1026	1031						10	Tsumura et al. (2012)
Cryptomeria japonica	Diversity	4	3744	13,413	0.0045	0.0042					Uchiyama et al. (2012)
Cryptomeria japonica	Adaptation	48	12	29	0.0012	0.0025	−0.2				Fujimoto et al. (2008)

Cryptomeria japonica	Diversity	16	7	14	0.0007	0.0007	0.46		5, 2		Kado et al. (2003)
Cryptomeria japonica	Diversity	8	4	41	0.0021	0.0016	0.44		4, 1		Kado et al. (2006)
Taxodium distichum	Diversity	40	10	403	0.0027	0.0038	−1.01		1, 9		Kusumi et al. (2010)
Taxodium distichum	Diversity	8	4	67	0.0026	0.0042	−0.99		1, 3		Kado et al. (2006)

appear to be nonneutral, most often less than 10%, a result consistent with the neutralist versus adaptive theories of evolution discussed in Chap. 9. In conifers, we might conclude that both are operative, but that the patterning of adaptive variation in conifer populations might be determined by variation in less than 10% of the genes. Knowing what these genes are and the complex adaptive traits they underlie (Chap. 11) is fundamentally important. We can also see that where studies use a small number of candidate genes versus genome-wide scans, the percentage of genes that are potentially nonneutral increases. This can simply be attributed to sampling and may reflect good prediction of candidate genes.

Can anything be inferred about which complex traits might be more highly adapted? Results from common-garden studies and the measurement of complex traits remain the superior approach to addressing this question. Do we see any differences among conifer species that might be associated with different life history strategies? Here F_{ST} outlier tests might yield more nonneutral genes in species with highly fragmented distributions than in those with very large and continuous distributions. Further, species at the leading edge of migrational fronts or those having passed through large bottlenecks might be the ones showing significant values of D that can be attributed to demographic versus adaptive processes. What still seems to be too early to infer with any amount of certainty is whether the same sets of genes (genetic loci) are generally nonneutral across species (parallel evolution) or whether local adaptation is the stronger evolutionary force. We return to a discussion of this debate in Chap. 12.

Detection of Nonneutral Genes in a Few Conifer Species

Pinus taeda

One of the first conifer species where resequencing of genes was conducted to discover SNPs, calculate measures of nucleotide diversity, and perform tests of neutrality was *Pinus taeda* (Table 10.1). The initial studies (2004–2013) were conducted using Sanger sequencing, while later studies used next-generation sequencing (NGS). The first study resequenced 19 wood-forming genes and found fairly low diversity and no genes that departed from neutrality (Brown et al. 2004). González-Martínez et al. (2006) and Ersoz et al. (2010) resequenced a small number of drought-tolerance and disease-resistance-related genes, respectively, in the same population sample and found about 10% of genes departing from neutrality. Thus, at this very early stage of discovery, it might be concluded that 10% of the coding loci were under selection, and the remaining 90% were selectively neutral. Next, Eckert et al. (2013b) conducted a large resequencing study (5773 genes) and performed several different neutrality tests. The D statistic was significant for up to 8% of the genes at a relaxed significance level. Furthermore, it was shown that genes departing from neutrality were also those more often associated with complex traits as determined from quantitative trait loci (QTL) mapping or association studies (Fig. 10.1). The tacit implication of these multiple and rather complicated analyses

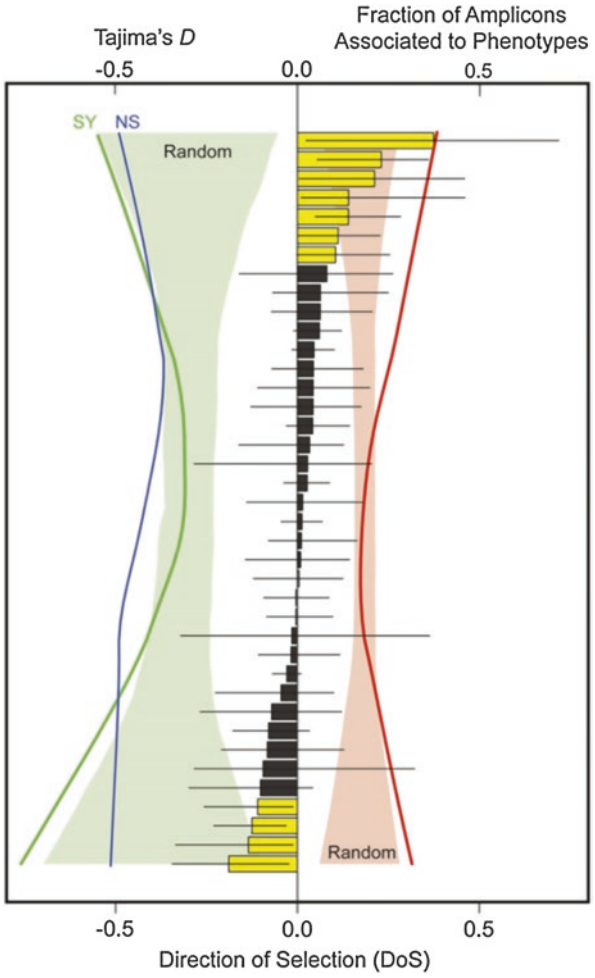


Fig. 10.1 The number of candidate genes for each of 40 functional gene classes (horizontal bars) and their associated average estimate of *D* for 5773 candidate genes in *Pinus taeda*. The distribution of values shows the proportion and functional classes of genes potentially under adaptive evolution. The bottom *x*-axis is for the DoS statistic, while the top *x*-axis is split between Tajima's *D* (left) and the fraction of amplicons associated with at least one phenotype (right). The red-colored area to the right gives the 95% confidence interval for the null distribution, based on 10,000 permutations of amplicons among categories, for the fraction of amplicons associated to at least one phenotype across functional categories. Similarly, the green area on the left does the same for the weighted average value of Tajima's *D* at synonymous (SY) sites. Lines give observed values (green, Tajima's *D* at SY sites; blue, Tajima's *D* at nonsynonymous (NS) sites; red, fraction of amplicons associated to at least one phenotype). All lines were loess smoothed, including those comprising the null distributions. The null distribution for the weighted average of Tajima's *D* at NS sites was similar to that for SY sites, so was omitted for clarity. (From Eckert et al. 2013b)

Table 10.2 Nucleotide polymorphism statistics at PHYP and PHYO loci in *Pinus sylvestris*

	Synonymous		Nonsynonymous		Intron I		Total	
Locus/site	π_s	θ	π_a	θ	π_s	θ	π (SD)	θ (SD)
PHYP								
Kolari	0.0007	0.0008	0	0	0.0037	0.0038	0.0006	0.0006
Lapinjarvi	0.002	0.0024	0	0	0.0021	0.0025	0.0007	0.0008
Puebla de Lillo	0.0013	0.0016	0.0008	0.001	0.0026	0.0025	0.0011	0.0013
All	0.002	0.0031	0.0003	0.0006	0.003	0.004	0.001	0.0016
PHYO								
Kolari	0.0019	0.0021	0	0			0.0004	0.0005
Lapinjarvi	0.0015	0.0018	0.0005	0.0005			0.0007	0.0008
Puebla de Lillo	0	0	0	0			0	0
All	0.0013	0.0025	0.0002	0.0004			0.0004	0.0009

π_s nucleotide diversity from synonymous sites, π_a nucleotide diversity from nonsynonymous sites, θ theta proportion of segregation sites
From García-Gil et al. (2003)

is that genes underlying complex traits, and moreover adaptive traits, are more likely to be nonneutral than genes that have more general functions. This observation is consistent with observed neutrality of allozyme markers which are for the most part from genes of general function. Eckert et al. (2013b) also performed outlier tests for many of these same genes and found only a small proportion of outliers. Exome capture technologies have also been applied to *P. taeda* to discover SNPs (Neves et al. 2013; Lu et al. 2016); however, in neither of these studies were estimates of nucleotide diversity or tests of neutrality performed.

Pinus sylvestris

Some of the very first studies to measure nucleotide diversity in conifer genes were done in *Pinus sylvestris* (Dvornyk et al. 2002; García-Gil et al. 2003) (Table 10.1). These studies found very low diversity in a phenylalanine ammonia-lyase (PAL) gene and two phytochrome genes (PHYP and PHYO, Table 10.2), respectively. Although no tests of neutrality were performed in these studies, the very low nonsynonymous diversity values suggested strong purifying selection. In a demographic study where only silent substitutions in a diverse set of genes were considered, Pyhäjärvi et al. (2007) nevertheless concluded there was little evidence for selection based on nonsignificant *D* values given that the silent sites were within genes and therefore linked to functional site changes that could have been under selection. Somewhat different results were found in a study of genes potentially related to cold tolerance (Wachowiak et al. 2009). Here, 33% of this group of genes had significant *D* values (Table 10.3), most of which were negative, suggesting positive selection at these genes for cold tolerance. A slightly different story emerged in a study of marginal populations in Scotland (Wachowiak et al. 2011b). Here, the Scottish (marginal) populations all had positive *D* values, two of which were significant, whereas all the mainland populations had negative values and none were significant (Table 10.4). These results are consistent with marginal populations arising from a recent population bottleneck. Pyhäjärvi et al. (2007) resequenced several loci that had been used as allozyme markers and found low diversity and none with significant *D* values. This result is consistent with allozymes being considered neutral genetic markers. Finally, a study seeking signals of clinal adaptation found low diversity in a small number of candidate genes and nonsignificant values of *D* (Kujala and Savolainen 2012). In summary, likely because *P. sylvestris* is found across such a large continuous range, finding individual genes coding for polygenic adaptive traits with significant departure from neutrality appears to be difficult due to the counterbalancing effects of gene flow. This situation might change once this genome has been sequenced, and a comprehensive SNP database becomes available for these types of studies.

Table 10.3 Neutrality tests at cold-related genes in *Pinus sylvestris*

Locus	D^a
dhn1	−0.234
dhn2	−0.643
dhn3	−0.254
dhn4	−1.570**
dhn5	−1.001
dhn7	0.043
dhn9	0.505
ef	−1.954***
lea	−0.724
lp3–2	−0.85
abaR	−1.400*
abaH	−0.166
gst1	−1.396*
gst2	−0.463
Average	−0.722**

[a]Tajima's D test (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$)
From Wachowiak et al. (2009)

Table 10.4 Estimates of θ and D at eight loci in Scottish and continental European populations of *Pinus sylvestris*

Groups	θ	D
Scottish		
West	0.0103	0.580[a]
South	0.013	0.107
East	0.0117	0.499*
All	0.0108	0.316
Continental European		
North	0.0095	−0.143
Central	0.0103	−0.359
North+Central	0.0096	−0.316
Spain	0.0098	−0.539
Turkey	0.0055	−0.279
All	0.0093	−0.379

*$P < 0.05$
From Wachowiak et al. (2011b)

Pinus mugo, *P. uncinata*, and *P. uliginosa*

Three species of the *Pinus mugo* complex have also been the subject of neutrality and outlier tests. Wachowiak et al. (2011a) estimated D for 17 genes in *P. mugo* and *P. uliginosa*, both of which had a negative average D value, but only that for *P. mugo* was significant (Table 10.5). In a similar study, this time with 12 genes, Wachowiak et al. (2013) got negative average D values but none were significant. Mosca et al. (2012, 2016) performed outlier tests from a large sample of genes (190, 383, respectively) from SNP genotype data from population samples from the Italian Alps and found 1.6% (Table 10.6) and 2.9% (Table 10.7) of the genes were significant

Table 10.5 Nucleotide diversity and *D* at nuclear loci in the *Pinus mugo* complex

Species	*N*	*L* (bp)	Nucleotide polymorphisms				
			S	π_{total}	π_{silent}	π_{nonsyn}	*D*
Pinus uliginosa	16.4	9814	157(71)	0.0042	0.0062	0.0018	−0.24
P. mugo	11.1	9909	176 (87)	0.0049	0.0067	0.0022	−0.61*
P. sylvestris	39.5	9880	275 (101)	0.0041	0.0057	0.0022	−1.17*

N average number of sequences analyzed per locus, *L* total length of sequence in base pairs excluding indels, *S* total number of polymorphic sites detected (number of singleton mutations), *D* Tajima's *D* test
[a]Significantly negative values ($P < 0.05$)
From Wachowiak et al. (2011a)

Table 10.6 List of the outliers from the standard neutral model (SNM) in candidate genes in *Pinus mugo*

Gene	Putative protein	S^a	$P_D{}^b$
0_13913	Exocyst subunit EXO70 family protein	4	0.1146
2_8627	Carbon-sulfur lyase	10	0.1131
2_8852	Galactokinase	8	0.0155

[a]Number of SNPs
[b]*P* value for *D*
From Mosca et al. (2012)

outliers. In contrast to the results from the related *P. sylvestris*, there seems to be a bit more evidence for genes under selection in these three species which have patchier distributions.

Pinus pinaster and *P. halepensis*

Mediterranean pine species, primarily *Pinus pinaster* and *P. halepensis*, have been studied rather extensively in search of genes underlying adaptation to drought. In the first reported study, Pot et al. (2005) estimated nucleotide diversity statistics and *D* for eight genes involved in wood formation and, like the similar study in *P. taeda* (Brown et al. 2004), found very low diversity and a positive, but nonsignificant, average value of *D*. Another *P. pinaster* study included 11 drought-tolerance candidate genes but a very large population sample (Eveno et al. 2008). A high proportion (45%) of these 11 candidate genes were significant outliers (Table 10.8), as might be expected if these candidates are truly genes responsible for drought adaptation. Grivet et al. (2009) estimated nucleotide diversity and *D* for ten candidate genes related to drought tolerance in a small number of *P. halepensis* samples from populations in Greece, Israel, and Italy. Average estimates of *D* were positive for the Greek and Israeli populations and negative for the Italian population. However, only one locus in one population (*dnh1,* Israel) was significant. The genes used in this study and the *P. pinaster* study (Eveno et al. 2008) were nearly the same as those first used in a *P. taeda* study (González-Martínez et al. 2006). There were no

Table 10.7 Outlier loci found in *Pinus mugo* with the Bayesian approach with a false discovery rate (FDR) = 0.01

ID	Locus	BLAST	SNP code[a]	All populations		
				F_{ST}	qval	P_{BayeScan}
11	2_4183_01-Pimg-401	Calcium-sodium antiporter (CAX11) NM_100767	SY	0.051	0.005	0.9838
144	2_9542_01-Pimg-254	1 -Aminocyclopropane-1-carboxylate synthase NM_124548	NS	0.049	0.002	0.9948
187	CL1029Contigl_01-Pimg-229	Imbibition protein NM_115612	NC	0.06	0	0.9996
202	CL4147Contig1_01-Pimg-175	Hexokinase (HXK1) NM_119057	NA	0.047	0.003	0.9906
300	CL1455Contig1 _06-Pimg-70	Outer arm dynein light chain 1 protein NM_106473	NC	0.047	0.001	0.9968
330	2_6317_01-Pimg-667	Heat shock protein 101 NM_106091	NA	0.059	0	1
332	2_6313_01-Pimg-172	Unknown protein NM_118832	NA	0.053	0.001	0.9976
341	CL4257Contig1_01-Pimg-345	Unknown protein NM_001084349	SY	0.048	0.004	0.9874
363	0_3483_01-Pimg-275	Glucosidase NM_125779	NA	0.055	0	1
437	2_5064_01-Pimg-410	MAP kinase (MPK16) NM_121906	NA	0.053	0	0.999
528	0_1116_02-Pimg-234	Alanyl-TRNA Synthetase NM_103905	NC	0.046	0.006	0.98
604	UMN_1263_01-Pimg-94	NA	NS			
649	CL3162Contig1 _02-Pimg-122	Rab GTPase homologue A4a NM_125925	SY	0.056	0.002	0.995

[a]The SNP codes are the following: *NA* no annotation, NC noncoding, SY synonymous, NS nonsynonymous.
From Mosca et al. (2016)

Table 10.8 Outliers detected at the SNP level for SNPs having a heterozygosity greater than 0.1 across ten contrasted *Pinus pinaster* populations, using the Fdist2, FstSNP, or BayesFst methods

			Fdist2	F_{ST}SNP		Bayes F_{ST}		
SNP	Gene region/Type of mutation[a]	F_{ST}	*P* value[b]	*P* value[b]	*q* value[c]	*P* value[b]	α_i[d]	SD[c]
PR-AGP4 SI82	Intron/—	0.383	0.08			0.05	0.95	0.5
S331	Intron/—	0.398	0.06			0.05	0.98	0.49
S472	Intron/—	0.382	0.08			0.08	0.62	0.5
S525	Intron/—	0.55	0.02	0.07	0.51	0.08	0.74	0.62
S666	Intron/—	0.341				0.09	0.95	0.55
S762	Intron/—	0.339	0.09			0.08	0.73	0.52
S1177	Exon/ns	0.333	0.09					
S1181	Exon/s	0.35	0.08					
S1409	3' UTR/–	0.365	0.08			0.04	0.97	0.52
S1480	3'UTR/–	0.394	0.06					
S1488	3'UTR/–	0.391	0.08					
GRP3 S1 10	Exon/ns	0.285	0.08					
S164	Exon/s	0.347	0.06			0.04	1.21	0.6
S633	3' UTR/–	0.277	0.1					
erd3 S42	Exon/ns	0.587	0.04	0.06	0.49	0.03	1.55	0.71
CCoAOMT 1832	Exon/ns	−0.034	0.01	0.01	0.33			
S980	Exon/ns	−0.031	0.02	0.02	0.33			
dhn-2 S202	Exon/ns	−0.047	0.04					
S332	Intron/–	−0.056	0.07					
S406	Exon/s	−0.045	0.08	0.05	0.49			
lp3-l S176	3' UTR/–	0.174		0.04	0.49			

(continued)

Table 10.8 (continued)

SNP	Gene region/Type of mutation[a]	F_{ST}	Fdist2	F_{ST}SNP		Bayes F_{ST}		
			P value[b]	*P* value[b]	*q* value[c]	*P* value[b]	α_i[d]	SD[c]
dhn-1 S310	Exon/s	0.063		0.06	0.49			
S725	Exon/ns	−0.043	0.02	0.01	0.33			

[a]s: synonymous polymorphism, ns: nonsynonymous polymorphism
[b]Significance level (*P* value) for detected outliers (in bold for $P < 0.05$)
[c]*q* values are the positive FDR analogues of *P* values. They were obtained for each outlier gene, by considering 11 independent hypotheses corresponding to the 11 candidate genes, and for each SNP, by considering all 94 hypotheses corresponding to the 94 polymorphisms tested
[d]Mode of the posterior distribution of α_i
[e]Standard deviation of the posterior distribution of α_i
From Eveno et al. (2008)

common genes in the three studies that had significant *D* values. Grivet et al. (2010) did an additional and more in-depth study of genes related to drought tolerance in both *P. pinaster* and *P. halepensis.* In this study, they estimated nucleotide diversity and a suite of neutrality tests, including *D*, from populations throughout the Mediterranean. With one exception, all measures of *D* were nonsignificant, and there was little similarity in the sign (+/−) or magnitude of the estimate of *D* between the two species (Table 10.9). In a final study, a search for genes under selection and adapted to drought was taken in a phylogenetic context using sequence data from six Mediterranean pine species Grivet et al. 2013). This study revealed two such genes, one defense- and one stress-related gene. In summary, these early studies have begun to identify individual genes under selection in response to drought stress, but like the work in many conifers, these studies suffer from an inability to conduct a genome-wide search.

Pinus radiata

Despite its huge economic importance in the Southern Hemisphere and ecological value in its native range in California and Mexico, there have been only two studies to date in *Pinus radiata.* As with *P. pinaster,* Pot et al. (2005) estimated nucleotide diversity statistics and *D* for eight genes involved in wood formation and found

Table 10.9 Estimates of *D* in *Pinus pinaster* and *P. halepensis* for phylogeographical groups defined with chloroplast microsatellites

Amplicon	*P. pinaster*	*P. halepensis*
Ip31-Pt		
a	2.172	1.689
b	1.834	0.535
Ip33-Pp	1.133	−0.629
dhn2-Pp		
a	0.269	nps
b	0.186	nps
dhn2-Ps		
a	0.631	na
b	−0.126	−0.595
dhnS-Ps	1.96	na
4cl-Pt		
a	0.422	0.752
b	−1.041	2.386
c	0.214	1.701
d	na	2.324

a through *d* indicate different amplicons from the same gene
From Grivet et al. (2010)

almost no diversity and no loci departing from neutrality. In a large study of California mainland populations, Dillon et al. (2013) genotyped 447 trees at 38 loci and found 6.7% of the loci to be outliers which potentially underlie patterns of adaptation to different environments. There has been no subsequent research since this publication.

Pinus contorta

Eckert et al. (2012) estimated nucleotide diversity and outlier analysis in two subspecies of *Pinus contorta* sampled over a 5 km gradient in northern coastal California. *Pinus contorta* ssp. *contorta* is found along the coast and *P. contorta* ssp. *bolanderi* is found just inland. A small number (24) of candidate genes were resequenced for 92 trees across the short sampling area and outlier tests were performed. All tests were insignificant except for two genes, an aluminum ABC-transporter gene and an inorganic phosphate-transporter gene. It seems quite likely that the difference in allele frequencies at the aluminum ABC-transporter gene was in response to higher concentrations of aluminum in the podzolic soils where the *bolanderi* subspecies is found.

Pinus massoniana and *P. hwangshanensis*

A single comprehensive study has been done in these species from China (Zhou et al. 2014). Many samples (60 and 44) and a modest number of genes (25) and SNPs (321) and 480) were used in *P. massoniana* and *P. hwangshanensis*, respectively. These data were used to estimate nucleotide diversity, *D,* and perform outlier tests (Table 10.1). Measures of diversity were higher than for most other conifer studies. The 25 candidate genes chosen were thought to relate to climate adaptation. Average estimates of *D* were negative for both species and a small number of loci had significant *D* values. However, there were no loci with significant *D* values in common between species. Five of the candidate genes had SNPs that were outliers in a test that combined data from both species (Fig. 10.2). In summary, based on the relatively large number of samples, the selection of candidate genes, the comprehensive analyses, and the comparative approach, this study begins to truly identify nonneutral genes and those potentially underlying climate adaptation.

Pinus lambertiana and Other Subgenus *Strobus* Species

Pines of the subgenus *Strobus* include several species which are found at high elevations and may be more vulnerable to changing climates. Consequently, they have been the focus of several studies, including a few landscape genomic studies (Chap. 12). In an early study to identify genes potentially controlling resistance to white pine blister rust, Jermstad et al. (2006) sequenced a CC-NBS-LRR gene

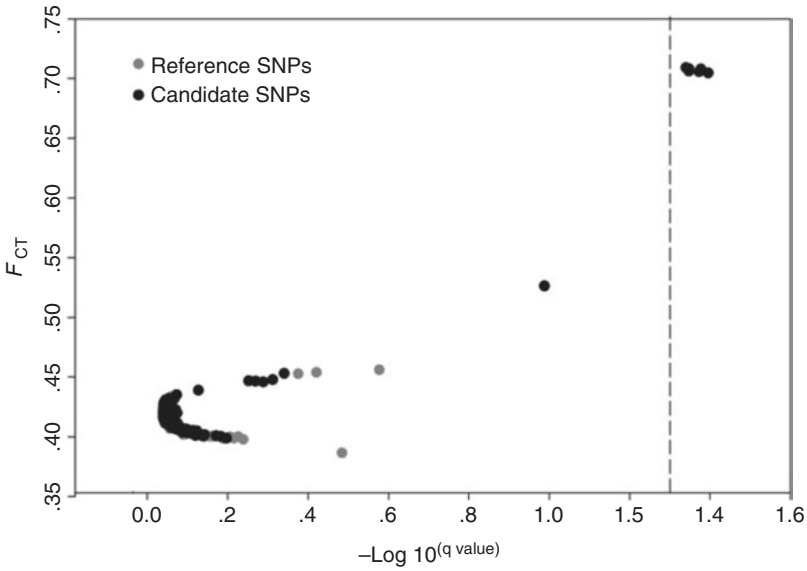


Fig. 10.2 F_{ST} outliers in *Pinus massoniana* and *P. hwangshanensis* detected using BAYESCAN. (From Zhou et al. 2014)

(see Chap. 5) in a sample of 24 diverse *Pinus lambertiana* trees and found absolutely no variation. Like the results with the phytochrome genes in *P. sylvestris*, they concluded that this disease-resistance-related gene must be under strong purifying selection. Eckert et al. (2013a) conducted a large resequencing study for 11 species of subgenus *Strobus* pines (Table 10.1). Estimates of nucleotide diversity and *D* were made for 163 genes in all 11 species. Estimates of diversity were generally quite low, and average *D* was always negative but nonsignificant. Eckert et al. (2013a) estimated an additional parameter, α, the fraction of nonsynonymous substitutions. This summary statistic provides an estimate of the proportion of the genome under selection (adaptive evolution). Surprisingly, they found that none of the 11 estimates of α were significantly different from zero (Fig. 10.3). However, they argued that the gene set sampled was likely highly biased toward highly conserved (and possibly general function) genes and that surveys with larger sets of genes or candidate genes for adaptive evolution should reveal significant estimates of α. Here again is the argument for reference genome sequencing and the need to develop complete gene catalogs (see Chap. 3 for discussion of the sequencing of the *P. lambertiana* genome).

Mosca et al. (2012, 2016) conducted two studies searching for genes under selection in *Pinus cembra* in the European Alps. In the first, nucleotide diversity was estimated for 171 candidate genes, and several tests of neutrality were performed (Table 10.10). They found a small number (1.8%) of the candidate genes were outliers based on several tests. In the second study, Mosca et al. (2016) did SNP genotyping in 673 trees from 18 populations across elevational gradients for 265 genes and

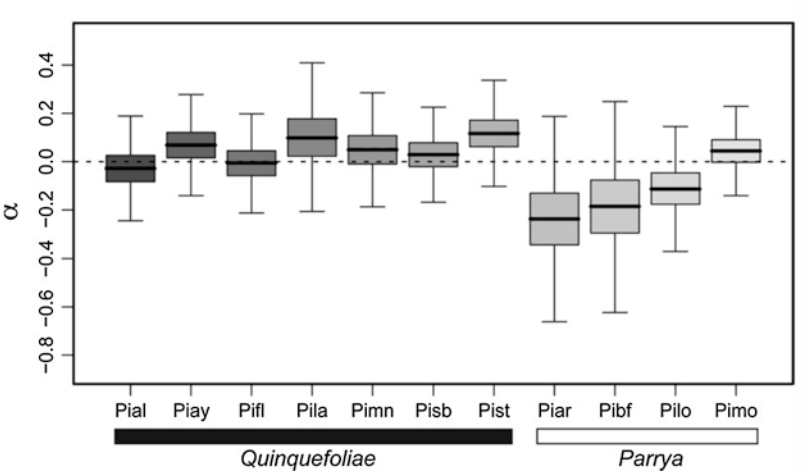


Fig. 10.3 The distribution of α for several subgenus *Strobus* species: Pial, *Pinus albicaulis*; Piay, *P. ayacahuite*; Pifl, *P. flexilis*; Pila, *P. lambertiana*; Pimn, *P. monticola*; Pisb, *P. strobiformis*; Pist, *P. strobus*; Piar, *P. aristata*; Pibf, *P. balfouriana*; Pilo, *P. longaeva*; Pimo, *P. monophylla*. (From Eckert et al. 2013a)

Table 10.10 List of the outliers from the standard neutral model in *Pinus cembra*

Gene	Putative protein	S[a]	P_D[b]
0_18619	Protein kinase family protein	6	0.0207
0_2775	spx domain-containing protein	2	0.118
0_8111	6-phosphogluconate dehydrogenase	3	0.0798
2_1528	Reduced epidermal fluorescence 4	3	0.0491
2_6731	E3 ubiquitin complex protein	3	0.0451
CL1659Contig1	Chloride channel-like protein	4	0.0786
CL1661Contig1	Acetyl-CoA carboxylase 2	3	0.0453

[a]Number of SNPs per locus
[b]*P* value for *D*
From Mosca et al. (2012)

found 4.9% of these were outliers. Therefore, in contrast to the results of Eckert et al. (2013a) with many species of subgenus *Strobus* where no outliers were found, these results in *P. cembra* may have been due to the choice of candidate genes or the deeper sampling across elevational gradients.

Pseudotsuga menziesii

Pseudotsuga menziesii is found across an environmentally diverse landscape, where genes under selection might more easily be identified. Two nucleotide diversity studies have been done in the coastal variety, *P. menziesii* var. *menziesii*

Table 10.11 A list of candidate genes in *Pseudotsuga menziesii* putatively affected by positive natural selection

Locus	Gene product
Pm_CL908Contig1	GRAM-containing/ABA-responsive protein
ES420171.1	Cold-regulated plasma membrane protein
ES420250.1	Dehydrin-like protein
CN634517.1	Luminal-binding protein
Polymorphism-to-divergence	
Pm_CL61Contigl	Cyclosporin A-binding protein
Pm_CL908Contigl	GRAM-containing/ABA-responsive protein
CN638556.1	Transcription regulation protein
Synonymous-to-nonsynonymous divergence	
Pm_CL922Contigl	Thaumatin-like protein
CN634677.1	LRR receptor-like protein kinase

From Eckert et al. (2009)

(Krutovsky and Neale 2005; Eckert et al. 2009). Krutovsky and Neale (2005) resequenced 15 cold-hardiness-related and 3 wood-quality-related genes in a sample of 24 trees and found two genes (11%) had significant *D* values (one positive, one negative). In a much larger study, again focusing on cold-hardiness-related candidate genes, Eckert et al. (2009) genotyped 24 trees for SNPs in 121 candidate genes and based on a variety of tests found 6.6% of these to be outliers (Table 10.11). These studies identify genes under selection that may control a complex adaptive trait such as cold hardiness. In Chaps. 11 and 12, we will see how QTL and association studies and landscape genomics studies, respectively, provide further insight into discovery of genes underlying complex adaptive traits.

Larix Species

A small number of studies have been done in this interesting taxonomic group, a genus of conifer having the rare deciduous form. The first resequencing study was done by Khatab et al. (2008) in five species and three varieties of *Larix* for just two wood-formation genes (4CL and C3H). A very low level of nonsynonymous substitutions was found for these genes, very similar to results from *Pinus* and *Picea* for these same genes. No neutrality tests were performed, but a general picture has emerged that genes in the phenylpropanoid pathway may be under strong purifying selection in conifers. Mosca et al. (2012) resequenced 61 genes in *Larix decidua* and got a positive average value of *D*, which is in contrast with most of all other studies where the average *D* is negative. They also performed a compound neutrality test and found no outliers. Semerikov et al. (2013) did a study in *Larix sibirica* and *L. occidentalis* by resequencing eight of the same cold-hardiness-related candidate genes used in the *Pseudotsuga menziesii* study of

Eckert et al. (2009). The average estimate of *D* was negative, and three loci with negative values were significant. These results in *Larix* seem to corroborate the results from *P. menziesii* that indicate candidate genes for cold hardiness are under positive selection.

Abies Species

There has been only one study searching for genes under selection in this important genus of trees that are often found in montane environments. Mosca et al. (2012) resequenced 70 candidate genes in *Abies alba* and got an average negative estimate of *D*, but none of the genes were detected as outliers. Clearly, more studies are needed in species where the effects of climate change could be quite important.

Picea Species

Researchers in Canada, Sweden, and Italy have conducted some of the most comprehensive work to discover genes under selection in species of the genus *Picea.* The first such study was done by Heuertz et al. (2006) in *Picea abies*. They resequenced 22 genes in 47 trees from Western Europe and found most genes had a negative *D* value of which 23% were significant. However, the authors attributed these results to recent population expansion more so than to purifying selection. Chen et al. (2012b) tested for outliers among 18 photoperiod-related candidate genes, also across Western European populations, and found 21% to be significant. It is not clear if any of the significant *D* estimates or outliers were the same between these two studies. Subsequently, Chen et al. (2014) tested for departure from neutrality and outliers in nine candidate genes in another species of *Picea*, *P. obovata*, and found four nonneutral genes in several tests (Table 10.12). Two large studies of *P. abies* in the Italian Alps have been done. Scalfi et al. (2014) compared outliers detected on a micro-geographical scale in Italy versus those from a macro-geographical scale across Western Europe and found a small number (3.8%) of outliers in just the micro-geographical sample (Table 10.13). In a larger study with 860 trees and a similar set of SNPs, Di Pierro et al. (2016) found 1.7% to be outliers, but it appears none were in common with the Scalfi et al. (2014) study. Many of the above studies also conducted landscape genomic analyses to discover the relationship between genes under selection and environmental factors. We will return to many of these studies in Chap. 12.

Four *Picea* species are commonly found in Canada (*P. glauca*, *P. mariana*, *P. rubens*, and *P. sitchensis*), and neutrality and outlier tests have been conducted in all of them. The first study found up to 14% of the 534 SNPs from 345 genes in 158 trees tested were outliers (Namroud et al. 2008) (Fig. 10.4). The genes tested were basically a random sample from the genome, so this gives a pretty good estimate of the portion of genes in the genome potentially under selection. Namroud et al. (2010) then did a study that targeted just five genes involved in transcriptional

Table 10.12 SNPs in *Picea obovata* candidate genes that are significant at the empirical threshold of 5% with different analyses of allele frequencies in the six Yenisei populations

Gene	SNP	#SNP linked	Methods[a]	Mutation[b]
Po*CCA1*	570	2	BEV, F_{ST}	Intron
	582	2	F_{ST}	Intron
	1016	2	LR, BEV	Intron
	2803	5	LR	Intron
	4113	1	LR, BEV	NS: Ser/Leu
Po*FTL2*	1567	1	F_{ST}	Promoter
Po*Gl*	*F2*_605	5	BEV, F_{ST}	NS: His/Tyr (*F*4_138, *F*4_810)
	*F4*_134	1	BEV	NS: Gly/Arg
	*F4*_638	2	LR, BEV, F_{ST}	NS: Phe/Leu
	*F6*_8	1	BEV	Intron
	*F6*_39	1	BEV	Intron
	*F6*_76	1	F_{ST}	Intron
Po*MFTL1*	273	1	BEV, F_{ST}	Intron
	322	3	LR, BEV	Intron
	416	12	LR, BEV	Intron
	1487	1	BEV	Intron
	2888	1	F_{ST}	Intron
Po*PHYN*	583	1	F_{ST}	Intron
	2695	5	LR	SYN (2478)
Po*PHYP*	562	1	LR, BEV	SYN
	726	1	LR, BEV	NS: Leu/Pro
	802	1	BEV	SYN
Po*PRR7*	4758	1	BEV	Syn
	5687	1	BEV, F_{ST}	NS: Asn/Asp
	6627	1	BEV, F_{ST}	Intron
	6656	1	BEV, F_{ST}	Intron
	6717	1	LR, BEV	Intron

[a]*LR* linear regression empirical significance, F_{ST} BayeScan empirical significance, *BEV* Bayenv empirical significance

[b]Mutation type of the SNP: *NS* nonsynonymous, *SYN* synonymous, positions of all nonsynonymous SNPs linked to are also indicated in the parentheses

From Chen et al. (2014)

regulation in *P. glauca*, *P. mariana*, and *P. abies* and found that all had negative values of *D* with just one gene each significant in *P. glauca* and *P. mariana* but four were significant in *P. abies* (Table 10.14). These results imply that genes controlling gene expression may be more likely under selection, but at the same time, confounding signatures of recent population expansion were found for all three species. In a study focused on estimating linkage disequilibrium (LD) in *P. glauca*, Pavy et al. (2012a) found that a high percentage (45%) of the genes with either high or low LD had significant and negative *D* values; however, they also attributed these results to demographic rather than adaptive processes. The most recent study in *P. glauca* employed NGS of 13,500 genes and estimation of the A/S ratio

Table 10.13 Outlier detection in *Picea abies* in the Alps using BayeScan

SNP	Putative function	F_{ST}
0_10267_01–274	R2R3-MYB transcription factor MYB8	0.0235
0_8642_01–166	translation elongation factor EF-G	0.1298
0_9922_01–345	UBX domain-containing protein	0.1309
2_10483_01–340	haloacid dehalogenase-like	0.2336
2_5073_01–321	Unknown	0.0183
2_8491_01–519	acyl-CoA thioesterase, putative	0.1304
CL813Contig1_03–235	sucrose synthase	0.1585
CL866Contig1_01–360	acetyltransferase component of pyruvate dehydrogenase	0.0183
1_3086_01–101	NA	0.1277
0_12021_01–161	ovule receptor-like kinase 28	0.0757
CL4578Contig1_02–154	NA	0.0545
UMN_4091_02–458	F-box family protein	0.0544

From Scalfi et al. (2014)

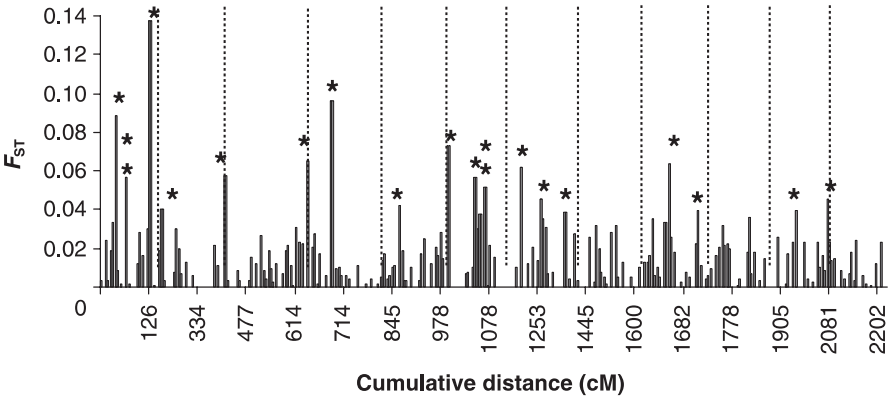


Fig. 10.4 Distribution of empirical F_{ST} values over the 12 linkage groups of *Picea glauca*. The vertical dotted lines indicate approximate boundaries between consecutive linkage groups. The identification of 20 outlier SNPs is indicated by asterisks above vertical solid lines. (From Namroud et al. 2008)

(nonsynonymous/synonymous) (Pavy et al. 2013). An exhaustive analysis of differences in A/S ratio and functional gene class was conducted (Fig. 10.5) and revealed a small number of functional classes with very high A/S ratios, gene classes that may more likely be under selection.

Two comprehensive studies to detect outliers in *P. mariana* have been done (Prunier et al. 2011, 2012). In the first of these, 583 SNPs from 313 genes were genotyped in 156 trees, of which 8.0% were found to be outliers. This value falls in the ±10% range that has been found with many other species (Table 10.1). The annotation of the outlier genes suggests these may be involved in adaptation to temperature and precipitation (Table 10.15). In a follow-on study, Prunier et al. (2012)

Table 10.14 Tajima's *D* values for five genes involved in transcriptional regulation in three *Picea* species

Species/Gene	Tajima's *D*
P. glauca	
KN1	−1.05
KN2	−0.69
KN3	−1.15*
KN4	−0.7
HB-3	−0.83
P. mariana	
KN1	−1.18
KN2	−0.63
KN3	−0.44
KN4	−0.95**
HB-3	−1.32
P. abies	
KN1	−1.64***
KN2	−1.14**
KN3	−2.28***
KN4	−0.4
HB-3	−2.08***

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$
From Namroud et al. (2010)
Note to publisher: Just the species and Tajima's D columns from Namroud et al. (2010) Table 10.4.

tested 47 of the genes identified as candidates in the previous study to a much larger sample (593) of *P. mariana* from across its natural range in Canada. This time nearly half (49.9%) of the candidate genes were outliers, suggesting the role of these genes in climate adaptation.

Two studies have been done using neutrality tests and outlier tests in *P. sitchensis* and *P. rubens*, respectively. Holliday et al. (2010) resequenced 153 genes in *P. sitchensis* and calculated an average *D* value, based only on silent sites, to be −0.56, thus arguing for recent population expansion following postglacial migration. In *P. rubens*, Bashalkhanov et al. (2013) found that 11.5% of 36 air-pollution candidate genes were significant in a population sample from eastern North America.

Finally, a couple of resequencing studies have been done with *Picea* species in Asia with the objective of inferring demographic history of populations. A small number of genes (12–16) were resequenced in *P. likiangensis*, *P. wilsoni*, *P. purpurea*, and *P. schrenkiana* where low levels of diversity and negative values of *D* were found, indicating recent population expansion from bottlenecks (Li et al. 2009b). In a study of *P. morrisonicola* in Taiwan (Bodare et al. 2013), 15 genes were resequenced and a positive average value of *D* was found, suggesting a population contraction in this species with a vulnerable International Union for Conservation of Nature (IUCN) rating.

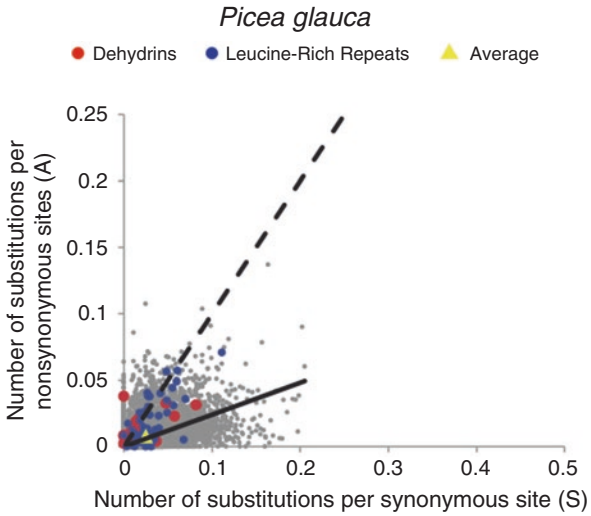


Fig. 10.5 Number of substitutions per nonsynonymous site (A) and number of substitutions per synonymous site (S) for *Pinus glauca* (13,498 genes). The yellow triangles represent the averages. The dashed line represents the null expectation if substitutions were randomly distributed (A 1/4 S). The solid line is the slope averaged over all data (average A for all coding sequences divided by average S for all coding sequences). Data points for genes from two families involved in response to biotic stress (LRR family) and abiotic stress (dehydrins) are colored. (From Pavy et al. 2013)

In summary, resequencing studies to find genes departing from neutrality and outliers are quite advanced in *Picea.* In the very widespread species, such as *P. glauca* and *P. abies*, large candidate-gene studies have already revealed many genes potentially underlying climate adaptation. Now with the availability of reference genome sequences for these species (see Chap. 3), very large studies that include nearly all coding genes should be possible. Instead of inferring candidates from function in non-conifer model species or from gene expression data, candidates might be identified from genome scans. Once candidates are identified in this way, more comprehensive studies including QTL or association mapping (Chap. 11) or landscape genomic studies (Chap. 12) can be conducted in parallel with candidates to improve accuracy in identifying genes underlying climate adaptation.

Cryptomeria japonica and *Taxodium distichum*

There are very few studies aimed at finding nonneutral genes in conifer species other than those in the Pinaceae. However, the notable exception to this is the work on *Cryptomeria japonica* from Japan and to a much lesser extent the related *Taxodium distichum* from the southeastern United States. The first published study in any conifer using gene resequencing and estimating nucleotide diversity and departures from neutrality was done in *C. japonica* (Kado et al. 2003). Here, seven

Table 10.15 List of outlier gene SNPs identified in *Picea mariana* and descriptive information

Climatic partition	Outlier SNP	LG	Gene family	Putative biological function	F_{ST}	SNP location	Amino acid change
Temperature	03870a	3	C2H2 zinc finger	PH	0.049	Exon	Val/Ala
	04312b	6	C2H2 zinc finger	GR, RE, AS	0.07	Exon	Asn/Asn
	04554n	7	R2-R3 MYB	GR, WF	0.044	Intron	–
	04671 m	7	Zinc-binding family	WF	0.046	Intron	–
	0581le	10	C3HC4 RING finger	PH	0.06	3'-UTR	–
	08080a	7	R2-R3 MYB	GR, WF	0.045	Exon	Leu/Leu
	08398a	5	AP2	GR, RE	0.055	3'-UTR	–
	09027f	3	WD-40 repeat family	GR, PH	0.045	Intron	–
	09455a	8	PHD finger family	PH	0.055	3'-UTR	–
	09573a	11	MBF	PH, AS	0.046	3'-UTR	–
	09644 m	5	Peroxidase	WF, AS	0.058	3'-UTR	–
	10,254 m	4	Glycosyl-hydrolase	Unknown	0.056	Exon	Ser/Ser
	13,855 m	9	R2-R3 MYB	GR, WF	0.046	Exon	Asn/Asp
	13855n	9	R2-R3 MYB	GR, WF	0.078	Exon	Iso/Thr
Precipitation	02991 m	10	C3HC4 RING finger	PH	0.049	3'-UTR	–
	03650 g	9	B-box zinc finger	Unknown	0.048	Intron	–
	03870a	3	C2H2 zinc finger	PH	0.043	Exon	Val/Ala
	04073 m	4	LIM	GR, WF, BS	0.044	Exon	Ser/Ser
	04312b	6	C2H2 zinc finger	GR, RE, AS	0.068	Exon	Asn/Asn
	04671 m	7	Zinc-binding family	WF	0.041	Intron	–
	06340a	12	Mov34 family	Unknown	0.063	Exon	Pro/Pro
	07074e	6	ATBZIP2/ GBF5	AS	0.056	Exon	Gln/Gln
	07681e	1	C3HC4 RING finger	PH	0.055	Intron	–
	09406f	5	LEA	AS	0.056	Intron	–
	09870a	6	Ubiquitin	Unknown	0.043	Intron	–
	09889a	3	NAC	GR, WF	0.05	3'-UTR	–
	13572n	7	HD-Zip	GR	0.048	Exon	Asn/Ser
	13855n	9	R2-R3 MYB	GR, WF	0.063	Exon	Iso/Thr

(continued)

Table 10.15 (continued)

Climatic partition	Outlier SNP	LG	Gene family	Putative biological function	F_{ST}	SNP location	Amino acid change
	15328f	12	R2-R3 MYB	GR, WF	0.044	Exon	Asn/Lys
	16364e	1	C2H2 zinc finger	GR, RE, AS	0.063	Exon	Asn/Arg

Abbreviations: *LG* linkage group, *PH* phenology, *GR* growth, *RE* reproduction, *WF* formation, *AS* abiotic stress, *BS* biotic stress, *UTR* untranslated region
From Prunier et al. (2011)

genes were resequenced which revealed very low estimates of nucleotide diversity and a positive average *D* value. Other neutrality tests were performed that suggested selection might be acting on some loci. This study was followed by a similar comparative study using *C. japonica* and *T. distichum. C. japonica* again had a positive value of *D*, whereas for *T. distichum* it was negative, an early indication that selection might be acting differently on these related species. Fujimoto et al. (2008) resequenced 12 genes in *C. japonica* and like the study of Kado et al. (2003) found an average value of *D* being negative. The final study using a small number of candidate genes (ten) was done in *T. distichum* where nine of the ten genes had a negative *D* value, consistent with the earlier study of Kado et al. (2006). The most recent studies on *C. japonica* have used much larger numbers of genes. Tsumura et al. (2012) conducted outlier tests on 1026 genes from *C. japonica* and found that 10% were significant, an estimate very much like that seen in species of Pinaceae. Finally, Uchiyama et al. (2012) resequenced 3744 genes from *C. japonica* and found estimates of nucleotide diversity similar to that found in the Pinaceae, although no tests of neutrality were reported.

Summary

The technical ability to identify genes and genetic markers potentially under selection, and thus having adaptive genetic potential, was enabled by the availability of DNA sequences of genes. This began in the early 2000s with expressed sequence tag (EST) sequencing, followed by resequencing of ESTs in population samples. Molecular population genetic parameters, notably neutrality tests and F_{ST} outliers, could be estimated from gene sequences and inferences made as to whether a gene might be neutral or nonneutral. Studies of this type have been done for 40 or more conifer species and for many hundred genes. Nevertheless, these studies are still quite preliminary as full genome-wide studies have not yet been conducted. Early inferences from these data suggest that on the order of 10% of the protein-coding genes in any one species might be nonneutral. However, it is much too early to conclude if such nonneutral genes are common among species (parallel evolution) or if there are great differences among species and environments (local adaptation). A deeper understanding of the genes responsible for the considerable adaptive genetic variation in conifers will emerge once full genomes are sequenced and resequenced from population samples taken from across heterogeneous environments.

Quantitative Trait Dissection

11

A Short History of Complex Trait Dissection in Conifers

The inheritance of phenotypes (traits) in any organism can generally be classified in either of two ways: (1) those inherited by a single gene and (2) those inherited from multiple genes. Multiple terms have been used to describe these two conditions. For single-gene traits, *qualitative, monogenic*, and *Mendelian* are used, while traits controlled by multiple genes are referred to as *quantitative, polygenic*, or *complex*. In forest genetics, qualitative and quantitative are most often used (White et al. 2007) and will be generally used in this chapter. In conifers, both qualitative and quantitative inheritance of traits are observed, but by far most traits of interest are quantitatively inherited. For example, traits related to yield, wood properties, and abiotic adaptation are just about always highly quantitative. There are far fewer examples of qualitatively inherited traits, one of the most notable being resistance to white pine blister rust in several species of *Pinus* subgenus *Strobus* (Chap. 14).

The study of the inheritance of quantitative traits has classically been accomplished using quantitative genetic theory and statistical methods (Falconer 1960). These methods are the basis of tree breeding theory and application (Zobel and Talbert 1984; White et al. 2007). Genetic parameters such has heritability (h^2) and variance components are used to describe the inheritance of quantitative traits. What cannot be derived from these methods are the specific genes that underlie quantitative traits. Furthermore, the number of genes controlling the quantitative traits and their individual contribution to the total phenotypic or genetic variance can only be very crudely estimated. It was then only natural for geneticists to wonder what specific genes were underlying quantitative traits, most notably the genes responsible for many diseases in humans.

The notion that a quantitative trait could be broken down to its individual gene components goes back to Sax (1923), working in common bean, and later Thoday (1961), working in *Drosophila.* The very general idea was that if a relationship could be established between the genotypes of a Mendelian genetic marker and the phenotypic values of some quantitative trait in a segregating population, then this

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_11

would be evidence that a quantitative trait locus (QTL) exists somewhere proximal to the map position of the genetic marker. We refer the reader to a series of seminal papers by Eric Lander and coauthors for a more detailed description of the basic idea of QTL mapping and complex trait dissection (Lander and Botstein 1989; Lander and Schork 1994; Kruglyak and Lander 1995). Simple explanations can also be found in White et al. (2007). Numerous background papers on QTL mapping in forest trees were written in the early 1990s (i.e., Neale and Williams 1991; Neale et al. 1992, 1994). Later, as the work progressed, several review type papers appeared (Sewell and Neale 2000; Kole 2007; Ritland et al. 2011). The primary motivation for QTL mapping was to enable marker-based breeding in forest trees, similar to what was already well advanced in several agricultural systems (Neale and Williams 1991; Williams and Neale 1992). Although QTL mapping contributed to a deeper understanding of quantitative trait architecture in forest trees, the application to applied tree breeding was rather limited. The reason for this was in the resolution of the mapping of QTLs in conifers and most forest trees. Forest trees are nearly all highly heterozygous, outbreeding, and found in populations with very low linkage disequilibrium (Strauss et al. 1992; Neale and Kremer 2011). Under these conditions, the linkage phase between genetic marker alleles and QTL alleles can differ among trees in a breeding population, greatly complicating and challenging marker-based breeding. The straightforward solution to this challenge was always to simply improve the resolution of the mapping of the QTL. This could be done by greatly increasing the size of pedigreed QTL mapping populations, but the time and expense of this solution limited its application.

QTL mapping studies in humans always had practical limitations, notably the inability to arrange crosses and F_1 family sizes. Human geneticists worked with data combined across many small families and the complications arising from differences in linkage phase among families. What emerged as an alternative in human studies was the *association mapping* approach where genotypic and phenotypic data were obtained from a large number (100 s to 10,000 s) of unrelated individuals. Conifer geneticists recognized quite quickly that this approach would be very amenable to conifers (Neale and Savolainen 2004; Wilcox et al. 2007; Khan and Korban 2012; Thavamanikumar et al. 2013) for a number of reasons including (1) conifers can be found in random mating and unstructured populations (Chap. 9), (2) adequate levels of nucleotide diversity could be found (Chap. 10), (3) low linkage disequilibrium, and (4) precise evaluation of phenotypes from common garden tests (Chap. 8). Beginning around 2005 there was a transition from the QTL approach to the association approach which continues to this day. In the following sections we will summarize the general findings from QTL and association studies in several conifer species.

Pinus taeda

Many studies, using both QTL mapping and association genetics, have been done with *Pinus taeda* (Table 11.1) to discover QTLs for physical and chemical wood properties, growth traits, form traits, resistance to abiotic stress, disease resistance,

Table 11.1 Summary of QTL and association studies conducted in multiple genera

Taxon	Study type	Mapping population (number)	Population (number)	Genetic marker (number)	Trait	References
Pinus taeda	QTL	3Gen (177)		RFLP (146)	Wood specific gravity	Groover et al. (1994)
		3Gen (171)		RFLP (119)	Wood specific gravity	Knott et al. (1997)
		3Gen (47, 172)		RFLP, RAPD, Isozyme (211, 336)	Height and diameter increment	Kaya et al. (1999)
		Self (279)		AFLP	Inbreeding depression	Remington and O'Malley (2000)
		3Gen (172)		RFLP	Wood specific gravity, microfibril angle, % latewood	Sewell et al. (2000)
		3Gen (172)		RFLP	Cell wall chemistry	Sewell et al. (2002)
		3Gen (672, 500, 77)		RFLP, ESTP	Wood specific gravity, microfibril angle, % latewood	Neale et al. (2002)
		3Gen		RFLP, RAPD, Isozyme, ESTP	Wood specific gravity, microfibril angle, % latewood, cell wall chemistry	Brown et al. (2003)
		3Gen (73–91)		SSR (95)	Height increment	Gwaze et al. (2003a)
		F_1 (217)		SNP (409)	Forking	Xiong et al. (2016)
	Association		Clonal, unrelated (480)	SNP (58)	Wood specific gravity, % latewood, microfibril angle	González-Martínez et al. (2007)
			Partial diallele (961)	SNP (46)	Water-use efficiency	González-Martínez et al. (2008)
			Clonal, unrelated (498)	SNP (7216)	Resistance to *Fusarium circinatum*	Quesada et al. (2010)
			Clonal, unrelated (380)	SNP (3938)	Water-use efficiency, height, foliar nitrogen	Cumbie et al. (2011)

(continued)

Table 11.1 (continued)

Taxon	Study type	Mapping population (number)	Population (number)	Genetic marker (number)	Trait	References
			Clonal, unrelated (445)	SNP (356)	Metabolome	Eckert et al. (2012a)
			Clonal diallele	SNP (4854)	Oleoresin flow	Westbrook et al. (2013)
			Clonal, unrelated (400)	SNP (3938)	Gene expression	Palle et al. (2013)
			Clonal, unrelated (498)	SNP (7216)	Gene expression, metabolome, disease resistance, water-use efficiency	Eckert et al. (2013b)
			Clonal, diallele (1706)	SNP (7216)	Height, diameter, volume, disease resistance, straightness, forking, wood specific gravity	Chhatre et al. (2013)
			Clonal, diallele	SNP (4853)	Fusiform rust resistance	Quesada et al. (2014)
			Clonal diallele	SNP (4854)	Resin canal number	Westbrook et al. (2015)
			Clonal, unrelated (384)	SNP (2.8 M)	Leaf area, branch angle, crown width, diameter, height, water-use efficiency, nitrogen, disease resistance	Lu et al. (2017)
Pinus palustris x elliottii	QTL	F1 (235)		RAPD (159)	Aluminum tolerance	Kubisiak et al. (1999)
		BC1 (258)		RAPD (133)	Early growth	Weng et al. (2002)
Pinus radiata	QTL	F1 (93)		RAPD (222)	Seedling growth	Emebiri et al. (1998)
		S1 (198)		RAPD (202)	Viability, inbreeding depression	Kuang et al. (1999a)
		F1 (93)		RAPD, AFLP, SSR (126)	Wood density	Kumar et al. (2000)
		F1 (100, 200)		RFLP, SSR (92)	Wood density, diameter growth	Devey et al. (2004a)
		F1 (7 families)		RFLP, SSR (250)	Dothistroma resistance	Devey et al. (2004b)

		F1 (86)		AFLP, SAMPL, SSR (787)	Pitch canker resistance	Moraga-Suazo et al. (2014)
	Association		OP families (240)	SNP (149)	Wood density, microfibril angle, modulus of elasticity, fiber wall thickness, fiber wall diameter, coarseness, fiber surface area, cell population	Dillon et al. (2010)
			F1, clonal, clonal	SNP (26)	Form traits, branch cluster frequency, stem straightness	Li et al. (2016)
Pinus sylvestris	QTL	BC1 (84)		RAPD (164)	Bud set, frost hardiness	Hurme et al. (2000)
		F1 (94)		AFLP (94, 155)	Height, stem diameter, branch diameter, branch angle, wood density, frost hardiness	Lerceteau et al. (2000)
		F1 (108)		RAPD (186)	Cold acclimation, growth rhythm	Yazdani et al. (2003)
		Open pollinated seed (120)		AFLP	Height, diameter, needle number, length, width and area	Nowicka et al. (2013)
		F1 (1000)		AFLP, SNP	% earlywood, wood density, fiber width, spiral grain angle	Li et al. (2014)
Pinus pinaster	QTL	F2 (126)		RAPD (120)	Height growth	Plomion et al. (1996)
		3Gen (90)		AFLP	Water-use efficiency, ring width	Brendel et al. (2002)
		F1 (90)		AFLP (219)	Growth, wood quality, wood chemistry, kraft pulp production, fiber properties	Pot et al. (2006)
		F2 (200, 477)		AFLP, RAPD, Protein (235, 127, 32)	Water-use efficiency, growth	Marguerit et al. (2014)
		F1 (162)		SSR, SAMPL, SNP (8, 54, 2485)	Water-use efficiency, photosynthesis	de Miguel et al. (2014)

(continued)

Table 11.1 (continued)

Taxon	Study type	Mapping population (number)	Population (number)	Genetic marker (number)	Trait	References
	Association		Unrelated, F1 (160, 162)	SNP (384)	Growth, stem straightness, wood chemistry	Lepoittevin et al. (2012)
			Unrelated (509)	SNP (384)	Serotiny	Budde et al. (2014)
			3Gen (197, 477)	SNP (2498)	Height growth, stem straightness	Bartholomé et al. (2016)
Pinus contorta	Association		Unrelated (98)	SNP (95,000)	Serotiny	Parchman et al. (2012)
Pseudotsuga menziesii	QTL	3Gen-clonal (190, 78)		RFLP (74)	Bud flush	Jermstad et al. (2001a)
		3Gen-clonal (184))		RFLP (74)	Cold-hardiness	Jermstad et al. (2001b)
		3Gen-clonal (460)		RFLP (72)	bud flush, bud set	Jermstad et al. (2003)
		3Gen-clonal (170, 383)		RFLP (74)	Cold-hardiness	Wheeler et al. (2005)
		Full sib families (320)		AFLP (120)	Growth, fiber length and coarseness, wood density, microfibril angle, wood chemistry	Ukrainetz et al. (2008)
	Association		Unrelated families (700)	SNP (384)	Phenology, cold-hardiness	Eckert et al. (2009a)
Picea	QTL	F1 family (260, 500)		SNP (603, 1261)	Phenology, height growth	Pelgas et al. (2011)
		Full sib families (210)		AFLP (?)	Extractive content	Markussen et al. (2005)
					Wood density	Markussen et al. (2004)
		BC1-clonal (283)		SNP (231)	Bud set, height	Prunier et al. (2013)

		F1 family (247)		SNP (686)	Root rot	Lind et al. (2014)
	Association		Open-pollinated families (492)	SNP (944)	Wood density, microfibril angle, modulus of elasticity, tracheid coarseness, cell wall thickness, cell diameter, ring width, % earlywood	Beaulieu et al. (2011)
			Open-pollinated families (1694)	SNP (7434)	Wood density, stiffness, microfibril angle, ring width	Lamara et al. (2016)
			−1355	SNP (83)	Bud set, height	Prunier et al. (2013)
			Unrelated-clonal (410)	SNP (768)	Bud set, cold-hardiness	Holliday et al. (2010)
			Unrelated-clonal (410)	SNP (768)	Bud set, cold-hardiness	Holliday et al. (2012)
Larix	Association				Wood density, height, diameter, volume, carbon content, carbon concentration, lignin content, cellulose content	Wang et al. (2015)
Larix kaempferi x *gmelini*	QTL	F1 (145)		RAPD (581)	Tracheid length, tracheid width, tracheid length to width ratio, wood specific gravity	Guan et al. (2011)
Cryptomeria japonica	QTL	3Gen (73)		RFLP, RAPD (85)	Growth, flowering, rooting	Yoshimaru et al. (1998)
						Goto et al. (2003)
		3Gen (73, 93, 111)		HRM, STS (113, 109, 91)	Male strobilus	Ujino-Ihara et al. (2012)
		Half sib (150)		SNP (1261)	Male sterility (single gene)	Moriguchi et al. (2012)
		Full sib F1 (147)		SNP, EST-SSR, SSR (2560)	Male sterility (4 genes)	Moriguchi et al. (2016)

oleoresins, metabolites, and gene expression. In the initial QTL mapping study, Groover et al. (1994) used a three-generation pedigree, multi-allelic RFLP markers, and a single-marker ANOVA approach to identify five regions of the genome harboring one or more QTLs for wood specific gravity. Because inbred pedigrees are not available in conifers, as they were for most early crop QTL mapping studies, each of the two parent trees could be segregating for two QTL alleles, and these might be different alleles in each parent (up to four alleles). This study showed how these QTL alleles might be detected using multi-allelic RFLP markers and how the effects of each and their interactions could be estimated (Fig. 11.1). Subsequently, Knott et al. (1997) reanalyzed these data using an interval mapping approach they modified from its use in livestock pedigrees. The study established the power of detecting QTLs of large and small effect. A series of studies followed using the interval approach. Kaya et al. (1999) mapped QTLs for height and diameter increment and Sewell et al. (2000, 2002) mapped QTLs for physical wood properties

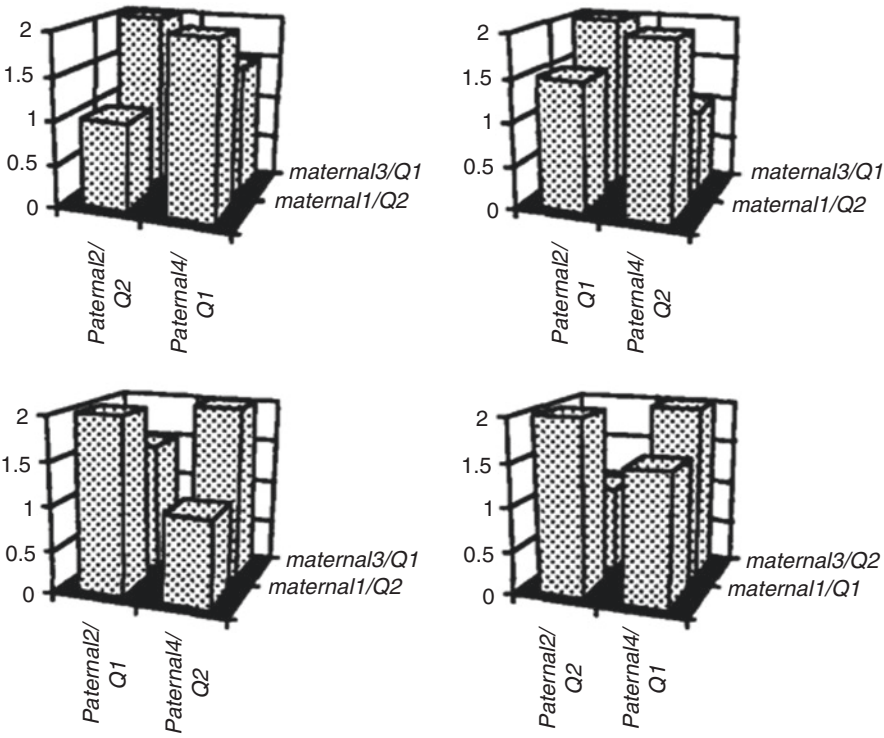


Fig. 11.1 Four possible two-locus phase relationships between a fully informative RFLP marker *(maternal* and *paternal* indicate marker alleles) and a heterozygous QTL where both parents segregate for the same two QTL alleles (*Q1* and *Q2*), assuming complete linkage. *Q1* increases the phenotypic score; *Q2* decreases it. The situation depicted displays over dominance, but the model is applicable regardless of QTL gene action. Note that, in every case, the two classes representing the QTL heterozygous class diagonally oppose each other. (From Groover et al. 1994)

(wood specific gravity, volume percentage latewood, and microfibril angle) and chemical wood properties (lignin, cellulose, and hemicellulose). In the final study in this series, Brown et al. (2003) repeated the wood-quality QTL mapping in a much larger (500) family to verify QTLs detected earlier in smaller (100 to 200) families (Fig. 11.2). This established the need for large families and the power to detect QTLs of small effect size. This body of work established approaches for QTL mapping in conifers and that genome regions harboring QTLs could be repeatedly detected. Still, the resolution of QTL mapping was not precise enough to move directly to marker-assisted breeding (Neale et al. 2002).

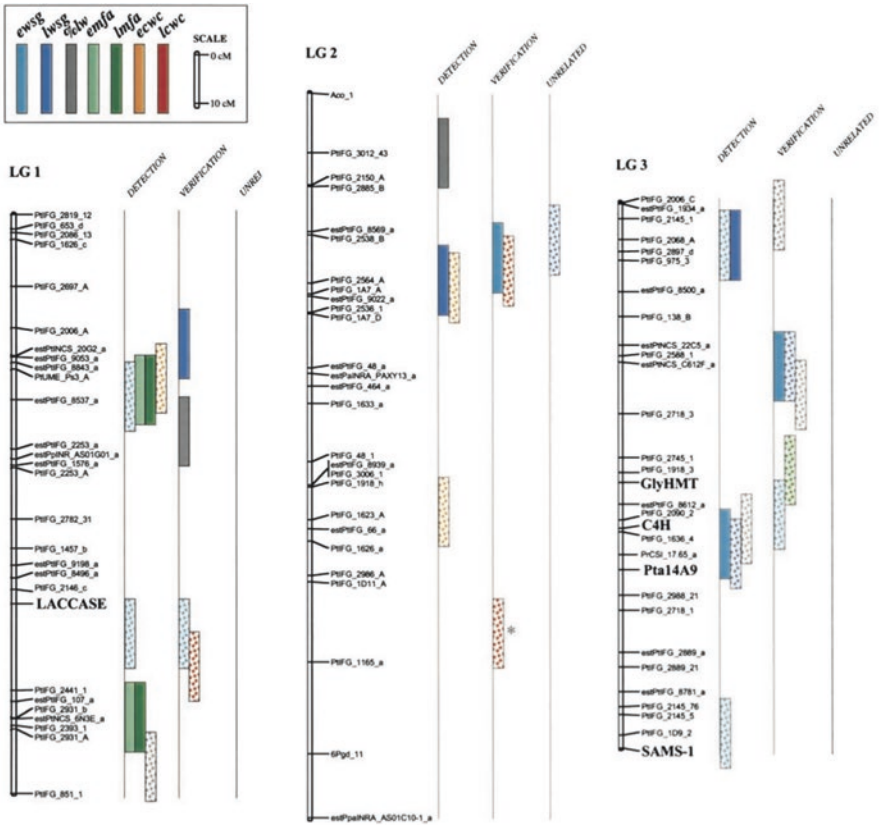


Fig. 11.2 Verification of QTLs influencing wood properties in three populations of *Pinus taeda*. Unique QTLs for earlywood and latewood traits in each pedigree are presented as 15-cM bars on the consensus genetic map at left. An evenly spaced subset of the markers shown on the consensus map was used for QTL analysis in each population. Shaded bars denote QTLs verified by repeated detection across multiple growing seasons; stippled bars represent QTLs observed or measured in only one growing season. Asterisks (*) show QTLs whose position is estimated due to insufficient markers in common to allow more accurate placement on the consensus map. Candidate genes are capitalized and in boldface type. (From Brown et al. 2003)

Beginning around 2005, the *P. taeda* community moved toward using association studies. Simultaneously, EST databases became available (Chap. 5) along with the ability to resequence candidate genes and discover SNPs (Chap. 10). Existing and newly created large, clonally replicated populations were used for genotyping, phenotyping, and testing for marker × trait associations. Several traits were evaluated in a series of studies (Table 11.1). The wood-property trait association study (González-Martínez et al. 2007) included all the same physical and chemical wood-property traits as the earlier QTL studies and found several positively associated candidate genes with genetic map positions close to the QTL positions (e.g., *4cl* and *cad*) (Fig. 11.3). This element of validation was satisfying since it demonstrated that the association approach would work well, and more importantly, greatly improve the mapping resolution relative to the QTL approach. Given the rapid decay of linkage disequilibrium that had already been observed (Brown et al. 2004), the results from association studies suggested that the associated SNP might actually be in the gene controlling the phenotype and not some distance away in a noncoding region. These early results provided optimism that the mapping resolution problem had been solved and there would now be a direct way forward to implementing marker-assisted breeding in conifers (Neale and Savolainen 2004). This first-generation association study was conducted with just 58 SNPs; thus the search of the gene space was limited, and the total amount of variation accounted for was limited. More of the variation would have to be accounted for if this approach was to be applied in a practical breeding program. In a second study to address genes, controlling abiotic stress resistance, González-Martínez et al. (2008) found SNPs in candidate genes associated with water-use efficiency with a panel of just 46 SNPs.

In the early part of the twenty-first century, conifer genome sequencing remained intractable; therefore genome-wide association studies (GWAS), as were being done in organisms having a genome sequence, were not yet feasible in conifers. The alternative path to accounting for more variation controlling a quantitative trait using association studies was simply to use SNPs from a much larger sample of candidate genes. At a significant expense using Sanger sequencing, ~7000 genes were resequenced and ~22,000 SNPs identified (Eckert et al. 2013b). With this SNP database available, it was now possible to construct a much higher density SNP array (7216 SNPs) for association studies. This SNP array was used in a series of association studies (Table 11.1).

Two pathosystems of *P. taeda* (pitch canker and fusiform rust) were used in association studies to discover genes associated with host resistance. SNPs in ten candidate genes for pitch canker resistance were identified (Quesada et al. 2010) (Table 11.2). For fusiform rust a slightly different approach was used. Here, Quesada et al. (2014) were searching for SNPs that mapped closely to single-gene determinants of fusiform rust resistance (*Fr* genes). Several associated SNPs mapped closely to *Fr* loci. The next step in this progression will be to annotate these regions and determine if a gene in a resistance pathway might be present.

The first-generation study by González-Martínez et al. (2008) to discover genes underlying water-use efficiency was repeated (Cumbie et al. 2011) with this large 7216 SNP array and a different association population. Several more candidate

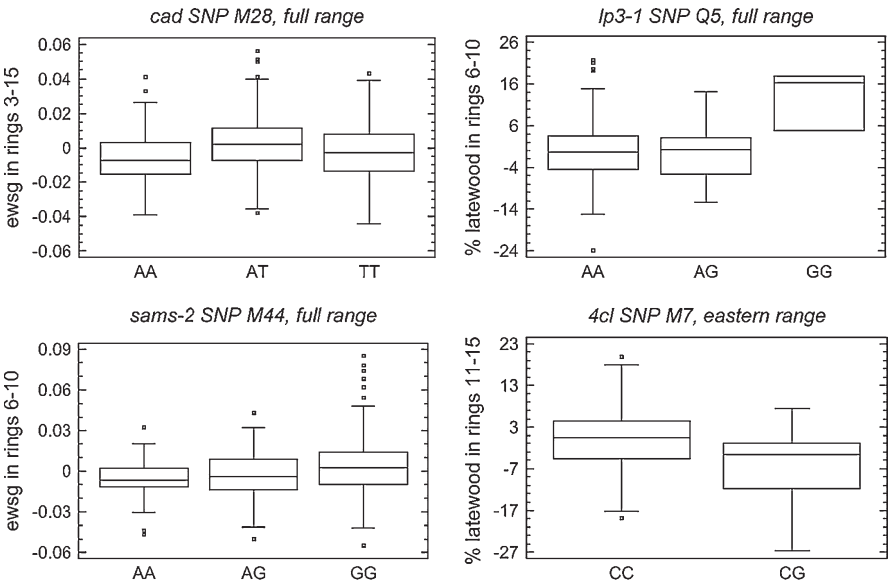


Fig. 11.3 Genotypic effects (box plots) of SNPs in *Pinus taeda* that showed significant genetic association (after correction for multiple testing) with earlywood specific gravity (*cad SNP M28* and *sams-2 SNP M44*) and percentage of latewood (*lp3-1 SNP Q5* and *4cl SNP M7* in the east of the Mississippi Valley range). (From González-Martínez et al. 2007)

genes were associated with carbon isotopes, although none that were positively associated in the initial study were tested again. This points out the need to carefully design experiments so that validation will emerge because association studies are known to have high false positive rates due to multiple testing.

Oleoresins play an important role in conifer defense against insects and pathogens. Westbrook et al. (2013, 2014) conducted two association studies to search for SNP associations for oleoresin flow and resin canal number in a clonal association population. A large number (231) of SNPs in genes in the terpenoid biosynthetic pathway were detected, some of which are shown in Table 11.3. Likewise, for resin canal number, 251 SNPs were positively associated. These authors go on to show how these associations could be used in breeding for increased oleoresin flow and resin canal number using genomic selection.

The final two association studies that used the 7216 SNP array sought to discover associations with the transcriptome (Chap. 6) and the metabolome (Chap. 7). Palle et al. (2013) measured the expression of 106 xylem development genes in a clonal association population using qRT-PCR (Chap. 6). Eighty SNPs were associated with the expression of these genes. This study was an early demonstration that the underlying causes of phenotypic variation in conifers also resulted from differential expression of genes in addition to structural variation among alleles which was already well established. Another interesting result from this study was that individuals homozygous for the rare SNP allele had reduced expression levels,

Table 11.2 SNPs significant for association with pitch canker resistance in *Pinus taeda* and best hits based on BLASTx search using the contig sequence as query

SNP_ID	Best hit (expect < l*e*-10)	Best hit (no cutoff)	Predicted SNP location	Effect on aa sequence	No. SNPs in LD
0_15227_01_159	ATP binding protein, lectin-like protein kinase	ATP binding protein, lectin-like protein kinase (expect = 7*e*-27)	Coding region	Synonymous	0
0_15382_01_99	Geranylgeranyl transferase type I beta subunit	Geranylgeranyl transferase type I beta subunit (expect = 2*e*-30)	Coding region	V to A	0
0_2234_01_128	Putative long-chain acyl-CoA synthetase	Putative long-chain acyl-CoA synthetase (expect = 5*e*-63)	Coding region	D to Y	0
0_6323_01_240	DELLA protein	DELLA protein (expect = 3*e*-59)	Coding region	Synonymous	0
0_9288_01_370	No hits found	Predicted protein *Populus trichocarpa* (expect = 4*e*-06)	Coding region	Synonymous	0
1_3327_01_113	No hits found	Unnamed protein product *(Vitis vinifera)* (expect = 0.23)	Coding region	C to Y	8
2_4484_02_622	Plastid hexose transporter	Plastid hexose transporter (expect = 8*e*-64)	Coding region	C to Y	0
2_6181_02_400	Hexokinase	Hexokinase (expect = 2*e*-31)	Noncoding, putative 3'-UTR	NA	0
2_8946_02_435	Cucumber peeling cupredoxin	Cucumber peeling cupredoxin (expect = 3*e*-10)	Non-coding, Putative 3'UTR	NA	1
CL4336Contigl_01_180	Unknown *(Picea sitchensis)*	Unknown *(Picea sitchensis)* (expect = 2*e*-72)	Coding region	Synonymous	1

From Quesada et al. (2010)
Predicted SNP location and effect on amino acid sequence are also shown on the basis of sequence alignments with genomic DNA sequences. aa, amino acid; LD, linkage disequilibrium; NA, not applicable

Table 11.3 Tests for associations between transformed oleoresin dry mass and SNPs within sequences similar to terpenoid biosynthetic genes in the *Pinus taeda* CCLONES population

Enzyme	Gene	Query GI#	#hits	#sig.
1-Deoxy-D-xyulose-5-phosphate synthase	DXS	215,478,267 215,478,265	2	0
1-Deoxy-D-xyulose 5-phosphate reductisomerase	DXR	215,478,269	0	–
2-*C*-Methyl-D-erythritol 4-phosphate cytidyl-transferase	MECT	73,672,048	0	–
4-(cytidine 5′-diphospho)-2-*C*-Methyl-D-erythritol kinase	CMEK	73,672,044 73,672,046	0	–
2-*C*- Methyl-D-erythritol 2,4-cyclodiphosphate synthase	MECS	40,849,972	1	0
1-Hydroxy-2-methyl-2-(*E*)-butenyl 4-diphosphate synthase	HDS	186,532,616	1	0
1-Hydroxy-2-methyl-2-(*E*)-butenyl 4-diphosphate reductase	IDS	126,697,259 126,697,261	1	0
Geranyl pyrophosphate synthase	GPPS	307,950,754	0	–
Geranyl-geranyl pyrophosphate synthase	GGPPS	17,352,450	3	0
Terpene synthase	TPS	28,894,481 59,800,271	1	0
Abietadienol/abietadienal oxidase (cyp450)	AO	59,800,265	7	2

From Westbrook et al. (2013)
Sequences from terpenoid biosynthetic genes were blasted against the *P. taeda* EST database used for SNP discovery. Query GI#, GenBank identifier of query sequences; #hits, number of BLASTX hits (e-value cutoff 10^{-8}) to ESTs containing SNPs; #sig, number of SNPs significantly associated with transformed oleoresin dry mass

suggesting that these are the deleterious forms that are being selected against in populations (Table 11.4). Moving one step beyond the transcriptome is the metabolome (Chap. 7) and using the same association population and SNP array as that used by Palle et al. (2013), Eckert et al. (2012a) found many SNPs associated with concentrations of 292 metabolites isolated from xylem tissue (Fig. 11.4). These "omic" phenotypes, combined with the earlier whole-plant physical and wood-chemistry phenotypes produce a developmental continuum of phenotypes determining wood properties in *P. taeda* and lead to using a broad network approach to understand the genetic control of phenotypic variation.

The major limitation of all the studies discussed so far in this section is the small portion of the genomic variation used in association studies. Assuming that *P. taeda* may have anywhere from 25,000 to 50,000 structural genes (Chap. 3), a SNP array with ~7000 SNPs can only account for a small portion of the segregating genetic variation accounting for phenotypic variation in populations. This situation began to change around 2010 with the arrival of next-generation sequencing technologies. The first association study to use a much greater number of SNPs was that of Lu et al. (2017) who used an exome-capture approach to discover 2.8 M SNPs among 384 clones of an association population. They tested these SNPs for association with a suite of phenotypes including specific leaf area, branch angle, crown width,

Table 11.4 Average gene expression values (ΔΔCt) of *Pinus taeda* individuals with different SNP alleles

Gene	Gene containing associated SNP	Average ΔΔCt values			Difference between ΔΔCT values of homozygotes
		Homozygous— common allele	Heterozygous	Homozygous—rare allele	
120PR	Unknown function	0.01 (270)	−0.09 (50)	1.76 (5)	1.75 [3.4]
AdeKin	Unknown function	0.10 (263)	−0.13 (62)	1.76 (3)	1.66 [3.2]
AIP	Phosphoribosylanthranilate transferase	−0.51 (348)	−0.56 (20)	1.75 (2)	2.26 [4.8]
αtub-1	Zinc finger domain protein	−0.83 (338)	−1.0 (26)	1.99 (2)	2.82 [7.1]
αtub-2	SEC14 cytosolic factor	0.16 (234)	−0.10 (88)	−1.28 (11)	1.44 [2.7]
βtub-1	SEC14 cytosolic factor	0.15 (234)	−0.13 (88)	−1.53 (11)	1.68 [3.2]
BQR	Fascilin-like arabinogalactan protein 18	−0.02 (220)	0.14 (95)	−1.06 (14)	1.04 [2.1]
CAD	Cystathionine γ-synthase-1	−0.15 (252)	0.08 (73)	1.7 (6)	1.85 [3.6]
CeSA9	FCA protein domain	0.08 (206)	0.1 (107)	−0.93 (22)	1.01 [2.0]
CeSA9	UDP-glycosyltransferase	0.08 (233)	−0.04 (88)	−1.8 (2)	1.88 [3.7]
GMP2	Dehydrogenase	−0.13 (196)	−0.08 (123)	1.43 (10)	1.56 [2.9]
GMP2	ATPase	−0.25 (169)	0.03 (124)	0.95 (18)	1.20 [2.3]
Importin	ATPase	−0.18 (169)	−0.10 (124)	0.85 (18)	1.03 [2.0]
NH-3	*PtMYB2*	−3.4 (328)	−1.48 (14)	0.61 (1)	4.01 [16.1]
NH-3	K homology RNA-binding domain protein	0.09 (279)	−0.16 (51)	−2.24 (2)	2.33 [5.0]
NH-3	Unknown function	0.05 (292)	0.11 (35)	−1.33 (5)	1.38 [2.6]
NH-6	DNA-binding protein	−0.14 (299)	−0.77 (30)	−2.4 (2)	2.26 [4.8]
NH-9	Predicted protein	0.08 (268)	0.18 (59)	−1.87 (4)	1.95 [3.9]

NH-9	Receptor protein kinase CLAVATA1 precursor	0.13 (285)	−0.18 (44)	−2.37 (2)	2.50 [5.7]
XET3	FLA8 of *Arabidopsis*	0.28 (290)	−0.68 (37)	−0.73 (3)	1.01 [2.0]
XGFT7	ATPase	−0.24 (169)	−0.07 (124)	0.82 (18)	1.06 [2.1]
UGP	Unknown function	0.33 (152)	−0.19 (146)	−0.76 (32)	1.09 [2.1]
UGP	Cysteine desulfurase/transaminase	−0.32 (189)	0.49 (101)	0.82 (13)	1.14 [2.2]

The numbers in () columns 3, 4, and 5 are the number of plants. The table includes all the significant associations that have both homozygotes, and the average expression difference between the homozygotes is at least one cycle. The numbers in [] in the last column indicate the fold difference between homozygotes. From Palle et al. (2013)

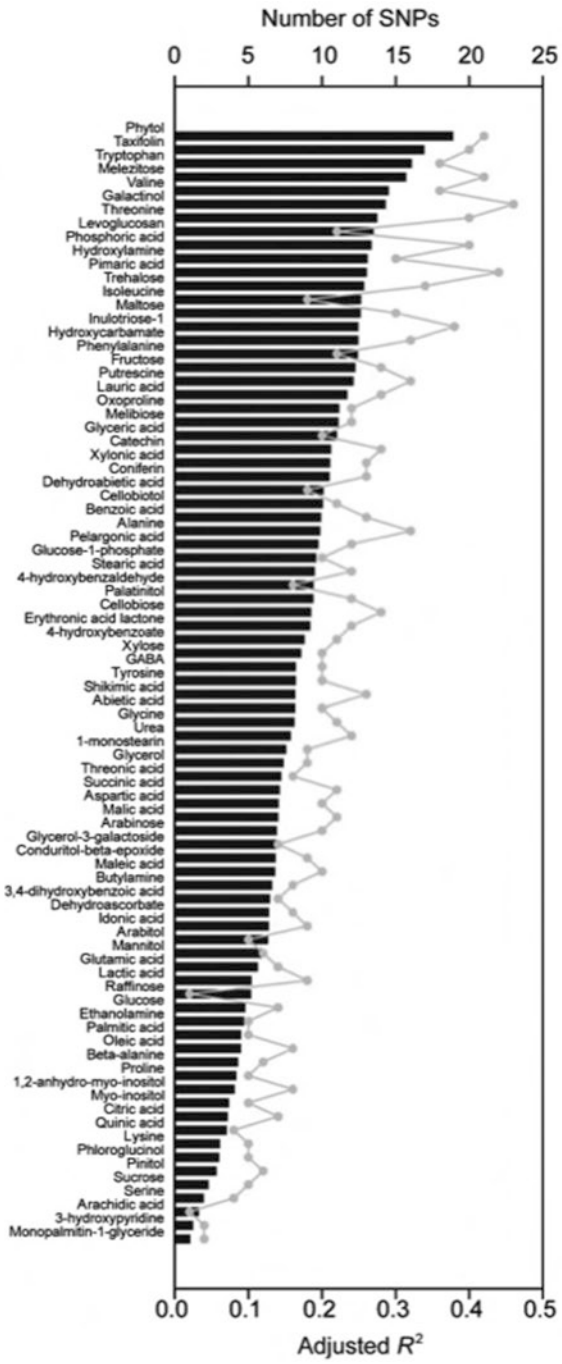


Fig. 11.4 Multilocus SNP models explain a large percentage of the phenotypic variance for many metabolites in *Pinus taeda*. Illustrated is the adjusted R^2 for marker effects from a linear model with population structure covariates and ancestry-corrected phenotypes as the dependent variable. The adjusted R^2 (R^2_{adj}) was calculated as: $R^2_{adj} = 1 - [(1 - R^2)((n - 1)/(n - k - 1))]$, where k is the number of independent predictors, n is the sample size, and R^2 is the coefficient of determination for the set of SNPs in the linear model. The gray line and points denote how many SNPs are in the linear model, which for each metabolite was the set of SNPs identified using the Bayesian mixed linear model in BAMD (Bayesian Association with Missing Data). The same patterns were also seen for the unknown metabolites. (From Eckert et al. 2012a)

stem diameter, total height, carbon isotope discrimination, nitrogen concentration, and pitch canker resistance. Given the very large number of SNPs used in association tests, the number positively associated with these phenotypes was surprisingly few, just 36. This may have resulted from over-correcting for false positives and have led to a high false-negative rate. De La Torre et al. (2018) used a different approach to discover and use more SNPs for *P. taeda*. Here, 10 trees were sequenced genome-wide and these sequences were mapped to the reference sequence (Neale et al. 2014) to call 455 M SNPs. These SNPs were then passed through a filtering process to select 635,453 SNPs that were used in an array to genotype all individuals in the association population. Tests of association were then done with 409 phenotypes including whole-plant phenotypes (height, DBH, carbon isotope discrimination, and pitch canker resistance) and molecular phenotypes (gene expression and metabolite concentration). Most of these phenotypic data were those used in earlier studies. This study found 2335 positive associations and used multiple test corrections similar to those of the Lu et al. (2017) study. Thus, even though the GWAS era has now arrived in conifers, it will take some time for studies to be replicated to be confident in distinguishing true associations from false discovery.

One final observation that has emerged from a summary of results from association studies and tests of neutrality of candidate genes (Chap. 10) was described by Eckert et al. (2013b). Here it was concluded that genes positively associated to phenotypes were also likely to be nonneutral, thus providing some experimental evidence that the genes underlying important phenotypes in *P. taeda* are also under selection in populations.

Pinus elliottii

This is the second most commercially important conifer species in the southeastern United States after *P. taeda*, yet very little has been done to dissect complex traits in it. *Pinus elliottii* has traditionally been used for lumber and for the extraction of resins. Two QTL studies have been done, both using RAPD markers. The first identified 14 chromosomal regions associated with aluminum tolerance (Kubisiak et al. 1999) and the second found 11 regions associated with early growth (Weng et al. 2002). Given the issues with repeatability with RAPD markers, this work is very preliminary and would need to be repeated if this research line were to go forward.

Pinus radiata

Native to California and Mexico, *Pinus radiata* is the most planted exotic conifer, grown primarily in Australia, New Zealand, and Chile. Given its advanced breeding and economic value, it would be expected that there would be many complex trait dissection studies; however, in this species much of this work has been done in a proprietary context and is not found in the literature. There are, however, several QTL mapping and association studies for wood quality, growth, and disease resistance phenotypes.

In the first reported QTL mapping study, Emebiri et al. (1998) searched for QTLs for height, diameter, and volume over a time course and classified the QTLs as being detected early, late, or in a curvilinear mode. They found that different QTLs were found at different time points. This result is not surprising based on classical quantitative genetic studies that trait variation and heritability can change over time and is a reminder that QTLs detected at a single time point may not be detectable at other times. Devey et al. (2004a) later mapped two QTLs for diameter in another family, although it was not determined if these were the same or different from those detected earlier.

QTL mapping and association studies for a variety of wood-property traits were done in several studies. Kumar et al. (2000) mapped QTLs for wood density in inner, middle, and outer rings and found one QTL. However, as with the diameter results, there were large developmental differences in the power of detection (Fig. 11.5). Devey et al. (2004a) also mapped QTLs for juvenile wood density and found several. These QTLs were also detected in another family, providing basis for validation which was not done in earlier studies. Dillon et al. (2010) were the first to conduct an association study in *P. radiata* which searched for association in a suite of wood-property traits including wood specific gravity, microfibril angle, modulus of elasticity, fiber coarseness, and more with 149 SNPs from wood-property candidate genes. They found some of the same associations as reported in *P. taeda* by González-Martínez et al. (2007). An interesting observation made by Dillon et al. (2010) was that a higher proportion of the associated SNPs were found in exons and 3'UTR regions than in introns or 5'UTR regions and a higher proportion were nonsynonymous than synonymous (Fig. 11.6). This result suggests that many of the positively associated SNPs might be functional nucleotide substitutions.

QTL mapping was also done for resistance to two pathogens, Dothistroma and pitch canker. Devey et al. (2004b) mapped four QTLs for Dothistroma resistance and, like the earlier work by this group for wood-quality traits, there was replication and validation. Moraga-Suazo et al. (2014) mapped QTLs for pitch canker resistance and discussed the importance of comparing their results with those found by Quesada et al. (2010) in *P. taeda*, although it was not clear if any associations were common to both studies.

The most recent association study reported for *P. radiata* was for two form traits, branch cluster frequency and stem straightness, using SNPs from wood-quality candidate genes (Li et al. 2016). Seven different SNPs were positively associated for each of the two traits, many of these coming from candidate genes associated with wood-property traits in *P. radiata* and *P. taeda.*

Pinus sylvestris

This is a widespread pine species of northern latitudes and is used in multiple breeding programs across that region. In contrast to this importance, the number of published complex trait dissection studies is rather limited (Table 11.1). Several growth,

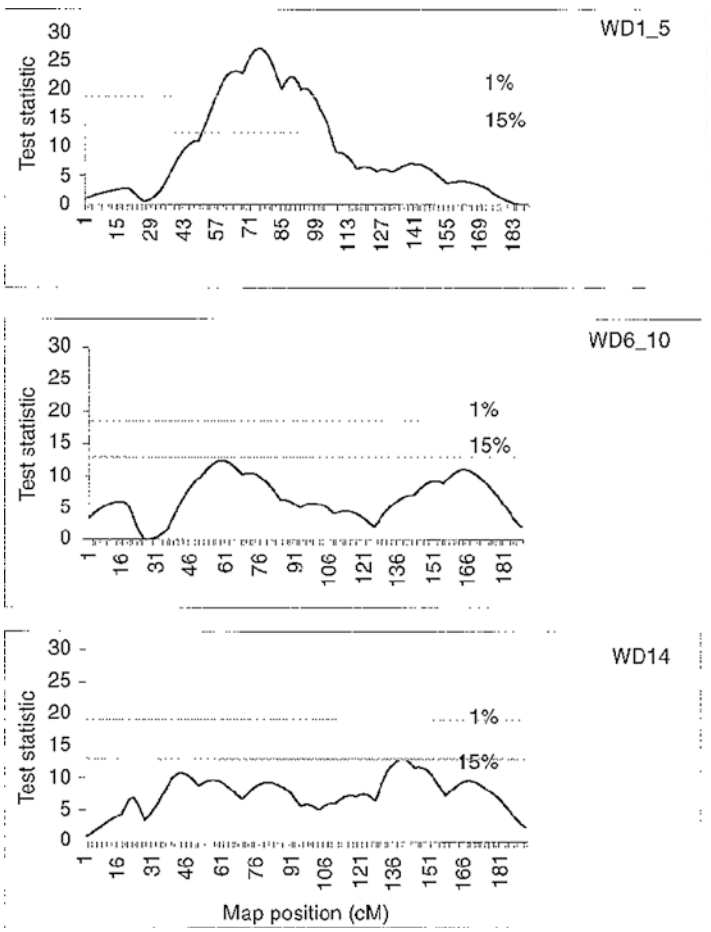


Fig. 11.5 QTL mapping for wood density in *Pinus radiata*. Test statistics for different positions (at every 1 cM) on chromosome 3 for different traits. Experiment-wise threshold levels (1% and 15%) are also given. (From Kumar et al. 2000)

wood-quality, growth rhythm, and, not surprisingly, cold-hardiness traits have been the subject of QTL mapping studies, although no association studies have yet been reported. All studies were done with RAPD or AFLP markers and had little or no replication within study and no cross-validation across studies; therefore the knowledge base for complex trait dissection in *P. sylvestris* is rather suspect and would seem to be far from application in marker breeding.

The first QTL mapping study in *P. sylvestris* was to map QTLs for bud set and frost hardiness (Hurme et al. 2000) using RAPD markers. To maximize segregation of alleles at QTL, they used an "open-pollinated backcross" from an F_1 tree that resulted from mating parental trees from the north and south of Finland. Four QTLs for bud set and seven QTLs for frost hardiness were found. Yazdani et al. (2003) did

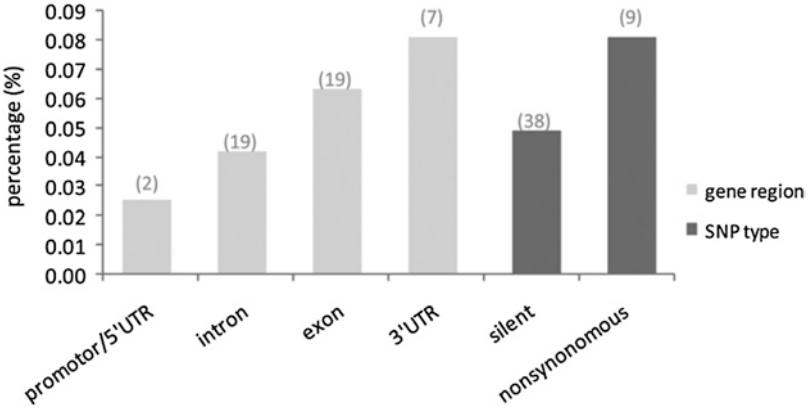


Fig. 11.6 Percentage of significant associations in *Pinus radiata* for wood-property traits, ($P < 0.05$) per gene region and SNP type. Differences in the proportion of significant associations for silent and nonsynonymous sites were significant ($P < 0.01$). (From Dillon et al. 2010)

a similar study with populations from Sweden, also using RAPD markers, and found QTLs for bud set and cold-hardiness but made no comparison to the work by Hurme et al. (2000). Lastly, growth trait QTLs were mapped in two studies (Lerceteau et al. 2000; Nowicka et al. 2013), both with AFLP markers and no replication, thus results from these studies must be considered preliminary. The most recent QTL study in *P. sylvestris* sought to map a suite of wood-property traits (percent earlywood, wood density, fiber width, and spiral grain angle) using statistical methods to integrate sampling over time (Li et al. 2014). The cross used in this study was the same as that used by Lerceteau et al. (2000), this time with many more progeny (1000). Many QTLs for these traits were mapped and are likely more robust than any of the earlier studies.

Pinus pinaster

Pinus pinaster is a Mediterranean pine of commercial importance in Europe for which there is a desire to apply marker-based breeding. Complex trait dissection using both QTL mapping and association studies has been conducted for a variety of traits including growth, wood properties, and drought tolerance. Plomion et al. (1996) began with a study to map height growth QTLs in seedlings three times over a two-year period. They found no QTLs in common across years which they argued was consistent with low juvenile-mature trait correlations for this trait. However, because this study was not replicated, it may be due to false discovery under low power of detection. Breeding for drought tolerance in *P. pinaster* is a high priority, and there have been several studies to discover QTLs for water-use efficiency. Brendel et al. (2002) conducted the first study where four significant and four suggestive QTLs for carbon isotope discrimination were mapped using AFLP markers (Table 11.5). Another QTL mapping study used a different and larger mapping

Table 11.5 Results of the composite interval mapping analysis for $\delta^{13}C$ and mean ring width (MRW) in *Pinus pinaster* using MultiQTL software

Trait	Map	Chr[a]	N[b]		Position ± SD[c]		LOD		*P*-value[d]	Dir.[e]	R^2	R_{total}^2
$\delta^{13}C$	Male	3a	149		4·3 ± 13·2		1·78		0·021*	+	0·047	0·268
	Male	6	164		102·7 ± 21·6		4·40		0·001***	+	0·124	
	Male	8	85		0·0 ± 15·7		1·85		0·021*	–	0·050	
	Male	9	183		104·6 ± 28·5		1·90		0·033*	–	0·047	
	Female	2	84		209·1 ± 56·4		2·30		0·019*	+	0·065	0·246
	Female	5	164		99·9 ± 22·0		1·98		0·003***	–	0·062	
	Female	12	153	1:	0·0 ± 36·7	I:	4·24	I:	0·002***	–	0·119	
				2:	135·4 ± 27·1	II:	1·88	II:	0·036*	+		
MRW	Male	2b	180	1:	47·2 ± 15·9	I:	2·49	I:	0·048*	–	0·181	0·429
			153	2:	51·8 ± 11·1	II:	1·57	II:	0·027*	+		
	Male	5	145		56·8 ± 33·4		2·41		0·022*	–	0·065	
	Male	6	174	1:	7·5 ± 31·0	I:	4·24	I:	0·001***	–	0·124	
			168	2:	133·7 ± 22·0	II:	2·68	II:	0·002*	+		
	Male	11	165		11·6 ± 30·9		1·73		0·040*	+	0·059	

The *P*-value associated with the LODs were calculated using 1000 permutations of the data, and standard deviation of position (SD) was calculated using the bootstrap method (1000 permutations); in case of a significant QTL pair for a chromosome, LOD and *P*-value are given for the tests of (I) two QTLs versus no QTLs and (II) two QTLs versus one QTL (difference of LOD for two and for one QTL). From Brendel et al. (2002)

[a]Chromosome ID

[b]Number of full-sibs with available data for QTL detection

[c]LOD score peak position (from the top) of the chromosome (cM)

[d]Probability for the null hypothesis of no QTL at the chromosome level. ***Probability corresponding to a 5% genome-wise type I error (significant QTL). *Probability corresponding to a 5% chromosome type I error (suggestive QTL)

[e]Direction of influence of presence of the allele for each QTL

population and found nine QTLs for carbon isotope discrimination (Marguerit et al. 2014). They discussed the importance of validation by comparing across studies, which they did for several *P. pinaster* maps (Fig. 11.7), but they made no comparison to what had been discovered in *P. taeda.* Finally, de Miguel et al. (2014) mapped five QTLs for water-use efficiency in another mapping population, and although there was discussion of comparative map validation, this was not done with that of Marguerit et al. (2014), presumably because both papers were published in the same year and there was no information exchange between groups.

QTL mapping and association studies have been done for wood-property traits as well. Pot et al. (2006) mapped QTLs for many physical and chemical wood-property traits in an F_1 family of 90 progeny and 219 AFLP markers. Fifty-four QTLs were reported but most were detected at a "suggestive" probability level. Only a small number were significant following multiple test correction. Several wood-property candidate genes were also positioned on the genetic map but only one, Korrigan, co-located with a QTL. Pot et al. (2006) did compare their results with those found by Brown et al. (2003) in *P. taeda* and found QTLs for lignin and alpha-cellulose content mapping to similar positions. Lepoittevin et al. (2012) conducted an association study for wood-chemistry traits (and also growth and form traits) using 184 SNPs from 40 candidate genes but found only one SNP in an HD-Zip III transcription factor positively associated with cellulose content (Fig. 11.8). The genetic map position of this gene was not given, so it could not be compared to the earlier QTL mapping study.

Bartholomé et al. (2016) conducted a large QTL mapping and association study for height growth and stem straightness. These studies were done with large mapping populations and many genetic markers so the power of detection was higher than in previous studies. Three QTLs and seven SNP associations were found for stem straightness and two QTLs and three SNP associations were found for height growth. Quite surprisingly, there was no co-location of QTLs and SNP association within the same trait (Fig. 11.9). This is a bit of a sobering result on the power and reliability of these types of studies.

One final association study in *P. pinaster* was conducted to discover genes underlying serotiny, a trait of great interest in several fire adapted species. Budde et al. (2014) found SNPs in several candidate genes associated with serotiny (Table 11.6). We will return to a discussion of the genetics of serotiny in the next section on *Pinus contorta.*

Pinus contorta

Pinus contorta is a widespread, pioneer species of conifer found in western North America. There are large active breeding programs in Canada and the United States; thus it is surprising that there has been almost no work in complex trait dissection (Table 11.1). The one published study used the association approach to identify genomic regions controlling serotiny (Parchman et al. 2012). Early research suggested that serotiny might be simply (qualitative) inherited, although aside from

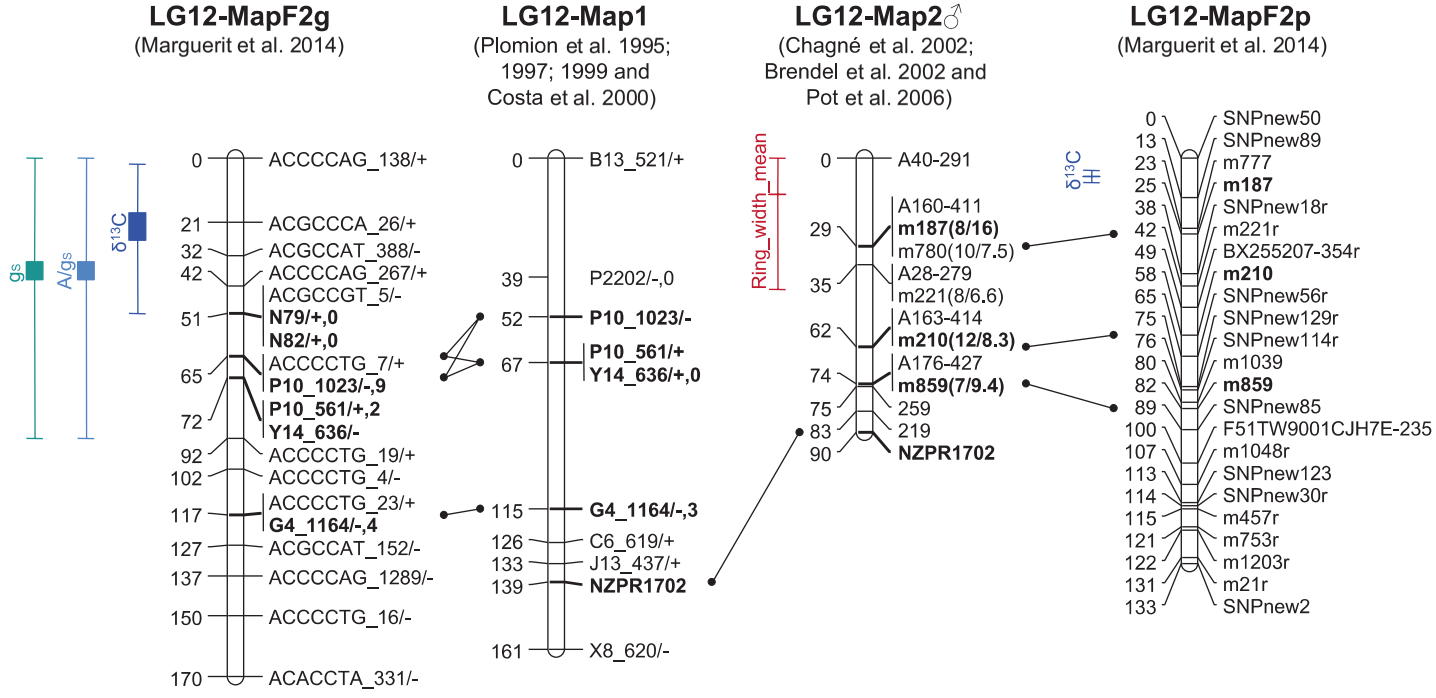


Fig. 11.7 Linkage group 12 (LG12) for the four genetic linkage maps of *Pinus pinaster* and associated QTLs. MapF2g corresponds to the genetic map published by Costa et al. (2000) from the F_2 greenhouse population. Map 1 corresponds to the genetic map published by Plomion et al. (1995) and used as bridge to align MapF2g to Map 2♂. Map 2♂ corresponds to the male map published by Chagné et al. (2002) on which SNP markers were added as accessory markers by Chancerel et al. (2011). The framework Map 2 was used for QTL detection of water-use efficiency by Brendel et al. (2002). MapF2p corresponds to the map established by Chancerel et al. (2013) from the analysis of the F_2 plantation population. Bars indicate the range of mean positions from *L* and L_{BS}. (From Marguerit et al. 2014)

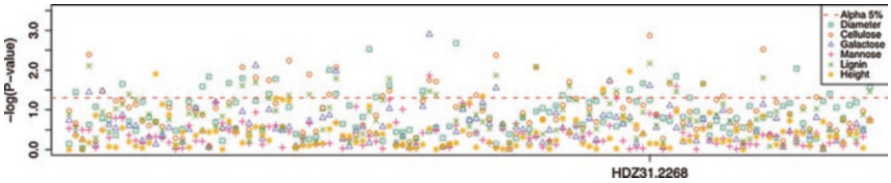


Fig. 11.8 *P* value for the SNP effect for 121 informative SNPs associated with several wood-quality and growth traits in *Pinus pinaster*. (From Lepoittevin et al. 2012)

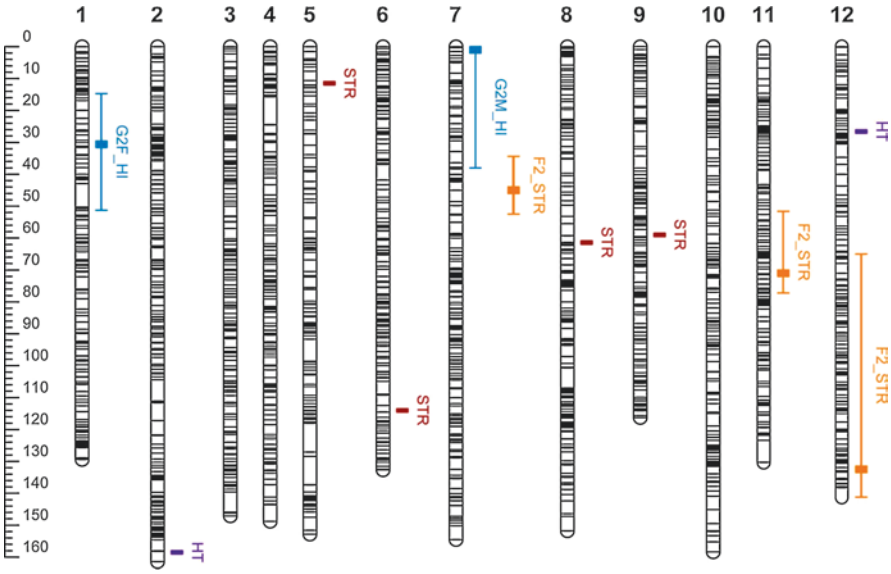


Fig. 11.9 Position of the QTLs detected by QTL mapping and association on the *Pinus pinaster* composite genetic map. For QTL mapping, QTLs for stem straightness (STR) are shown in orange and QTLs for height increment (HI) are shown in blue. The whiskers indicate the 95% credible interval around QTL peaks. The locations of markers significantly associated with height (HT, purple) and stem straightness (STR, red) are also indicated. On linkage group 2, the two significant associations for HT are co-located (158 cM and 158.1 cM). (From Bartholomé et al. 2016)

trees in populations that are entirely serotinous or entirely nonserotinous, trees with both cone types are also found. This suggested that the inheritance of serotiny might be more quantitative. Parchman et al. (2012) used a genomic sequencing approach to type 98 trees, 48 serotinous and 50 nonserotinous. They performed an association analysis with these data and found 11 SNPs that were associated with the phenotype (Table 11.7). None of the SNPs were in linkage disequilibrium, suggesting these are independent loci. Annotation of these sequences did not reveal genes of any known function. It is possible with the much-improved transcriptomic and genome sequence resources since 2012, these associated sequences might be successfully annotated to a gene of known function.

Table 11.6 Significant marker effects of common SNPs (minor allele frequency (MAF) > 0.10) on serotiny in *Pinus pinaster*, as identified by a two-step approach based on mixed-effects linear models (MLMs) and Bayesian genetic association

SNP	Annotation	SNP motif	Site annotation[a]	LG	MAF	*N*	Genetic model	Marker effects MLM			Bayesian model
								F	*P*	R^2	Mean γ (95% CIs)
m15	Defectively organized tributaries 2 (DOT2)	T/C	nc		0.3706	197	A	6.052	0.003	0.062	−0.2467 (−0.4653, −0.0282)
m594	Pyrophosphate-energized vacuolar membrane proton (AVP)	T/C	syn	8[b,c]	0.1878	197	D	6.915	0.009	0.035	−0.3678 (−0.6264, −0.1105)
m692	Unknown	A/G	unk	3[b]	0.4133	196	D	12.932	0.0004	0.067	−0.4696 (−0.7964, −0.1426)
m696	Arabinogalactan-like protein (AGP)	C/G	nc	3[b,c]	0.4031	196	D	5.722	0.018	0.029	−0.3206 (−0.5773, −0.0658)
m698	Nascent polypeptide-associated complex subunit alpha-like protein (NAC-alpha)	T/C	syn		0.2864	199	A	ns	ns	ns	0.2938 (0.1130, 0.4751)
m705	Carotenoid cleavage dioxygenase (CCD)	A/G	nc		0.1231	195	D	6.737	0.010	0.034	−0.3083 (−0.5993, −0.0197)
m816	Receptor protein kinase clavatal (CLVI)	C/G	syn	1[b,c]	0.4924	197	O	ns	ns	ns	0.3391 (0.1080, 0.5706)
m912	Peroxidase 72-like (PER72)	A/T	non-syn		0.3795	195	D	5.791	0.017	0.031	0.2601 (0.0205, 0.4992)
m955	Unknown	A/G	unk	3[b,c]	0.2475	198	A	3.049	0.050	0.031	0.2294 (0.0469, 0.4160)
m974	1-Aminocyclopropane-1-carboxylate synthase (ACC)	A/G	syn	11[c]	0.1231	199	D	5.387	0.021	0.027	0.2946 (0.0137, 0.5731)

(continued)

Table 11.6 (continued)

SNP	Annotation	SNP motif	Site annotation[a]	LG	MAF	*N*	Genetic model	Marker effects MLM			Bayesian model
								F	*P*	R^2	Mean γ (95% CIs)
m1194	Cell division-related protein	C/G	syn		0.1439	198	A	3.714	0.026	0.038	0.2501 (0.0275, 0.4744)
m1196	Peptidyl-prolyl cis-trans isomerase (PPI)	A/C	syn		0.3266	199	A	ns	ns	ns	−0.1994 (−0.3871, −0.0086)

From Budde et al. (2014)

Bayesian mean allelic effects (γ) and 95% confidence intervals (CIs) were obtained from the distribution of the last 20,000 iterations in BAMD (Bayesian Association with Missing Data; Li et al. 2012a). Allelic effects are provided for the genetic model (A, additive; O, over-dominance; D, allele dominance) with higher effect on fire phenotype. Marker names and linkage groups (LG) as reported in Chancerel et al. (2011) and de Miguel et al. (2014); ns, not significant for that particular genetic model

[a]Site annotation: nc, noncoding (untranslated regions or introns); nonsyn, nonsynonymous; syn, synonymous; unk, unknown

[b]LG from Chancerel et al. (2011)

[c]LG from de Miguel et al. (2014)

Table 11.7 Identifiers for genetic regions containing nucleotide polymorphisms associated with serotiny in *Pinus contorta*, minor allele frequency for each locus (MAF), genotype-specific parameter estimates for the probability of serotiny conditional on genotype

		Probability of serotiny			
Identifier	MAF	AA	AA'	A'A'	Accession
65,253	0.18	0.542 (0.16–0.87)	0.429 (0.34–0.52)	**0.704** (0.51–0.86)	AC241283
112,487	0.16	0.485 (0.13–0.85)	**0.760** (0.51–0.93)	0.456 (0.37–0.54)	AC241311
54,398	0.22	0.488 (0.13–0.86)	**0.734** (0.50–0.90)	0.455 (0.37–0.54)	NA
1428	0.2	0.550 (0.21–00.86)	0.433 (0.35–0.52)	**0.781** (0.57–0.92)	NA
2539	0.31	0.444 (0.12–0.80)	0.385 (0.28–0.50)	**0.649** (0.51–0.77)	AC241288
103,454	0.18	0.480 (0.13–0.86)	**0.274** (0.10–0.48)	0.539 (0.45–0.63)	AC241359
17,466	0.19	0.494 (0.13–0.86)	**0.339** (0.20–0.49)	0.566 (0.46–0.67)	AC241351
1994	0.26	0.497 (0.14–0.86)	0.359 (0.24–0.48)	**0.605** (0.50–0.71)	AC241292
64,526	0.25	0.483 (0.13–0.85)	**0.307** (0.15–0.47)	0.567 (0.46–0.66)	AC241284
1853	0.18	0.484 (0.39–0.57)	0.708 (0.48–0.88)	**0.188** (0.04–0.44)	EU998740
40,518	0.2	0.571 (0.18–0.89)	0.601 (0.48–0.71)	**0.347** (0.23–0.48)	NA

The significant parameter for the locus is in bold. From Parchman et al. (2012)

Pseudotsuga menziesii

Pseudotsuga menziesii is a widespread and extremely important conifer species of western North America. It is the primary timber species of this region and is the object of many advanced breeding programs in the United States and Canada. For these reasons it was the target of early complex trait dissection studies in conifers. The species is found across a highly heterogeneous landscape; therefore in addition to standard traits of interest such as growth and wood properties, adaptation to the environment is of high importance to breeders and resource managers. The adaptive traits studied in QTL and association studies include bud phenology (bud flush and bud set) and cold-hardiness (Table 11.1). Drought tolerance would also be important, but this has not yet been a target of study.

Jermstad et al. (2001a) estimated QTLs for both terminal and lateral spring bud flush over 4 years at two test sites. Each individual of the mapping population was clonally replicated to allow testing at two sites and replication within site; thus the precision of phenotyping was much higher than studies where there was no replication of members of the mapping population. Twenty-one highly significant QTLs were detected (Table 11.8), many of which were detected across multiple years. However, the same QTLs were not often found across sites. These results suggest temporal stability of QTL expression but spatial heterogeneity. In a second study using essentially the same mapping population, Jermstad et al. (2001b) estimated QTLs for spring and fall cold-hardiness. Multiple QTLs were detected in both spring and fall, a few of which were co-located on the genetic map (Fig. 11.10). Cold-hardiness evaluation was done on three tissue types; buds, needles, and stems. The same QTLs were generally detected across tissue types in the spring but not in the fall. This result is consistent with common garden studies that have shown synchronization across tissues during de-acclimation in the spring. Finally, there were

Table 11.8 Bud-flush QTLs detected in *Pseudotsuga menziesii* from sites in the US states of Washington and Oregon

Trait	LG	Map position (cM)	*F*-value	Effect			Proportion var. (%)
				Pat. (SE)	Mat. (SE)	Pat. × Mat. (SE)	
Washington							
wlat95	3	45	4.71**	−0.09 (0.03)	0.01 (0.03)	0.01 (0.03)	5.6
wlat95	4	89	5.70**	–[a]	0.05 (0.05)	–	2.4
wlat95	9	0	5.71**	0.05 (0.02)	–	–	2.4
wlat96	4	89	10.22**	–	0.07 (0.02)	–	4.7
wlat96	9	0	6.91**	0.05 (0.02)	–	–	3.0
wlat97	2	115	4.28*	0.00 (0.03)	0.07 (0.05)	0.15 (0.05)	4.9
wlat97	4	89	6.22**	–	0.06 (0.03)	–	2.7
wter95	3	45	5.42**	−0.11 (0.03)	0.01 (0.03)	−0.01 (0.03)	6.6
wter95	4	89	5.56**	–	0.06 (0.03)	–	2.4
wter95	9	5	7.11**	0.07 (0.03)	–	–	3.1
wter96	4	89	7.45**	–	0.07 (0.03)	–	3.3
wter96	15	5	3.62*	0.01 (0.03)	−0.09 (0.04)	−0.09 (0.04)	4.0
wter97	8	55	4.19*	−0.10 (0.03)	0.01 (0.04)	−0.03 (0.04)	4.8
wflu98	2	115	4.47**	−0.05 (0.04)	0.09 (0.08)	0.28 (0.08)	5.2
wflu98	4	89	8.72**	–	0.13 (0.02)	–	3.9
wflu98	9	50	5.02*	0.10 (0.05)	–	–	2.1
wflu98	10	10	4.17*	0.00 (0.05)	0.16 (0.05)	0.01 (0.05)	4.8
Oregon							
oflu98	7	31	4.00*	−0.07 (0.09)	0.11 (0.09)	−0.31 (0.01)	9.7

oflu98	10	40	4.63**	−0.16 (0.14)	0.43 (0.14)	−0.09 (0.23)	11.5
oflu98	11	28	4.56*	–	0.19 (0.09)	–	4.1
oflu98	12	11	4.55**	0.78 (0.08)	−0.23 (0.08)	−0.05 (0.08)	11.2

Linkage group (LG) and map position in centimorgans (cM) are presented for each QTL along with *F*-values, paternal (Pat.), maternal (Mat.), and paternal×maternal interaction (Pat. × Mat.) effects with standard errors (SE) in parentheses, and the proportion of total phenotypic variance explained by the QTL. In the trait names, *lat* is for lateral bud flush, *ter* is for terminal bud flush, and *flu* is for both lateral and terminal; the two digits are for the year it was measured from 1995 to 1998. From Jermstad et al. (2001a)

* $p \leq 0.01$, ** $p \leq 0.005$

[a]Markers on this linkage group provide information for only one parent

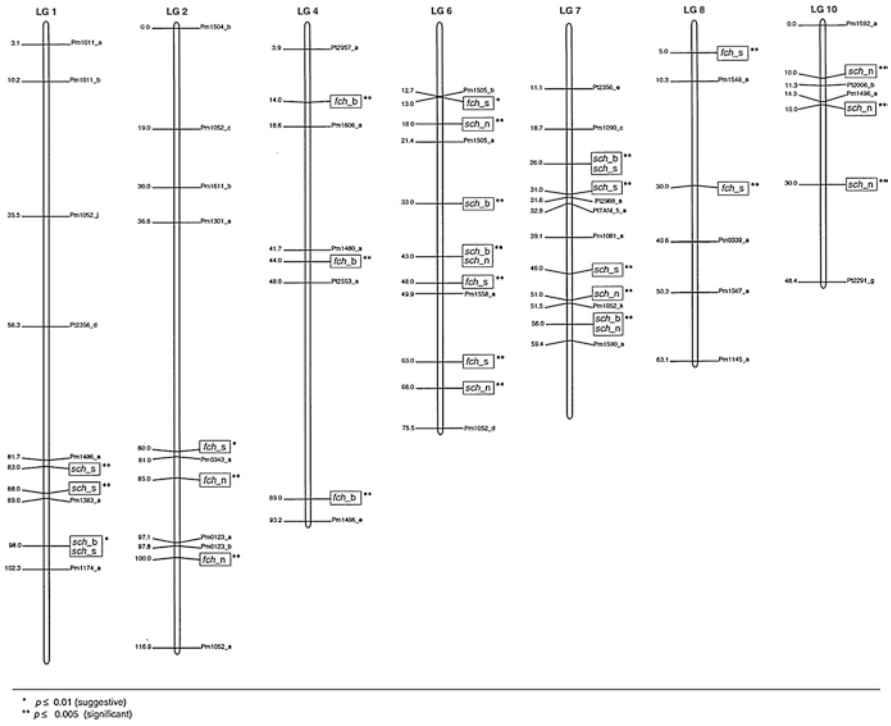


Fig. 11.10 QTL map positions for genes influencing fall and spring cold-hardiness in *Pseudotsuga menziesii*. Shown *within boxes* are the positions of 11 unique QTLs for fall cold-hardiness (fch) and 15 unique QTLs for spring cold-hardiness (sch). QTLs were estimated at 5-cM intervals. QTLs are labeled either suggestive * ($p \leq 0.01$) or significant ** ($p \leq 0.005$); if one or more traits were associated with the QTL at the significant level, then the QTL is labeled significant. (From Jermstad et al. 2001b)

also cases of co-location of bud phenology QTLs and cold-hardiness QTLs suggesting that some aspects of cold-hardiness are a function of bud phenology.

Jermstad et al. (2003) then conducted a validation study for bud phenology using a much larger mapping population. The unique aspect of this study was that QTLs were estimated under experimental treatment conditions, something rarely done in non-model systems. This was possible because all members of the mapping population were cloned allowing for multiple experiments. Spring bud-flush QTLs were estimated under two levels of winter chill and three levels of spring heat sum. Fall bud set QTLs were estimated under two levels of day length and two levels of moisture stress. These treatments were chosen as they reflect cues from the environment which affect the timing of bud flush and bud set. This experimental design also allowed the estimation of QTL × treatment interactions. For example, for the trait elapsed time between bud flushes under lammas growth, a QTL was found on linkage group 3 that also had a significant interaction with moisture stress (Fig. 11.11). This relationship with the environmental signal could aid in eventually identifying

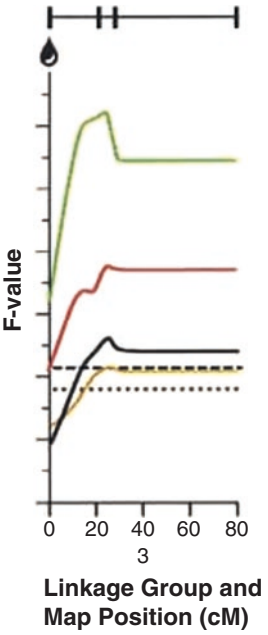


Fig. 11.11 QTLs for elapsed time between bud flushes (EBF) during lammas growth in *Pseudotsuga menziesii*. This QTL, found on linkage group 3, was detected in all four experimental conditions: winter chill, spring heat sum, day-length, and moisture stress. (From Jermstad et al. 2003)

the functional gene underlying the QTL. Wheeler et al. (2005) also did QTL mapping to validate cold-hardiness QTLs. Many of the spring cold-hardiness QTLs were detected from the earlier study in addition to several new QTLs. Several potential candidate genes controlling cold-hardiness were also positioned onto the genetic map and several co-located with QTLs (Fig. 11.12).

One association study has been conducted to identify genes controlling cold-hardiness in *P. menziesii*. Eckert et al. (2009a) found 30 SNPs from 12 candidate genes associated with different cold-hardiness phenotypes. The position of these positively associated candidate genes was not known relative to cold-hardiness QTLs, but now that the *P. menziesii* genome has been sequenced (Neale et al. 2017a), this should be possible.

A large suite of wood-quality and growth traits have also been QTL mapped in *P. menziesii* using multiple families and AFLP markers (Ukrainetz et al. 2008). Effect sizes for individual QTLs were quite small as is consistent with wood-quality QTL mapping in other conifers. There was no attempt to compare map position of these QTLs with those found in *Pinus taeda,* even though a comparative map between these two species was available (Krutovsky et al. 2004).

Picea ssp.

Complex trait dissection has been done in *Picea* species, although in just four of the major commercial species, *P. abies*, *P. glauca*, *P. mariana*, and *P sitchensis* (Table 11.1). Traits of interest are the same as for other species; growth, wood

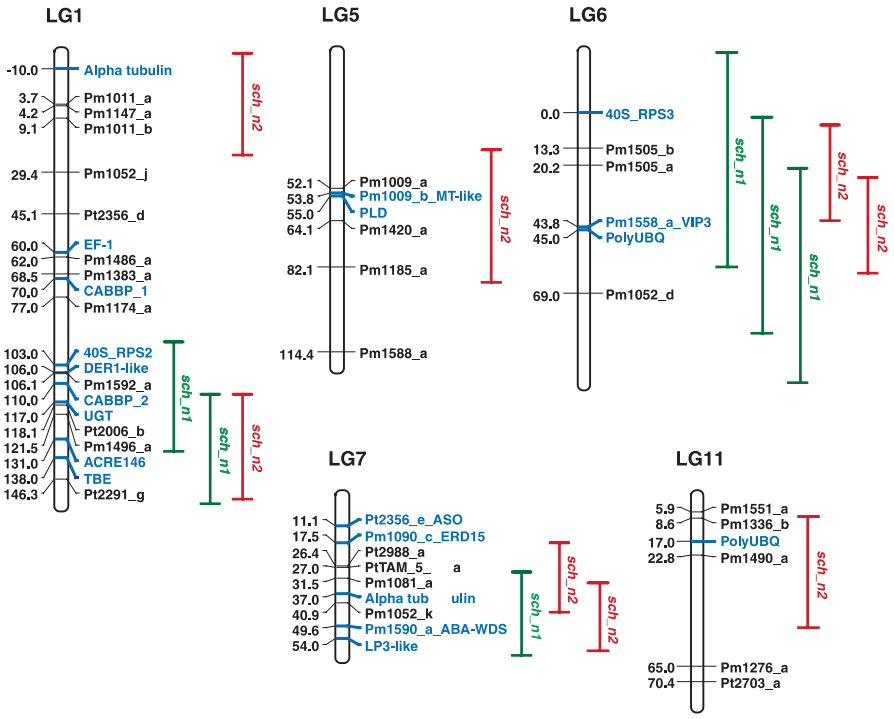


Fig. 11.12 Unique QTLs for spring needle cold-hardiness in *Pseudotsuga menziesii* are presented with 95% CI bars on the consensus map for Cohorts 1 (*sch_n1*) and 2 (*sch_n2*). Framework markers used for QTL mapping and candidate genes, noted in blue, bold type, are indicated. (From Wheeler et al. 2005)

properties, bud phenology, cold-hardiness, and disease resistance. The first studies used bulked segregant analysis (Michelmore et al. 1991) to identify AFLP markers linked to wood density (Markussen et al. 2004) and wood extractive content (Markussen et al. 2005) in *P. abies*. The linked AFLP markers were converted to SCAR (sequence characterized amplified region) markers for ease and reliability in application. The relationship of these QTLs to those found in *Pinus* and *Pseudotsuga* is unknown.

Two very large and comprehensive association studies have been done in *Picea glauca* to identify SNPs in candidate genes associated with a suite of wood-property traits. Beaulieu et al. (2011) tested 944 SNPs from 549 candidate genes with 25 wood-property traits and found 25 different associations (Table 11.9). They made a significant effort to compare their discovery with that found in *Pinus taeda* and did find several common associations, a rare example of cross-species validation. Later, Lamara et al. (2016) used the association study data from Beaulieu et al. (2011) and gene expression data from Raherison et al. (2015) (Chap. 6) to validate many of these associations based on the addition of co-expression networks.

Table 11.9 Significant associations between gene SNPs and wood traits in *Picea glauca* after correction for multiple testing

Trait[a]	SNP	Gene ID	Functional annotation
Earlywood			
ARW	*PGWD1-1094*	*GQ03806_H09*	LACS9
CWT	*PGWD1-1396*	*GQ02829_F04*	Glycoside hydrolase family 28
	PGWD1-1096	*GQ04010_I15*	Pectinesterase
	PGWD1-1282	*GQ03113_N22*	MYB4
PCT	*PGWD1-0511*	*GQ0224_E23*	Galactosyl-transferase XT2
	PGWD1-1313	*GQ03211_O01*	Receptor-like kinase
	PGWD1-0560	*GQ03010_F13*	ACC oxidase
	PGWD1-1035	*GQ0172_O22*	β-Expansin
Latewood			
ARW	*PGWD1-0863*	*GQ03005_C12*	Tubulin, TUA2
	PGWD1-1313	*GQ03211_O01*	Receptor-like kinase
	PGWD1-0581	*GQ03006_P17*	Unknown
	PGWD1-0560	*GQ03010_F13*	ACC oxidase
Total wood			
ARW	*PGWD1-1094*	*GQ03806_H09*	Acyl lipid metabolism
CWT	*PGWD1-1396*	*GQ02829_F04*	Glycoside hydrolase family 28
	PGWD1-1096	*GQ04010_I15*	Pectinesterase
Earlywood			
CWT	*PGWD1-1282*	*GQ03113_N22*	MYB4
CRS	*PGWD1-1282*	*GQ03113_N22*	MYB4
PCT	*PGWD1-0354*	*GQ04010_P08*	Tubulin, TUB3
	PGWD1-1035	*GQ0172_O22*	β-Expansin
Latewood			
ARW	*PGWD1-0863*	*GQ03005_C12*	Tubulin TUA2
	PGWD1-0354	*GQ04010_P08*	Tubulin TUB3
MFA	*PGWD1-0107*	*GQ0133_K12*	Glycosyl hydrolase family 10
RCD	*PGWD1-1070*	*GQ02908_P24*	β-TIP
Total wood			
CWT	*PGWD1-1282*	*GQ03113_N22*	MYB4
CRS	*PGWD1-1282*	*GQ03113_N22*	MYB4

From Beaulieu et al. (2011)

[a]Trait abbreviations: *ARW* average ring width, *CWT* tracheid cell wall thickness, *PCT* percentage of earlywood, *CRS* tracheid coarseness, *MFA* microfibril angle, *RCD* tracheid cell diameter in radial direction

As with most other north temperate conifer species, the adaptive traits, bud phenology, and cold-hardiness have been the subject of study in several species of *Picea*. Pelgas et al. (2011) mapped many QTLs for bud phenology and height growth in two clonally replicated populations using SNP markers. This study was quite successful in validating QTLs across pedigrees, years, and environments. They also compared their overall results quite thoroughly with earlier results from *Pseudotsuga menziesii*, although not down to the level of identifying QTLs

common to both species. Now that the *Picea glauca* and *Pseudotsuga menziesii* genomes have both been sequenced (Chap. 3), it should now be possible to investigate common associations more thoroughly. Prunier et al. (2013) conducted a combined QTL mapping and association study in *Picea mariana* to discover QTLs and SNP associations for bud set and height growth. A unique aspect of this study was that "candidate SNPs" were identified from candidate genes using outlier analysis (Chap. 10). SNPs in candidate genes were associated with bud set timing, and some genes were common to those found in *P. glauca.* Furthermore, they also tested for SNP association with climate variables (moisture and temperature) and found many SNPs associated with the environment as well as with the phenotype. One difference in this study, relative to the general result from *P. glauca* and *Pseudotsuga menziesii*, was a lack of QTL stability across years (temporal). In *Picea sitchensis*, Holliday et al. (2010) searched for candidate gene SNPs associated with bud set timing and fall cold-hardiness and with environmental variables. Results from this study were similar to that from Prunier et al. (2013) in *P. mariana* which suggests parallel evolution for adaptive traits in related species. Finally, Holliday et al. (2012) conducted a Random Forest analysis (Breiman 2001) to predict adaptive phenotypes from the SNP data. This approach is similar to using genomic selection to predict phenotypes from genetic marker data in breeding applications.

There has been just one study in *Picea* to dissect complex disease resistance (Lind et al. 2014). They did QTL mapping for four different traits (lesion length at inoculation site, fungal spread within sapwood, exclusion of the pathogen from the host after initial infection, and ability to prevent infection) related to resistance to the fungus *Heterobasidion parviporum* using a large F_1 family and SNP markers. They found QTLs for all four traits (Fig. 11.13) but none common to more than one trait. This result supports the general conclusion of multiple and quantitatively inherited resistance mechanisms to plant disease, in this case to a root rot.

Larix ssp.

Complex trait dissection in species of the genus *Larix* have been done in only two cases. Guan et al. (2011) used 145 offspring from a *L. kaempferi* × *L. gmelini* hybrid cross with RAPD markers to map QTLs for height, diameter, wood specific gravity, tracheid length, and tracheid width. They found QTLs for all these traits, but given there was no replication and RAPD markers were used, these results are unverified. Wang et al. (2015) conducted an association study in *L. olgensis* with just one

Fig. 11.13 (continued) levels of significance for the individual trait and group. Wide colored areas between curve and group show the QTL confidence interval based on a 1 LOD drop from the QTL peak. The colored marker names denote SNPs within the QTL confidence intervals, the asterisks signify the level of significance according to the Kruskal-Wallis test (ranging from $p \leq 0.1$ (*) to $p \leq 0.????$ (**)). The marker names in boldface indicate the designated cofactors. "D" after a marker name indicates a segregation pattern deviating from the expected Mendelian ratios of 1:1 or 1:2:1 ($0.005 > P < 0.05$, χ^2). (From Lind et al. 2014)

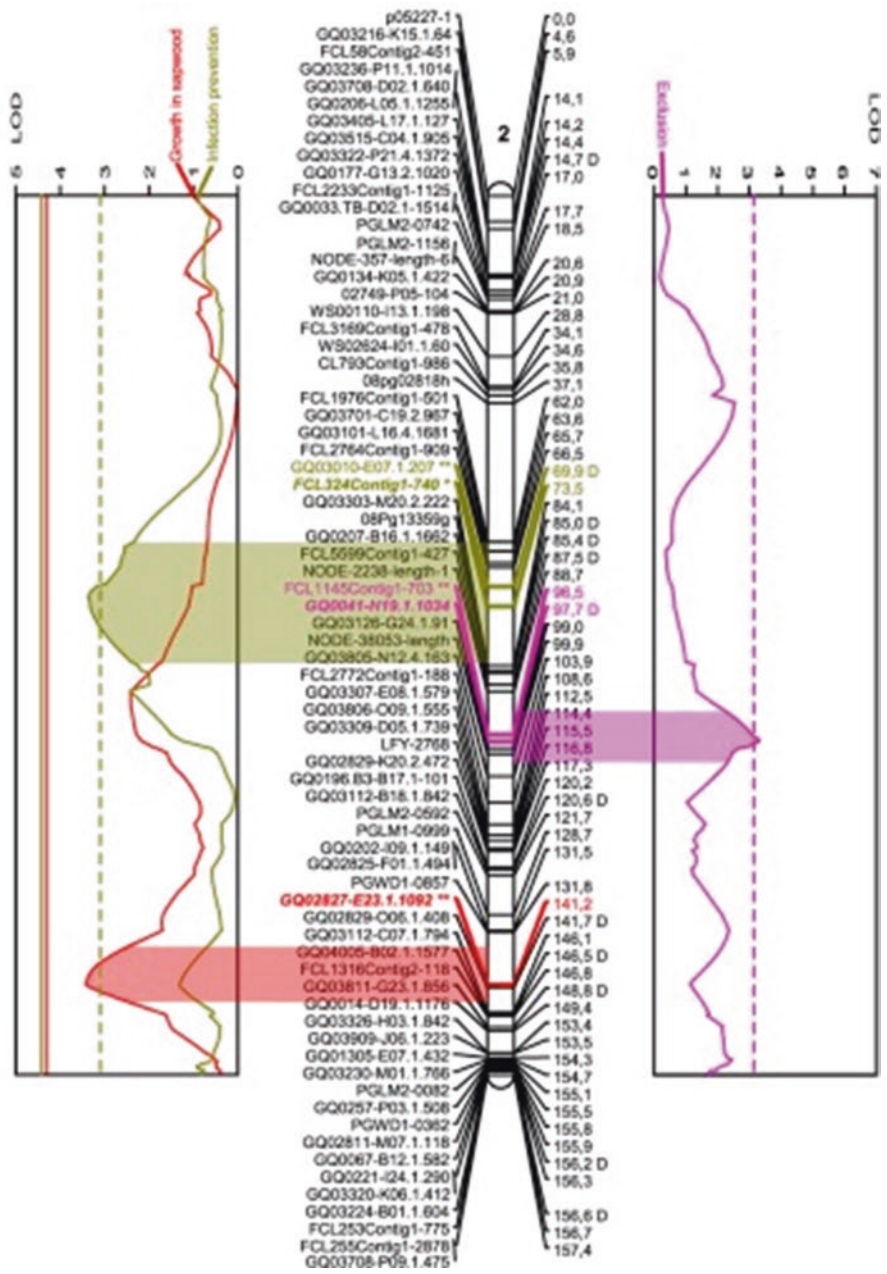


Fig. 11.13 Linkage group 2 of the *Picea abies* genome and the QTLs for various resistance traits. Names of the SNP markers are displayed on the left of the linkage groups. Genetic distance (cM) is indicated on the right. The red curve on the left indicates fungal growth within sapwood, the purple curve on the right indicates fungal exclusion, and the orange [greenish?] curve on the right indicates infection prevention. The complete and dashed vertical lines describe the 0.1% and 5%

candidate gene, cinnamyl alcohol dehydrogenase (CAD), and a suite of growth and wood-property traits (height, diameter, volume, wood specific gravity, carbon content, carbon concentration, lignin content, and cellulose content). There were 47 SNPs across the CAD gene, and they found significant associations for height, lignin content, carbon content, and wood density but not for the other traits. Associations of the CAD and some of these traits had previously been found in *Pinus taeda* and *P. radiata.*

Cryptomeria japonica

Cryptomeria japonica is the most important conifer species in Japan both in terms of artificial plantations and natural forests. Timber from *C. japonica* is used for building construction, furniture, and many other uses. An important negative issue with *C. japonica*, however, is that its pollen has serious allergenic properties for the human population. For this reason, there is great interest in understanding the genetic control of pollen production and male sterility.

The first complex trait dissection study was done by Yoshimaru et al. (1998) who mapped QTLs for both male and female flowering as well as for growth and rooting ability. This early study suggested that genes controlling male and female flowering were different (Fig. 11.14). Goto et al. (2003) mapped a cDNA clone, Cry j 1, that was known to be involved in the allergenic response, but its map position relative to the QTLs mapped earlier was not described. Ujino-Ihara et al. (2012) discovered one major QTL for male strobilus production and showed by comparative mapping that this QTL was likely the same as one first reported by Yoshimaru et al. (1998) (Fig. 11.15). Genes controlling for male sterility had been discovered earlier based

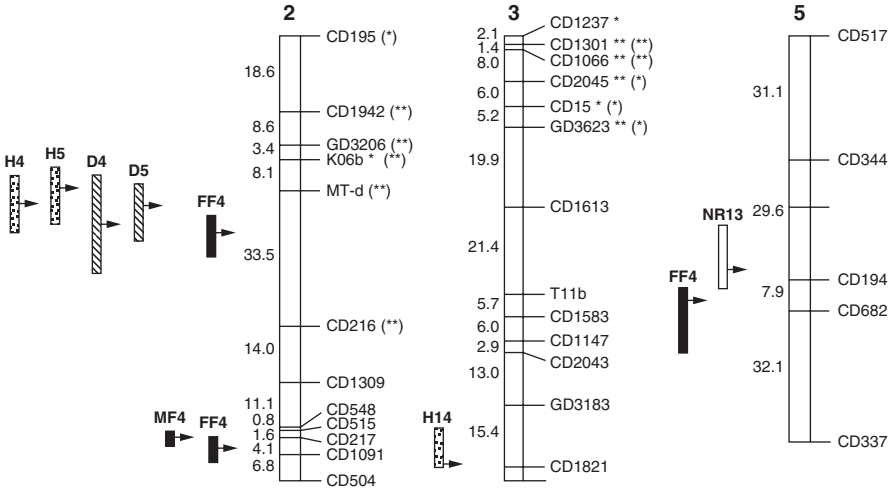


Fig. 11.14 Map position for QTLs for height (H), diameter (D), female flower (FF), and male flower (MF) at years 4 and 5 in *Cryptomeria japonica*. (From Yoshimaru et al. 1998)

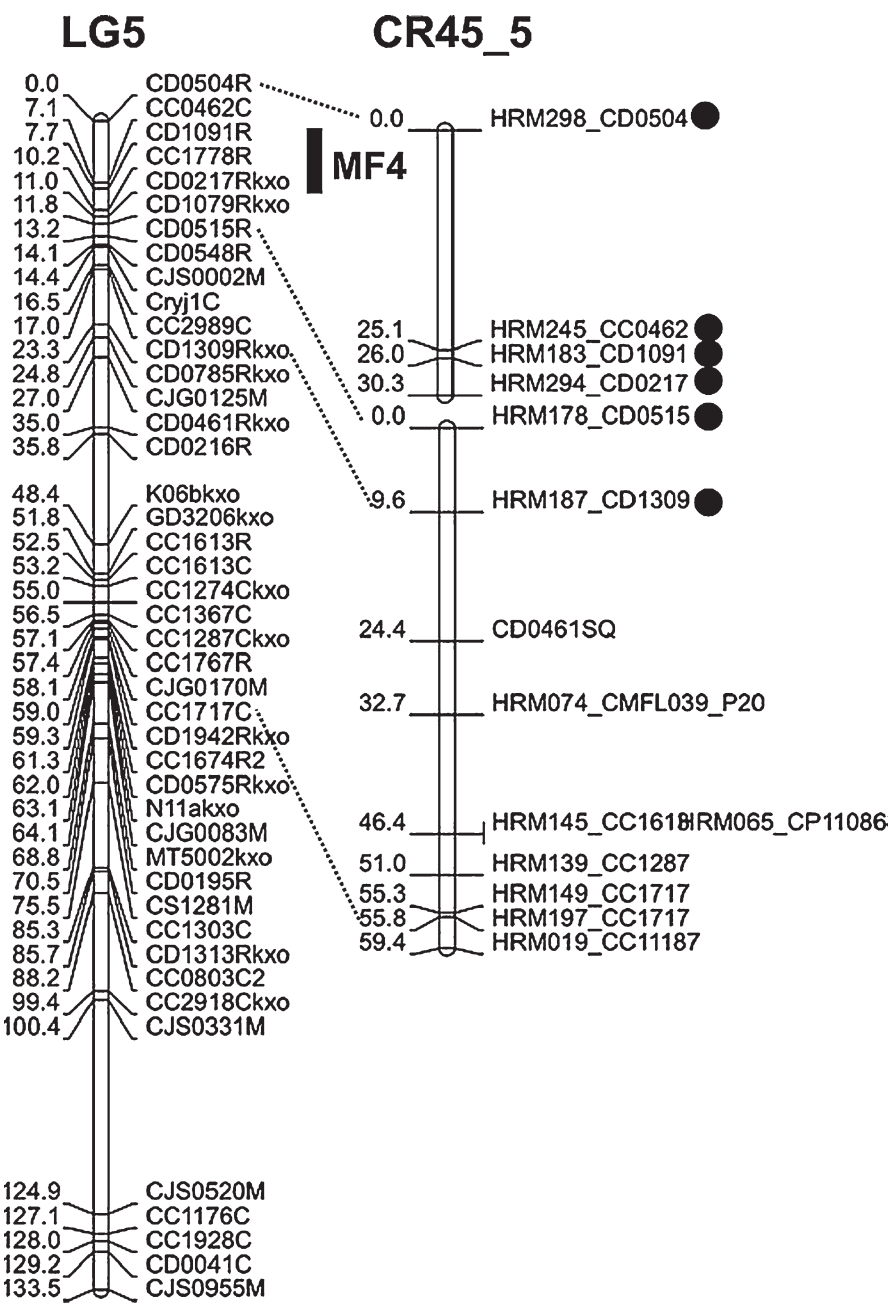


Fig. 11.15 Markers indicated by closed circles on the CR45-5 map of Ujino-Ihara et al. (2012) were significantly associated with male strobili production in *Cryptomeria japonica* as was a QTL for male strobili production (MF4) discovered by Yoshimaru et al. (1998) in the same region. (From Ujino-Ihara et al. 2012)

Table 11.10 SNP locus annotations and significance values for wood property and quantity of male strobili in *Cryptomeria japonica*

Trait	SNP locus	SNP	SNP type	Chromosome position	*P* value	FDR Q value	Marker effect	MAF	Annotation
wood property	gSNP01986	[A/T]	5'UTR	–	8.81 e-05	0.091	0.047	0.334	microtubule-associated protein RP/EB family member 1-like
	gSNP04252	[T/C]	synonymous	5	1.89 e-03	0.976	0.030	0.175	lecithin cholesterol acyltransferase-like protein
	gSNP03140	[A/C]	5'UTR	11	4.43 e-03	1.000	0.025	0.269	RNA-binding protein 40-like
	gSNP01022	[A/G]	synonymous	10	4.51 e-03	1.000	0.025	0.269	cytochrome P450
	gSNP0l196	[T/C]	intron	–	9.98 e-03	1.000	0.020	0.026	fatty acyl-CoA reductase
male strobili	gSNP00856	[T/C]	synonymous	–	2.73 e-03	1.000	0.022	0.155	CLIP-associating protein

From Uchiyama et al. (2013)
FDR false discovery rate, *MAF* minor allele frequency

on phenotypic segregation but had not yet been placed onto genetic maps. Moriguchi et al. (2012) mapped one locus, *ms1*, and then mapped three more loci, *ms-2*, *ms-3,* and *ms-4*, which all mapped to different linkage groups (Moriguchi et al. 2016). The next step in the progression of this line of research will certainly be the positional cloning of these genes which will be greatly facilitated once the *C. japonica* genome has been sequenced.

One association study has been completed to date in *C. japonica*. Uchiyama et al. (2013) associated 1032 SNPs from 1032 candidate genes with several wood-property traits and the quantity of male strobili (Table 11.10). In summary, Japanese researchers have made excellent progress toward developing knowledge and technology toward controlling the abundance of allergenic pollen affecting their human population.

Summary

Complex trait dissection studies have been an active and important area of research in conifer genetics since the early 1990s and continues to this day. The original work was motivated by the need to develop marker × trait associations that could be used in marker-based breeding programs. This work continues for a small number of commercially important species. What has developed more recently is studies in noncommercial species where marker-by-trait associations might be used in conservation and restoration programs. The initial approach used for complex trait dissection was QTL mapping in segregating populations from defined pedigrees. This approach was successful in many species for a variety of traits. Limitations were generally related to the lack of replication across pedigrees, years, and environments, largely due to the significant cost of doing such studies. However, in a few cases, adequate replication and validation was employed and it is from these studies that there can be most confidence in the results. Validation across species was much less frequently done, largely due to the absence of comparative maps and genome sequences. These resources are increasingly available and should be applied as resources allow. What remain absent are good comparative genomic bioinformatics tools to conduct such analyses quickly and with precision. The other great limitation of the QTL mapping approach was the resolution of the mapping of the QTL. This limitation has now been largely overcome by large association studies. The limitation of association studies has been the density of the genome search space due to the lack of reference genome sequences, resequencing databases, and low-cost SNP genotyping technologies. These limitations are now also slowly being overcome. What then lies ahead is the opportunity to conduct GWAS and environmental association studies for many species and many traits. The limitation going into the future will likely be generating resources to grow and maintain populations and to conduct the phenotyping.

Landscape Genomics

12

A Short History of Landscape Genomics Studies in Conifers

We have noted several times in this volume the importance of understanding how trees are adapted to their environments. Conifers are found across diverse environments that include extremes for moisture, temperature, soil types, available sunlight, and so on. Not only do we find different species adapted to diverse environments but also within many species we see adaptation to different environments. Having a deep understanding of the genetic basis of adaptation is important for successful reforestation after harvesting, for conservation and restoration programs, and for potentially coping with climate-induced species range changes.

In Chap. 8, we discussed how the phenotypic component of variation (Var P) varies across diverse environments and also how the genetic component of phenotypic variation (Var G) can be indirectly estimated and distinguished from the environmental component of variation (Var E) in common garden tests. Common garden testing of all kinds has been used extensively in conifers, although it has been applied primarily to a small number of commercially important species because of available resources for this rather expensive and long-term activity. For the vast number of conifer species there is little or no information at all about the genetic component of adaptation to the environment.

In Chap. 9, we discussed the use of neutral genetic markers to, in part, understand the partitioning of genetic variation across landscapes, but we noted that neutral markers only provide insight into demographic processes (mating system, drift, and gene flow) and do not provide information about adaptive processes (selection). In Chap. 10, we introduced nonneutral genetic markers and tests that can be applied to determine if a genetic marker is in fact nonneutral and thus informative for understanding adaptive processes in populations. In Chap. 11, we discussed the topic of complex trait dissection (QTL mapping and association studies), whereby complex adaptive traits, both biotic and abiotic, could be broken down to their individual genetic components. What we have not done is discuss the relationship between the patterning of nonneutral variation in genes and variation in the environment. This

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_12

would be the final step in understanding the relationship between the variation in genotype (Var G) and the variation in the environment (Var E) and is the focus of this chapter.

Population genetic approaches have long been used with neutral genetic markers to understand the partitioning and distribution of genetic variation across landscapes (Chap. 9). These same population genetic approaches could also be used with non-neutral genetic markers. In fact, the F_{ST} outlier approach introduced in Chap. 10 has been the primary population genetic method for testing for a relationship between Var G and Var E. However, this approach requires an a priori assignment to subpopulations of all individuals across the landscape. These subpopulations might be classified as being different for an environmental variable (i.e., wet sites versus dry sites), but in natural populations, these subpopulations might vary for a whole suite of environmental variables in complex ways. Traditional population genetic approaches might provide some evidence for adaptation to the environment, but it is less clear which environmental factors are providing the selective forces for adaptation.

In 2003, Manel et al. (2003) introduced the concept and discipline of *landscape genetics* whereby a direct relationship between genetic variation and environmental variation of one or more types could be directly established. Manel et al. (2003) defined landscape genetics as the combined application of the disciplines of landscape ecology and population genetics. Forest geneticists were quick to appreciate the power of this approach and many background papers were written on the potential for using this approach in forest trees (Sork et al. 2013; Wheeler and Neale 2014; Bragg et al. 2015). Sork et al. (2013) extended the definition of *landscape genetics* first proposed by Manel et al. (2003) to be primarily the study of demographic processes using neutral markers versus *landscape genomics* being the simultaneous study of both demographic and selective processes using genome-wide marker data, thus including both neutral and nonneutral markers. Furthermore, Sork et al. (2013) tried to illustrate how the relationships between the *genotype*, the *phenotype*, and the *environment* can be established using the disciplines *ecological genomics*, *ecological genetics*, and *landscape genomics* (Fig. 12.1).

The basic workflow of a landscape genomics study (sometimes called *environmental association analysis*, EAA) is to sample a large number of geo-referenced trees across a landscape, perform candidate gene or genome-wide SNP genotyping of all samples, obtain environmental data for geo-reference locations either directly or from GIS databases, and then perform an association analysis between the genotypic and environmental data (Fig. 12.2). Rellstab et al. (2015) have written a review that describes the advantages and disadvantages of (1) sampling designs, (2) environmental data, (3) genotypic data, and (4) methods of analysis. These topics will not be discussed here; rather, we refer readers to this excellent review. In the following sections of this chapter, we will review the literature in conifer landscape genomics studies organized around sections on related species (Table 12.1). A short review was published in 2016 (Ćalić et al. 2016) on all forest tree landscape genomics studies published up until that time. The overall summary from that review was that forest landscape genomics offers enormous potential for rapidly, and

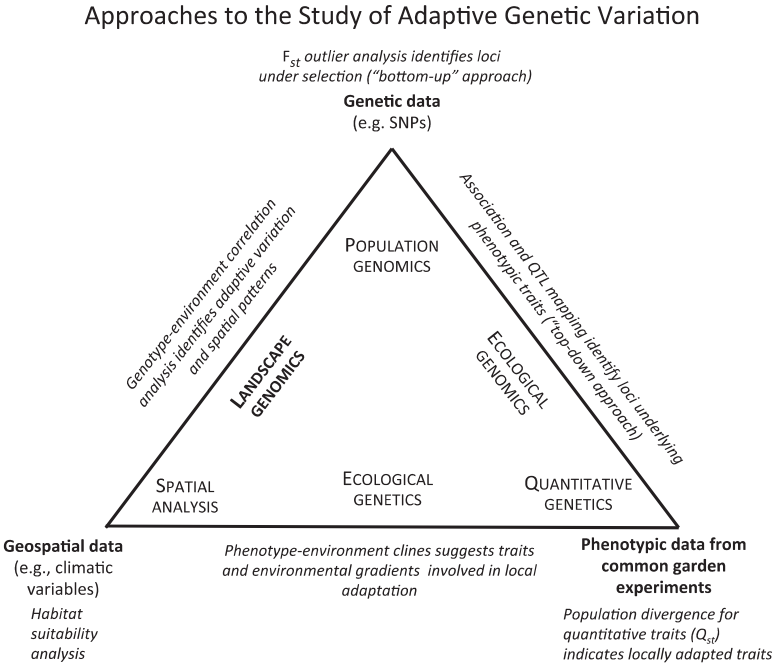


Fig. 12.1 Schematic diagram of how phenotypic, genomic, and geospatial data are combined and analyzed to identify adaptive genetic variation. Disciplinary research areas are in CAPS. Analytical methods named along sides of triangles combine two data types; methods named at triangle vertices involve analysis of just one data type. (From Sork et al. 2013)

increasingly quite inexpensively, inferring the relationship between genotypic and environmental variation across the landscape but that this discipline is still in its infancy (Table 12.2). There are simply not enough results currently to make any broad generalization as to the classes of genes and biochemical pathways most associated with specific environmental variables. As reference genome sequences, genome-wide resequencing, and large SNP databases all become available, the resolution of these studies will improve and gradually most, if not all, of the genetic loci controlling adaptation to the environment will be discovered. Once the discovery reaches a mature state, managers of forests affected by changing climate can begin using this tool.

Pinus Subgenus *Pinus*

The first landscape genomics studies to be done in conifers, in fact in all forest trees, were done in *P. taeda* (Eckert et al. 2010b; Table 12.1). In this study, 907 trees from across the natural range of *P. taeda* were genotyped for 3059 SNPs, one each from 3059 genes. These genes were assumed to be a random set from the genome. The environmental variable tested in the association analysis was an aridity index (AI)

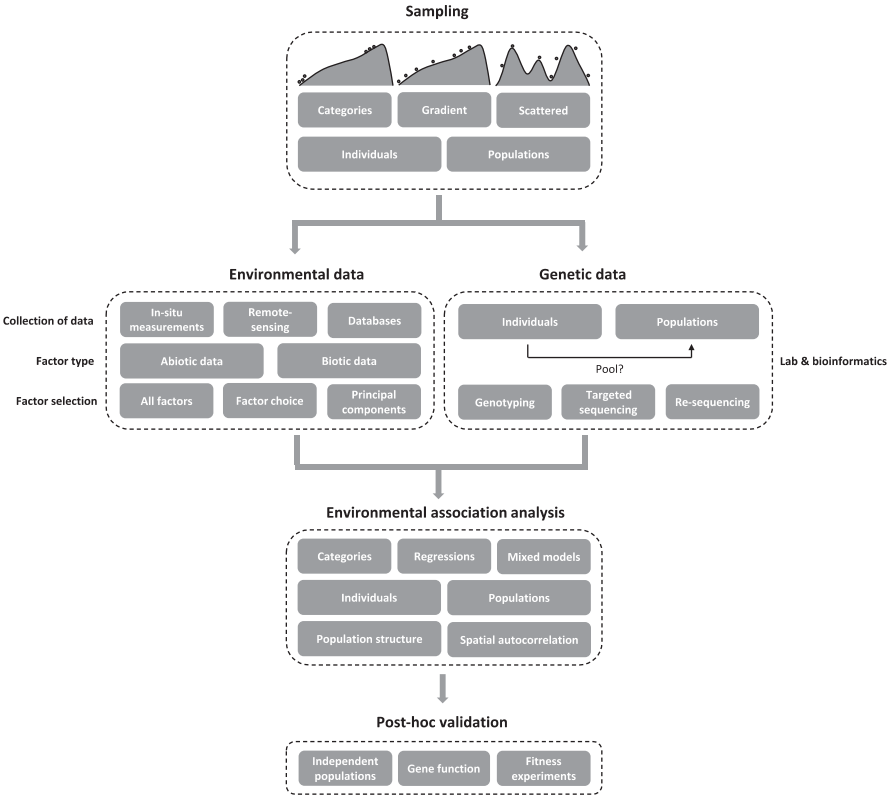


Fig. 12.2 A typical workflow in environmental association analysis (EAA). The three most important options per step are horizontally aligned. Genetic and environmental data are collected at the same sampling locations, processed separately, and jointly analyzed in EAA. The results can subsequently be validated with complementary approaches. (From Rellstab et al. 2015)

which was generated from measures of temperature and moisture taken from GIS databases. The test for association was essentially the same as that used in genotype x phenotype associations (Chap. 11), except AI was substituted for phenotype. Five loci were shown to be associated with AI (Fig. 12.3). All five loci have been functionally implicated in stress responses in *Arabidopsis thaliana.* This seminal landscape genomics study in forest trees provided great enthusiasm that the landscape genomics approach would discover individual gene's underlying adaptation to the environment in tree populations.

In a follow-up study, Eckert et al. (2010a) genotyped 682 trees from 54 populations for 1730 SNPs from 1730 genes. The population samples again covered the natural range of *P. taeda* and the genes were in general a random set. Maximum and minimum temperature and precipitation values for geo-referred points were obtained from GIS databases, and from these data measures of temperature, growing degree days, precipitation, and aridity were generated. Furthermore, all the environmental data were subjected to principal component analysis (PCA) to generate

Table 12.1 Landscape genomics studies conducted in conifers through 2017

Species	No. of pops. *a priori*	No. of pops. *a posteriori*	No. of samples	No. of genes	No. of SNPs	Sig. associations (%)	References
Abies alba	37	2	1183	155	249	3 (1.2)	Mosca et al. (2012b)
Abies alba	36	2	1108	150	231	4 (1.7)	Mosca et al. (2014)
Larix decidua	27	2	935	151	267	5 (1.9)	Mosca et al. (2012b)
Larix decidua	22	2	824	151	233	16 (6.9)	Mosca et al. (2014)
Picea abies (macro scale)	27	4	546	247	237	38 (16.0)	Scalfi et al. (2014)
Picea abies (micro scale)	12	0	300	224	226	2 (0.8)	Scalfi et al. (2014)
Picea abies	24	2	826	285	288	10 (3.5)	Di Pierro et al. (2016)
Picea abies		2	392	240	384	3 (0.8)	Ćalić et al. (2016)
Picea glauca	2	1	22	290	384	20 (0.05)	De La Torre et al. (2014a)
Picea engelmannii	3	3	40	290	384	20 (0.05)	De La Torre et al. (2014a)
Picea hybrid	5	5	756	290	384	20 (0.05)	De La Torre et al. (2014a)
Picea hybrid	254		566	24,644	810,802	3706 (0.46)	Yeaman et al. (2016)
Pinus albicaulis	8		244	unknown	116,231	many, polygenetic	Lind et al. (2017)
Pinus cembra	24	2	860	260	459	11 (2.4)	Mosca et al. (2012b)
Pinus cembra	18	2	673	265	455	20 (4.4)	Mosca et al. (2016)
Pinus contorta	281		620	24,388	547,872	3790 (0.69)	Yeaman et al. (2016)
Pinus lambertiana	10		241		475	31 (6.5)	Eckert et al. (2015)
Pinus lambertiana	13	2	313		186	9 (4.8)	Vangestel et al. (2016)
Pinus monticola	61		348	127	158	1 (0.6)	Nadeau et al. (2016)
Pinus mugo	27	3	935	380	693	14 (2.0)	Mosca et al. (2012b)
Pinus mugo	20	3	678	383	663	20 (5.2)	Mosca et al. (2016)
Pinus strobus	133		831	120	153	26 (17.9)	Nadeau et al. (2016)
Pinus strobus	29		638	25	44	2 (4.5)	Rajora et al. (2016)
Pinus taeda	54	3	682	1730	1730	394 (22.8)	Eckert et al. (2010a)
Pinus taeda	167	3	907	3059	3059	5 (0.16)	Eckert et al. (2010b)

The very small number of significant associations reported might be a function of the power of experimental designs but is also suggestive of the small proportion of genes that may underlie adaptation to the environment

Table 12.2 Genes associated with the same or similar environmental factors in more than one independent landscape genomic study

Gene product	Associated variable	Species	References
ABC transporter	TD	*Picea glauca, P. engelmannii*	De La Torre et al. (2014a)
Flavodoxin family protein	AI	*Picea glauca, Pinus taeda*	De La Torre et al. (2014a), Eckert et al. (2010b)
Glycosyl hydrolase family protein	MWMT, EXT	*Picea glauca*	De La Torre et al. (2014a)
	T (MAT, TS, MTDQ, MTCQ)	*Picea abies*	Scalfi et al. (2014)
Heat shock protein 101	P	*Pinus mugo, P. cembra*	Mosca et al. (2012b)
No apical meristem (NAM) protein	MWtQT	*Picea abies*	Di Pierro et al. (2016)
	MWMT	*Picea glauca, P. engelmannii*	De La Torre et al. (2014a)
	MWtQT, MDQT	*Picea abies*	Di Pierro et al. (2016)
Putative circadian clock genes	LT	*Picea abies*	Chen et al. (2012b)
Ubiquitin specific protease	LG	*Pinus taeda*	Eckert et al. (2010b)
Vernalization family proteins	T	*Picea abies*	Scalfi et al. (2014)
4-coumarate: CoA ligase	MTCM, TS	*Pinus pinaster, P. halepensis*	Grivet et al. (2010)

From Ćalić et al. (2016)
AI aridity index, *EXT* extreme maximum temperature over the 30-year period, *LG* longitude, *LT* latitude, *MAT* mean annual temperature, *MCMT* mean coldest month temperature, *MDQT* mean temperature of driest quarter, *MTCM* minimum temperature at the coldest month, *MWMT* mean warmest month temperature, *MWtQT* mean temperature of wettest quarter, *T* temperature, *TD* continentality, *TS* temperature seasonality, *P* precipitation, *WT* temperature during the warmest month

environmental principal components (PCs). The EEA was done using the Bayesian approach developed by Coop et al. (2010) and implemented in the software *Bayenv*. The authors found that 22.8% of the SNPs were associated with one or more environmental factors (Fig. 12.4). The functional annotations of the associated genes were identified to get a preliminary idea of the genes and pathways that might underlie adaptation to these environmental variables (Table 12.3). Many of these genes had also been previously shown to be involved in stress-related responses in *Arabidopsis thaliana.*

The other two species of subgenus *Pinus* having landscape genomics studies to date are *P. mugo* in the European Alps and *P. contorta* in North America. Mosca et al. (2012b) sampled 935 *P. mugo* trees from 27 populations across the Italian Alps and genotyped these trees for a large number of genes and SNPs. A suite of moisture and temperature variables were obtained from GIS databases and the

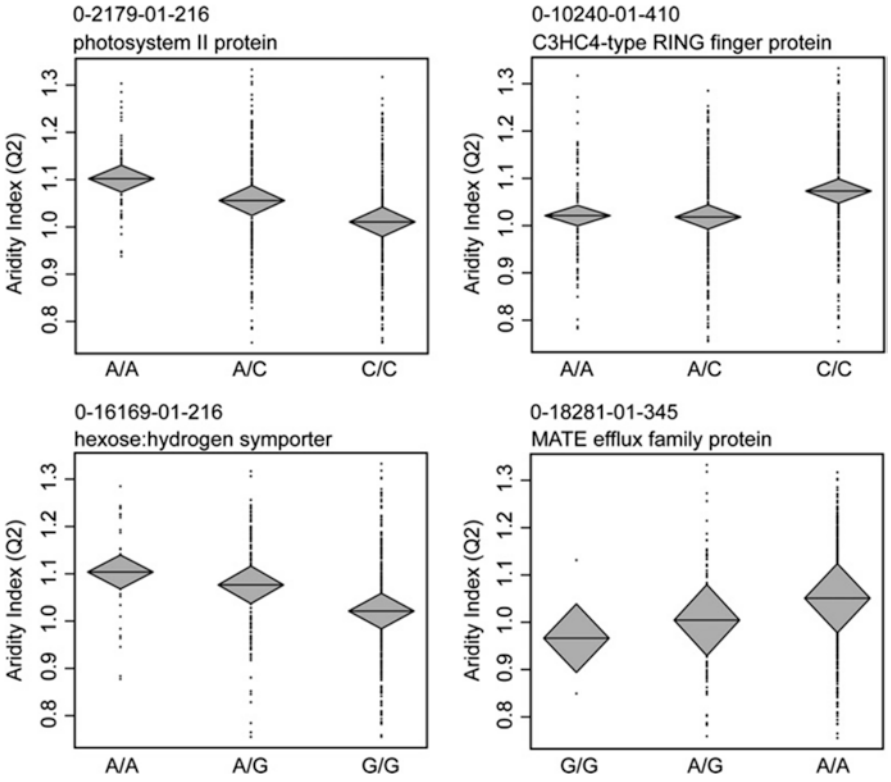


Fig. 12.3 Genotypic associations in *Pinus taeda* for four loci (the fifth locus is not shown) with the aridity index during the second quarter (April–June) illustrate small yet significant correlations. Horizontal lines within diamonds denote the genotypic mean, with diamonds representing the upper and lower bounds of the 95% confidence interval for the mean. Ancestry-corrected environmental indices are plotted on the *y* axis. For clarity, the SNP genotypes are plotted on the *x* axis, although ancestry-corrected SNP genotypes are used in the association analysis. (From Eckert et al. 2010b)

dimensionality reduced by PCA. Using *Bayenv* analysis, 2.0% of the SNPs were associated with one or more of the PCs (Table 12.4). Several of these associated genes had been implicated in stress response in *Arabidopsis thaliana* including those determining cell wall and membrane properties. In a second study using a subset of the populations and trees that formed elevational gradients and nearly the same SNPs, Mosca et al. (2016) used two EEA approaches, *Bayenv2* (Gunther and Coop 2013) and the latent factor mixed model (*LFMM*) approach (Frichot et al. 2013) to detect associations to a suite of moisture and temperature variables across the elevational gradients. Using these more advanced methods and the elevational gradient design, more associations were found (5.2%, Table 12.5), with a higher percentage associated with temperature than with precipitation. These results point to the classes of genes that may be important to follow closely as montane tree populations adapt to changing climate.

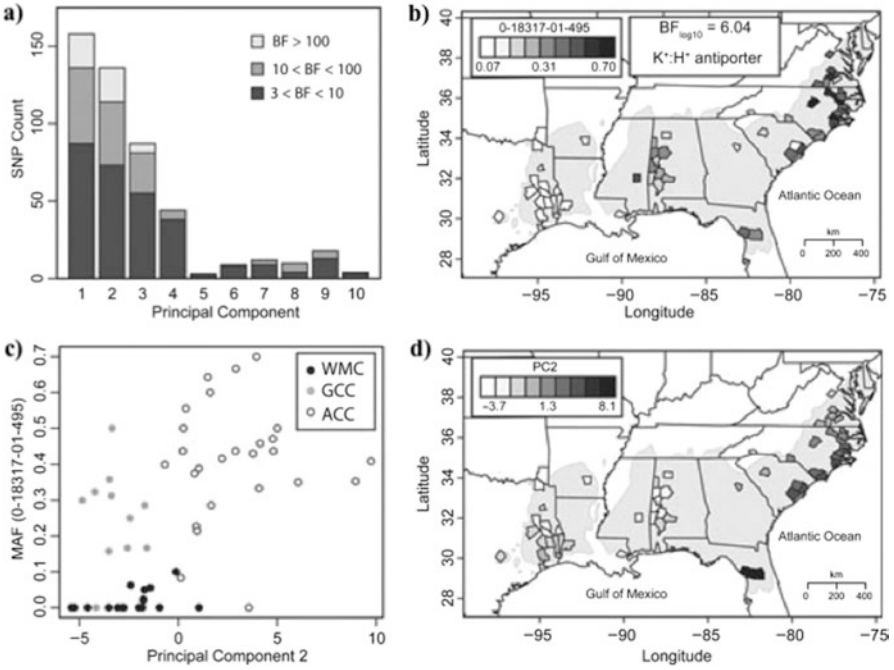


Fig. 12.4 Application of Bayesian geographical analysis discovered 441 supported climate associations representing 394 unique SNPs. (**a**) The number of supported SNPs per geoclimatic PC grouped by category of support; (**b**) the geographical distribution of the minor allele frequency (MAF) at a SNP located in locus 0–18,317; (**c**) a plot of the MAF at SNP 0–18,317-01-495 and PC2. Shading denotes the three regions (i.e., genetic clusters) identified previously for loblolly pine—WMC: west of the Mississippi cluster, GCC: Gulf Coast cluster, and ACC: Atlantic Coast cluster; (**d**) the geographical distribution for the scores on PC2. This component is largely composed of spring and summer precipitation, summer and fall aridity, and longitude. (From Eckert et al. 2010a)

The study involving *P. contorta* was a very large landscape genomics study using more than one million SNPs to genotype trees from 281 populations in the Canadian provinces British Columbia and Alberta and associate them with 22 climate variables (Yeaman et al. 2016). The primary environmental associations were with temperature. A parallel study was conducted with *Picea glauca*, *P. engelmannii*, and their hybrids, and very similar environmental and phenotypic associations with genotypes were found (Yeaman et al. 2016). The *P. contorta* study is significant because it provides the first evidence for convergent local adaptation in conifers. Previous studies have argued for independent (nonconvergent) evolution. However, as Ćalić et al. (2016) argued, these earlier studies generally had insufficient experimental design and power to make such claims.

Table 12.3 Genes in *Pinus taeda* positively associated with environmental variables

PC no. and description	BF[a]	Annotation
1. Latitude, longitude, temperature, GDD5, winter aridity	3.32	Hypothetical protein (Atlg01500)
	3.3	Hypothetical protein (At4g24090)
	3.29	Ca^{+2} dependent kinase (Atlg05410)
2. Longitude, spring-fall aridity, precipitation	6.04	K^{+}:H^{+} antiporter (At2gl9600)
	5.28	TIFY domain-containing protein (At4g32570)
	5.1	BAG protein (At5g07220)
3. Winter and summer precipitation, summer aridity	3.71	Dehydratase-like protein (Atlg76150)
	3.37	Hypothetical protein (At3gl2650)
	2.95	Thioredoxin-related protein (XM_002283585)
4. Spring and fall precipitation and aridity	1.64	PTAC2-like protein (Atlg74850)
	1.56	Hypothetical protein (At4gl0430)
	1.37	LIM transcription factor (Atlgl0200)
5. Winter aridity and GDD5	1.18	Hypothetical protein (AM432844)
	0.53	PPR protein (At4g02750)
	0.48	Histone 2A protein (Atlg54690)

From Eckert et al. (2010a)
[a]Bayes factors (BFs) are given on a $\log_{10}$ scale

Pinus Subgenus *Strobus*

Many EAA studies in species of the subgenus *Strobus* have been completed, likely due to the interest in this taxonomic group related to changing climate. The first of these was done by Mosca et al. (2012b) in *P. cembra.* They genotyped 860 trees from 24 populations across the Italian Alps for 459 SNPs from 260 genes. Climate data (temperature and precipitation) were reduced to a small number of PCs. A small number of genes (2.4%) were associated with four different PCs (Table 12.6); however, most were associated with PC1 which was a function of temperature. Thus, at this early phase of landscape genomics in conifers, a pattern emerges that northern latitude and high-elevation conifers are adapting through genes related to temperature whereas more southern latitude and low-elevation conifers are adapting through genes related to moisture stress. This observation is consistent with results obtained from common garden studies. In the follow-up study in *P. cembra*, Mosca et al. (2016) reanalyzed these SNP data with populations along an elevational gradient and found genes associated with both temperature and moisture; however, this time there were a greater number associated with moisture.

Two studies have been done in *P. lambertiana* in California, one at the micro-geographical scale, while the other on a macro-geographical scale. Eckert et al. (2015) sampled 241 trees from 10 populations around Lake Tahoe in the US states of California and Nevada and genotyped each of the trees for 475 SNPs from an

Table 12.4 Environmental principle components (PC*s*) associated (Bayes factor values) with genes in *Pinus mugo* in the Italian Alps

PC1	PC2	PC3	PC4	PC5	PC6	Putative function[a]
3.87	11.6					Polygalacturonase (QRT3)
					4.3	Heat shock protein 101
10.1						NA
213						NA
					16.8	NA
10.2						NA
1230						Protein kinase family protein
16.4	3.46					F-box family protein (FBW2)
	41.2					Tetratricopeptide repeat domain-containing protein
88.7						Cinnamoyl-CoA reductase (CCR1)
		29.9	18.6			AFC1–ATP binding/kinase/protein
			126			Hexokinase (HXK1)
456,000						Putative beta-galactosidase
	49.1					Imbibition protein

From Mosca et al. (2012b)

[a]*NA* no annotation

Table 12.5 Loci highly (Bayes factor > 10) associated with climatic variables (T_{m_y}, annual mean temperature; $T_{min,\ Q1,2}$ minimum temperature for the first/second quarter; $T_{max,\ Q2,3}$ maximum temperature for the second/third quarter) or elevation (Ele) according to *Bayenv2* simulation in *Pinus mugo*

	All sites					Elevational				
Protein	Ele	$T_{min\ Q1}$	$T_{min\ Q2}$	$T_{max\ Q3}$	T_{m_y}	Ele	$T_{min\ Q1}$	$T_{min\ Q2}$	$T_{max\ Q3}$	T_{m_y}
Tetratricopeptide repeat (TPR)- containing protein										4
Ring zinc finger protein					3.4					
Glucosidase	4.3	3.4	111.3	3.8	19.3	10.9		44.3		18.8
Glycosyl transferase family protein					3			6.3		12.6
Unknown protein	9	3.7	19.2		16.2	19.3	4.2	76.8		15.4
Heat shock factor protein, HSF4								4.2		11.2
NA	3.7		5.4	3.2	43			3.1		8.9
NA	9.1		10.7		171.8	6.5		35.3	3.2	59.8
HAT3 transcription factor	9.1				13.9	5.2		13.4		4.8
DEAD/DEAH box helicase	3.2					7.8	3.5	27.6		6.5
MAP kinase (MPK16)	10.8	3.4	14.4		18.9	23.7	4.9	110.3		20.7
NA	3	3.7	103		6.7	5.8		17.6		6
NA								7.3		15.7
qrt3 (quartet 3)/polygalacturonase (QRT3)			7.7		6.3			10.8		6.7
Beta-galactosidase 8						10.4		32.4		11.3
1-Aminocyclopropane-1-carboxylate synthase			12		13.8					
Imbibition protein	9.8	12.6	635.4	4.6	30.8	22.2	5.7	177.1		27.7
Endonuclease exonuclease phosphatase family protein	4.1	3.1	37.5		8.9					
NA	3.5				15					
Kinesin light chain related			15.9		3					

From Mosca et al. (2016)

Table 12.6 Environmental association analysis in *Pinus cembra* from the Italian Alps

PC1	PC2	PC3	PC4	Putative function	Code
31				Arabidopsis non-ATPase subunit 9	NA
36.9				Structural constituent of ribosome (P40)	NS
		13.2		Serine-type endopeptidase	NA
29.2				MYB61 transcription factor myb8	NS
			7.63	Heat shock protein 101	NA
			10.4	Splicing factor	NC
	70.2			Protein serine/threonine phosphatase (PP2A-3)	NC
167				Hypothetical protein	NA
21				Hypothetical protein	NA
301				Carbohydrate transmembrane transporter	NA
281				60S acidic ribosomal protein P0	NS

Several SNPs had a moderate to strong support (10 < BF < 100) for association with the first four climatic PCs. From Mosca et al. (2012b)
SNP single-nucleotide polymorphism, *BF* Bayes factor, *NA* no annotation, *NS* nonsynonymous, *SY* synonymous, *NC* noncoding

equal number of genes. SNPs were then associated with PCs that were generated from four environmental variables: precipitation, temperature, available water supply, and percent sand (Fig. 12.5). Each of the PCs were a function of multiple variables and could not easily be classified as a "moisture" PC or a "temperature" PC. Nevertheless, a significant number (6.5%) of associations with the PCs were found. This study was one of the first to demonstrate genotype by environment associations occurring at a very local scale. In the second study done at the macrogeographical scale, Vangestel et al. (2016) also found a few candidate genes associated with environmental PCs, but most notably a gene also found by Eckert et al. (2015) annotated as *mitochondrial import inner membrane translocase subunit TIM14 precursor* in *Arabidopsis thaliana* which has been implicated in drought-stress response. Therefore, even though these studies are still quite preliminary, the repeated detection of a few genes across studies supports the parallel evolution model of genetic adaptation to the environment.

A comparative environmental association study was conducted in two closely related species, *P. strobus* from eastern North America and *P. monticola* from western North America (Nadeau et al. 2016). Many trees (831 and 348) were genotyped for 153 and 158 SNPs in *P. strobus* and *P. monticola*, respectively. They tested for association under the *isolation by environment* (IBE) model while controlling for *isolation by distance* (IBD) and *isolation by colonization* (IBC) using both *Bayenv2* and *LFMM*. They found five significant SNPs in *P. strobus* but just one in *P. monticola* (Table 12.7). All but one of these genes is thought to be a candidate gene for growth and phenology. This study was comprehensive in terms of population sampling and methods for testing for environmental associations but again very limited in terms of the genome search space. Rajora et al. (2016) have also searched for environmental associations in *P. strobus.* They sampled many trees (638) across 29 populations but used only 44 SNPs from 25 candidate genes thought to be involved

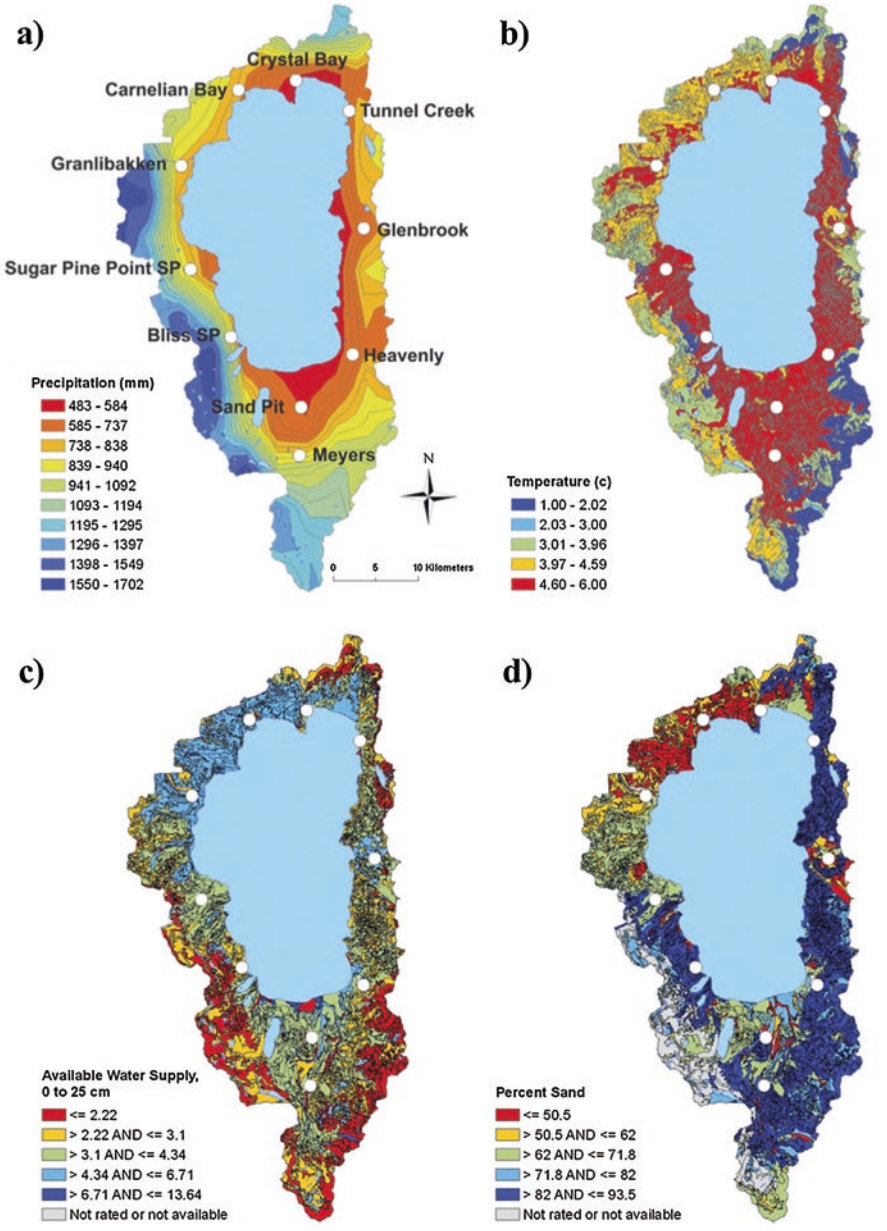


Fig. 12.5 Distribution of *Pinus lambertiana* sampling locations (white circles) and environmental gradients in (**a**) annual precipitation, (**b**) average annual temperature, (**c**) available water capacity at 25 cm, and (**d**) relative sand content. (From Eckert et al. 2015)

Table 12.7 Environmental association analysis in *Pinus strobus* and *P. monticola* in Canada

SNP	Gene	*Bayescan*	*Bayenv2*	LFMM	SNP annotation	Putative gene function	Candidate for growth/ phenology in *Picea glauca*
Pinus strobus							
N-029	0_6047_02	div***	DD5, TD, bFFP, eFFP, PAS, CMD****	DD5, bFFP, eFFP, PAS, TD, CMD, MSP, Elev****	Na	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein	No
G-014	GQ0081. BR.1_D09	div***	DD5, MSP, PAS, Elev**	DD5*	NS	Plas d movement impaired1-related1 (PMIR1); plant-speci c C2 domain containing gene family	No
M-015 0_8683_01	0_8683_01	Ns	DD5, bFFP, PAS, CMD****	PAS, CMD, DD5, bFFP***	NS	Serine–threonine-protein kinase at1g18390-like[b]	Yes
M-016 0_8683_01	0_8683_01	Ns	TD*	CMD, PAS*	NS	Serine–threonine-protein kinase at1g18390-like	Yes
M-017	0_8844_01	Ns	eFFP, bFFP, DD5**	DD5*	Intron	Galacturonosyltransferase 13-like	Yes
Pinus monticola							
S-007	CL3539-Con g1_01	div*	Elev*	Ns	Intron	TOM1-like protein 2	Yes

From Nadeau et al. (2016)

ns nonsignificant, *S* synonymous SNP, *NS* nonsynonymous SNP, *na* not annotated (no blast hit). BayeScan and LFMM: *$q < 0.05$; **$q < 0.01$; ***$q < 0.001$, ****$q < 0.0001$. *Bayenv2*: *BF > 3; **BF > 10; ***BF > 32; ****BF > 100

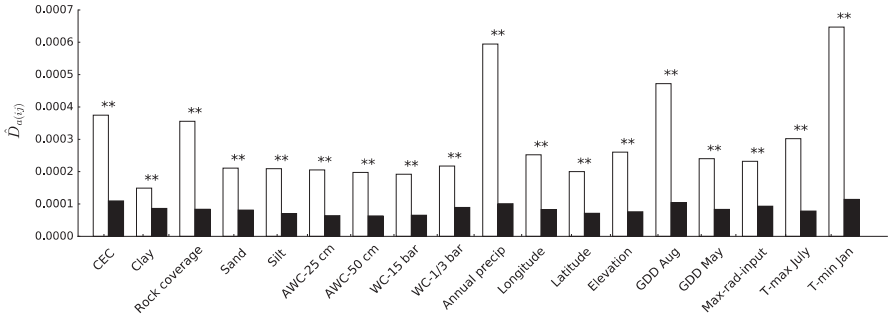


Fig. 12.6 Environmental association analysis in *Pinus albicaulis* in California. Allele frequency covariance ($\hat{D}_{a(ij)}$) among loci associated with environment by *Bayenv2*. In white are the median values from $\hat{D}_{a(ij)}$ calculated among focal SNPs associated with environment. Black bars display the 95th percentile of the null distribution of median $\hat{D}_{a(ij)}$. Environmental variables are grouped by those related to soil (cation exchange capacity (CEC) through water capacity (WC)-1/3 bar) and those related to either climate or geography (annual precipitation through T-min Jan), with variables related to water availability grouped together in the center of the figure (available water capacity (AWC)-25 cm through annual precipitation). All sets of focal loci had median $\hat{D}_{a(ij)}$ greater than the 100th percentile of the null distribution, as indicated by two stars (**). (From Lind et al. 2017)

in adaptation. They tested for single- and multi-locus associations using *redundancy analysis* (RDA). The single-locus RDA identified two SNPs associated with environmental PCs.

The final landscape genomics study to date in subspecies *Strobus* was in *P. albicaulis*. The sampled populations were at the same locations as the *P. lambertiana* study of Eckert et al. (2015) that was discussed earlier in this section. A large number (244) of individual trees were genotyped but this time with a *genotyping by sequencing* (GBS) method called *ddRadseq* that produced 116,231 SNPs (Lind et al. 2017). A comprehensive list of environmental factors was used in a *Bayenv2* analysis. Because of the large number of SNPs tested, they calculated another statistic, the allele frequency covariance, to show the polygenic relationship between the genotype and the environment (Fig. 12.6). They showed that this covariance structure tracks patterns of water availability among these local populations and that this environmental factor is primarily responsible for the patterns of local adaptation observed.

Picea

In contrast to the large number of studies in *Picea* that have used neutrality and outlier tests (Chap. 10), there are relatively few landscape genomics studies. As part of a study focused on studying admixture in hybrid populations of *P. glauca* and *P.*

engelmannii, De La Torre et al. (2014a) also conducted an environmental association analysis (*Bayenv*) just in the hybrid population and a second analysis that included the parental species. They found that 5% of the SNPs had an association with one or more environmental factors related to temperature or moisture and a few of these were also detected by outlier analysis (Table 12.8). As we noted earlier in this chapter, Yeaman et al. (2016) also conducted a very large environmental association study in hybrid populations of *P. glauca* and *P. engelmannii* in Canada and found many genes with signatures of adaptation. However, it should be noted that their study was focused on detecting signatures of convergence between the studied species of *Pinus* and those of *Picea*, so there was no correction for population structure within each species. Thus, many of the putatively locally adapted genes with signatures in only one of the two species are likely false positives.

Picea abies is the other *Picea* species for which a few landscape genomics studies have been done (Scalfi et al. 2014). In this study, investigators searched for environmental associations at both a micro-geographical (short elevational transect) scale and a macro-geographical scale (European Alps) using essentially the same set of SNPs. A regression approach was used to test for environmental associations using several different models that were functions of moisture and temperature variables. They found only two (0.8%) significant associations in the micro-geographical population but 38 (16.0%) associations in the macro-geographical populations. F_{ST} values among micro-geographical populations were insignificant, whereas among macro-geographical populations, many were significant. These results would suggest that detecting differences in environmental association at a very local scale might be difficult, although the results from the Lind et al. (2017) study in *Pinus albicaulis* suggest this might be possible if genome-wide data are used. Subsequently, Di Pierro et al. (2016) conducted an environmental association study with many *Picea abies* trees (826) and populations (24) across the Italian Alps using a SNP genotyping array with significant overlap with the one used by Scalfi et al. (2014). Ten (3.5%) associations were found with temperature and moisture variables (Table 12.9), one of which was also detected by outlier analysis. However, none were in common with those detected by Scalfi et al. (2014). Finally, Ćalić et al. (2016) also did an environmental association study in the Italian Alps with *P. abies* populations differently used by Di Pierro et al. (2016) but very similar genes/SNPs and environmental variables and found a small number (0.8%) of associations and none in common with the earlier studies. This lack of commonality among similar studies led Ćalić et al. (2016) to call into question the overall power of these first-generation landscape genomics studies and not rush to conclude that there are different genetic causes for local adaptation. The genome-wide and comparative study of Yeaman et al. (2016) argues that convergent local adaptation should be expected, at least for a proportion of the total number of genes driving observed responses.

Chen et al. (2012b, 2014) have done studies in *Picea* species in northern Europe to better understand the genetic basis of clinal variation. They conducted a *Bayenv* analysis with 137 SNPs from 18 candidate genes and some environmental factors (which are unclear) and found several (13.1%) associations, many in genes that had been implicated in adaptation by earlier gene-expression studies. This study is a

Table 12.8 Environmental association analysis in hybrid populations of *Picea glauca* and *P. engelmanni*i in Canada

SNP ID	F_{ST} outlier[a]	Environmental associations[b]	Annotation
208pgl 2875c	Yes	MWMT, SHM, EXT, Eref	Glycoside hydrolase family 28 protein/polygalacturonase (pectinase)
295_78	Yes	MWMT	No apical meristem
WS-2.0-GQ03105. B7–012.3-654	No	MWMT,MSP, SHM, Eref, CMD	Fructose-1,6-bisphosphatase
WS-2.0-GQ0064. B3.r-I13.1–1236	Yes	MWMT	Acid phosphatase
14_248	No	MCMT,TD	ABC transporter
WS-2.0-GQ0041. BR-J07.2–36	No	MAP, AHM, PAS	Unknown
0_13680-contig2. Cl-149	No	MAP, AHM, PAS	Hypothetical protein
WS-2.0-GQ0021. BR.1-G04.1–641	No	MAP, AHM, PAS	Unknown
WS-2.0-GQ0168. B3-N16.1–556	No	SHM, CMD	Flavin reductase
144_441	No	CMD	Phytochrome 4
C2270-contigl. NCI-384	Yes	–	CCAAT-binding transcription factor
C6522-contigl. NCI-269	Yes	–	Unknown
WS-2.0-GQ0024. B3.r-D12.1–239	Yes	–	Peroxisomal membrane protein
69_753	No	MAT,DD_0, DD 18	CBL-interacting protein kinase
68_286	Yes	MWMT, MSP, AHM, SHM, DD5, DD18, EXT,Eref	Glycosyl hydrolase
206_435	No	MCMT,TD,DD_0	Isoflavone reductase
13_496	Yes	MCMT,TD,DD5, bFFP, PAS	FK506-binding protein
45_1067	Yes	MCMT,Eref	Alpha-amylase
288_628	No	MAP, PAS	Late elongated hypocotyl
288_302	No	MAP, PAS	Late elongated hypocotyl

From De La Torre et al. (2014a)

[a]Only outlier loci suggesting diversifying selection in the *BayeScan* analyses were considered

[b]Environmental associations based on *Bayenv* are as follows: Mean annual temperature (MAT), precipitation as snow (PAS), mean warmest month temperature (MWMT), summer heat—moisture index (SHM), continentality (TD), annual heat—moisture index (AHM), mean annual precipitation (MAP), mean summer precipitation (MSP), degree-days below 0 °C (DD_0), mean coldest month temperature (MCMT), Hargreaves reference evaporation (Eref)), Hargreaves climatic moisture deficit (CMD), degree-days above 5 °C (DD5), bFFP (Julian date on which frost-free period starts), degree-days below 18 °C (DD_18), degree-days above 18 °C (DD18), extreme maximum temperature over a 30-year period (EXT)

Table 12.9 Environmental association analysis in *Picea abies* in the Italian Alps. (a) Summary of the single-nucleotide polymorphism (SNP) detected using outlier detection in *BayeScan* (b) and the environmental association analysis in *Bayenv2.0*

(a)

SNP		*BayeScan*			
Full dataset	Reduced dataset	*q* value	Alpha	F_{ST}	Annotation
L070 ns	L070 ns	0.0106; 0.0004	1.2342; 1.3441	0.0888; 0.0477	*Phosphoenolpyruvate carboxykinase*
L033 ns		0.0442	0.9859	0.0737	*Ubiquitin-conjugating enzyme*
	L109	0.0092	1.1211	0.0396	–
	L154	0.0193	1.0289	0.0367	*No apical meristem (NAM) protein*
	L059	0.0396	−1.1769	0.0051	*Poly-adenylate binding protein*

(b)

SNP		*Bayenv*		
Full dataset	Reduced dataset	BF	Climatic variables[a]	Annotation
L015 ns	L015	3 < BF < 10	AP, WtQP, *DQP*; *CQP*	*SNF2 family DNA-dependent ATPase*
L048 ns		3 < BF <11	MWtQT, MDQT, WtQP, CQP	*Manganese-stabilizing protein/ photosystem II polypeptide*
	L065	3 < BF < 10	MWtQT	*Hydrolase activity*
L070 ns	L070	3 < BF < 10	*WmQP*	*Phosphoenolpyruvate carboxykinase*
L126	L126	3 < BF < 10	*WmQP*	–
L154	L154	3 < BF < 12	CQP, *MWtQT*	*No apical meristem (NAM) protein*
	L165	3 < BF < 10	MDQT	–
L172		3 < BF < 10	WmQP, DQP	*Pleckstrin homology domain (PH domain)*
L183		3 < BF < 10	MWtQT	*Plasma membrane protein*
L196	L196	3 < BF < 10	WmQP	*HEAT repeat domain*

From Di Pierro et al. (2016)

[a]*AP* annual precipitation, *WtQP* precipitation of wettest quarter, *DQP* precipitation of driest quarter, *CQP* precipitation of coldest quarter, *MWtQT* mean temperature of wettest quarter, *MDQT* mean temperature of driest quarter, *WmQP* precipitation of warmest quarter

nice example of how multiple omic approaches and traditional common garden studies can be used in parallel to more deeply investigate and validate the genetic basis of adaptation. Chen et al. (2014) essentially repeated the *P. abies* study in *P. obovata* in Siberia to determine if some of the same candidate genes would also be implicated in determining patterns of clinal variation. This turned out to be the case for two well-known genes, *FTL2* and *Gigantea*.

Abies and *Larix*

Landscape genomics studies have been done in one species each of the genera *Abies* and *Larix* reported in two papers by Mosca et al. (2012b, 2014). In the first (Mosca et al. 2012b), a large number of trees (1183 and 935) and populations (37 and 27) were sampled in *Abies alba* and *Larix decidua*, respectively. The trees were genotyped with 249 and 267 SNPs, respectively. Environmental associations with PCs from several moisture and temperature variables were tested with *Bayenv.* The number of significant associations found, however, was much less in these species than was found with *P. cembra* and *P. mugo*, that were reported in the same study with nearly identical experimental parameters (Table 12.10). It might be tempting to conclude that the pines are more narrowly adapted than the spruce and larch in the montane environment. In the follow-up study, Mosca et al. (2014) were the first to use the *isolation by adaptation* (IBA) versus *isolation by distance* (IBD) approach to discover environmental associations in *Abies alba* and *Larix decidua*. This study used a subset of the tree samples and populations and essentially the same SNPs and environmental data. This time, however, slightly more associations were found, providing merit for using this more advanced statistical approach. Furthermore, the associations detected indicate that both species are likely adapted to precipitation whereas *A. alba* might be highly adapted to temperature and soil type.

Table 12.10 Environmental association analyses in *Abies alba* and *Larix decidua* in the Italian Alps

Species	SNP	PC1	PC2	PC3	Putative function	Code
Abies alba	2631301-160			5.57	Unknown protein	NA
	0_7009_01-212		7.56		Kelch repeat-containing F-box family protein	SY
	0_14221_01-394			5.11	Seryl-tRNA synthetase	NC
Larix decidua	2_3113 01-121	22.7			Dimethyladenosine transferase	NA
	0_9284_02-470	103			Phosphate translocator-related	NA
	0_6465_01-217	22.8			Hydroxyproline-rich glycoprotein family protein	NA
	0_17790_01-159	10.8			Beta-glucuronidase	NA
	CL1224Contigl_01-215	27.6			Alpha-N-acetylglucosaminidase family	NA

In each species, several SNPs had a moderate to strong support ($10 < BF < 100$) for association with the first three climatic PCs. From Mosca et al. (2012b)

BF Bayes factor, *SNP* single-nucleotide polymorphism, *NA* no annotation, *SY* synonymous, *NC* noncoding

Summary

Landscape genomics (environmental association analysis) offers enormous potential to discover the underlying genetic causes of adaptation to the environment. Such studies in conifers began only in 2010. The basic approach involves sampling trees across a broad landscape, SNP genotyping these trees for as many genetic loci as practically and economically possible, and then associating the genetic variation with environmental variation obtained from GIS databases. Genetic variation at loci associated with variation in the environment is statistically suggestive of a gene underlying adaptation to the environment. Once associations are identified, it is possible to develop diagnostic tools based on the genetic markers to assist breeders and gene resource managers in developing and managing adapted populations and specifically under rapidly changing climate scenarios. The path to discovery, however, has many challenges. The primary challenge is the scale at which discovery experiments must be performed: (1) large sample sizes, (2) multiple and diverse environments, and (3) genome-wide genotyping. Overcoming these obstacles is largely a financial challenge and will take time. An important fundamental question yet to be resolved is whether evolutionary adaptation is primarily local or parallel. Clearly both mechanisms must occur but the relative contribution of each, and the genes and environmental variation attributable to each, awaits understanding.

Conservation Genetics

13

A Brief Introduction to Conservation Genetics in Forestry

The topic of conservation genetics in forestry dates to nearly the beginning of breeding and genetic research in forest trees. In the early literature, the topic was often labeled *gene conservation*, whereas more recently it is generally labeled as *conservation genetics*, the distinction being that it is not individual genes that are conserved but entire genomes found in individual trees. The desire and justification for conservation genetics in forestry emerged both from the perspective of preserving genetic diversity for future breeding applications (Zobel 1978) and more broadly to avoid extinction of species due to the loss of genetic diversity (Ledig 1988; Eriksson et al. 1993). In the latter case, the driving principle is to *maintain adaptive evolutionary potential* and to a lesser extent *limit inbreeding* in naturally outcrossing species.

White et al. (2007) provides a broad overview of conservation genetics in forestry; thus, that general background will only be briefly presented here. They begin by describing all the threats to genetic diversity in forest tree populations: deforestation and overharvesting, habitat fragmentation and loss, forest management practices, disease, insects, pollution, invasive species, hybridization, and climate change. In the sections that follow in this chapter, we will provide some examples of how some of these threats are impacting genetic diversity in conifer species and what conservation genetic strategies have been proposed to mitigate these threats. White et al. (2007) defines the two approaches to conservation genetics: in situ and ex situ. In situ conservation approaches focus on the preservation of populations of trees in their natural or near-natural state, and ex situ approaches focus on the collection of biological materials that are maintained outside of the natural state (in such repositories as seed banks, arboretums, or clone banks). In this chapter, we will focus almost entirely on published studies related to in situ conservation.

There is a rather extensive literature in conservation genetics of forest trees, most of which includes conifers. The literature, however, is primarily background and review-type papers on the justification and suggested application of conservation

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_13

genetics. Some notable examples include Ledig (1992), Rajora and Mosseler (2001), Hamrick (2004), Alfaro et al. (2014), Ratnam et al. (2014), Dumroese et al. (2015), Aravanopoulos (2016), and Potter et al. (2017). There are very few empirical studies designed explicitly to evaluate the impact of various threats on genetic diversity. In the sections that follow we will feature a few such studies.

There is also an extensive literature on within- and between-population genetic diversity in conifers based almost entirely on neutral markers (Chap. 9). Many of these papers include "conservation genetics" as a key word. The conclusions reached regarding conservation genetics from these studies relate to descriptions of populations with either limited or highly differentiated variation that then become candidates for conservation genetics programs (Table 13.1). However, as we have noted many times (Chaps. 9, 10, and 12), neutral marker studies can only inform demographic patterns of diversity (drift, migration, mating system) and have limited capacity to inform adaptive (selective) patterns of diversity. We began this chapter by saying that *maintaining adaptive evolutionary potential* is the primary goal of conservation genetics, so this calls into question the practical value of the large body of neutral marker studies for informing conservation genetics strategies.

Beginning in the early part of the twenty-first century, a few background papers appeared which described the potential for genomic data to better inform adaptive evolutionary potential of tree populations (Neale and Kremer 2011; González-Martínez et al. 2006b; Neale 2007; Savolainen and Pyhäjärvi 2007; Neale and Ingvarsson 2008; Savolainen 2013). These papers generally describe methods illustrated in Chaps. 10 and 12 to develop and apply genetic markers tracking adaptive patterns of variation. Aravanopoulos (2016) describes how the metric trait approach (common garden trials) can inform many, but not all, aspects of adaptive genetic diversity and how non-neutral marker studies complement metric studies through their ability to estimate important population genetic parameters such as H_e and N_e (Table 13.2). Furthermore, there are limitations on number of species and time required to obtain metric trait information relative to marker-based studies. Aravanopoulos (2016) articulates a clear vision as to how empirical conservation genetics studies might be conducted in the twentieth century.

Fragmentation

Abies alba (silver fir) is an important forest tree found in a large continuous distribution in central Europe and also in many isolated stands surrounding the central continuous distribution (Fig. 13.1). Notable populations from the fragmented part of the distribution are found throughout the Italian peninsula and southwest France and have been the target of numerous studies. The question related to conservation genetics is whether the isolated populations in the fragmented part of the distribution possess unique genetic variation not found in the large continuous part of the distribution. Many neutral marker (isozyme, cpDNA, mtDNA, and nDNA SSR) studies have been conducted to address this question (Bergmann et al. 1990; Konnert and Bergmann 1995; Parducci et al. 1996; Fady et al. 1999; Vendramin et al. 1999;

Table 13.1 Number of genetically distinct groups and distribution patterns of genetic diversity in North American conifers by region

Species	Family	Genetic markers	Number of groups	Pattern	References
Southeastern North America					
Abies fraseri	Pinaceae	cpDNA SSRs	1 (3?)	No clear pattern. Possible north-south subdivision along the Appalachian Mountains	Potter (2006)
Chamaecyparis thyoides	Cupressaceae	Allozymes	3	Differentiation between Gulf and Atlantic populations	Mylecraine et al. (2004)
Pinus clausa	Pinaceae	Allozymes	2	Differentiation between central Florida and Gulf stands	Parks et al. (1994)
Pinus echinata	Pinaceae	Allozymes	1	Putative expansion from a single refuge in central Texas	Schmidtling (2003)
Pinus virginiana	Pinaceae	Allozymes	2	Differentiation between populations southeast and northwest of the Appalachians	Parker et al. (1997)
Pinus palustris	Pinaceae	Allozymes	2	Possible differentiation east-west of the Mississippi Valley with postglacial expansion from west to east	Schmidtling and Hipkins (1998)
Pinus rigida	Pinaceae	Allozymes	1	No clear pattern. Putative recolonization from a single glacial population	Guries and Ledig (1982)
Pinus taeda	Pinaceae	Allozymes, nuclear SSRs	2	Differentiation east-west of the Mississippi Valley	Schmidtling et al. (1999), Al-Rabab'ah and Williams (2004)
Boreal North America					
Picea glauca	Pinaceae	cpDNA sequences	2	Putative isolated refuge in Alaska	Anderson et al. (2006)
Picea mariana-P. rubens complex	Pinaceae	mtDNA sequences and restriction sites	5	Differentiation among eastern, central-continental, and western stands and putative northeastern glacial refuge for *Picea mariana*. Putative mtDNA recombination in the zone of contact between *Picea mariana* and *Picea rubens*	Gamache et al. (2003), Jaramillo-Correa and Bousquet (2005)

(continued)

Table 13.1 (continued)

Species	Family	Genetic markers	Number of groups	Pattern	References
Pinus banksiana	Pinaceae	mtDNA STR	4	Differentiation among the eastern and central-continental stands. Putative northeastern glacial refuge. Introgression with *Pinus contorta* in the west	Godbout et al. (2005)
Pinus resinosa	Pinaceae	cpDNA SSRs	2	Differentiation among eastern and central-continental stands. Putative northeastern glacial refuge.	Echt et al. (1998), Walter and Epperson (2001)
Pacific Northwest					
Abies procera	Pinaceae	Allozymes	2 (?)	Divergence of coastal isolated stands. North-south differentiation along the Cascades possibly produced by introgression with Abies magnified in southern populations	Yeh and Hu (2005)
Chamaecyparis nootkatensis	Cupressaceae	Allozymes	3	North-south discontinuity among the coastal stands	Ritland et al. (2001)
Larix lyallii	Pinaceae	Nuclear SSRs	2	Putative east-west differentiation in the northern Rocky Mountains	Khasa et al. (2006)
Larix occidentalis	Pinaceae	Nuclear SSRs	2	Putative east-west differentiation in the northern Rocky Mountains	Khasa et al. (2006)
Picea engelmannii	Pinaceae	Allozymes	2	North-south differentiation along the Rocky Mountains. Introgression with *Picea glauca* at the north of its range	Ledig et al. (2006b)
Picea sitchensis	Pinaceae	Nuclear SSRs and ESTPs	1 (?)	Rapid decline of genetic diversity toward the species margins. Isolation by distance	Gapare et al. (2005), Mimura and Aitken (2007)
Pinus albicaulis	Pinaceae	mtDNA, cpDNA SSRs	6-Mar	Discontinuity between the Sierran, Cascades, and Rocky Mountains populations. North-south differentiation within the last two regions	Richardson et al. (2002a, b)

Pinus contorta complex	Pinaceae	Allozymes, RAPDs, mtDNA, cpDNA SSRs	5	No correlation between genetic structure and differences among subspecies. Discontinuity between coastal and montane stands. North-south differentiation along the Pacific Coast and the Rocky Mountains. Quick postglacial expansion with isolation by distance. Putative cpDNA recombination in the zone of contact with *Pinus banksiana*	Wheeler and Guries (1982a), Marshall et al. (2001, 2002), Fazekas and Yeh (2006), Godbout et al. (2008)
Pinus flexilis	Pinaceae	Allozymes, mtDNA	7	Discontinuity between Sierran and Rocky Mountains populations. Modern stands from the Rockies derived from several isolated refugia. Restricted Holocene gene flow	Mitton et al. (2000), Jørgensen et al. (2002)
Pinus monticola	Pinaceae	Allozymes	2	North-south discontinuity along the Cascades	Steinhoff et al. (1983)
Pinus ponderosa complex	Pinaceae	mtDNA	4	Sierran-Cascade - Rocky Mountains discontinuity. Holocene introgression between varieties. North-south subdivision along the Rocky Mountains. Ancient vicariance.	Latta and Mitton (1999), Johansen and Latta (2003)
Pseudotsuga menziesii	Pinaceae	Allozymes, cp DNA SSRs	4	Divergence between coastal and interior varieties. North-south subdivision in interior variety. Divergence of Mexican stands	Li and Adams (1989), Aagaard et al. (1998b), Viard et al. (2001)
Thuja plicata	Cupressaceae	RFLPs, nuclear SSRs	3	Subdivision among northern-coastal, central and southern-interior stands. Putative recolonization from a single glacial population	Glaubitz et al. (2000), O'Connell et al. (2008)
California					
Abies bracteata	Pinaceae	Allozymes	1	No apparent differentiation among isolated stands.	Ledig et al. (2006a)
Abies magnified	Pinaceae	cpDNA sequences	2	Discontinuity between the Sierra Nevada and the Klamath - southern Cascades stands. Putative introgression from Abies *procera*	Oline (2008)
Chamaecyparis lawsoniana	Cupressaceae	Allozymes	2 (?)	Putative differentiation between inland and coastal stands	Millar and Marshall (1991)
Picea breweriana	Pinaceae	Allozymes	1	No clear pattern. Population divergence mainly modeled by genetic drift	Ledig et al. (2005)

(continued)

Table 13.1 (continued)

Species	Family	Genetic markers	Number of groups	Pattern	References
Pinus attenuata, *P. radiata*, *P. muricata*	Pinaceae	Allozymes, RAPDs, cpDNA and mtDNA restriction sites	_	Ancient vicariance among species with putative south-north migration. North-south differentiation along costal population observed with cytoplasmic markers for *Pinus muricata*	Hong et al. (1993), Wu et al. (1998, 1999)
Pinus balfouriana	Pinaceae	Allozymes, cpDNA, mtDNA and nuclear DNA sequences	1	No clear genetic structure reflecting the geographic differentiation between varieties or among populations of the same variety. Ancient vicariance	Oline et al. (2000), Eckert et al. (2008)
Pinus coulteri	Pinaceae	Allozymes	1	Migration from a single southern refugia with clinal loss of diversity. Putative introgression with *Pinus jeffreyi* in sympatric populations	Ledig (2000)
Pinus jeffreyi	Pinaceae	Allozymes	2	Discontinuity between Sierran and Coastal ranges populations	Furnier and Adams (1986)
Pinus lambertiana	Pinaceae	cpDNA sequences	2	North-south disjunction along the Sierra Nevada. Putative introgression from *Pinus albicaulis* in the northern part of the distribution	Liston et al. (2007)
Pinus sabiniana	Pinaceae	Allozymes	1	Holocene collapse and fragmentation. No clear geographic pattern	Ledig (1999)
Mexico and northern Central America					
Abies flinckii, *A. guatemalensis*, *A. hickeli*, *A. religiosa*	Pinaceae	Allozymes, mtDNA and cpDNA SSRs	9	Divergence of isolated stands along the Transverse Volcanic Belt and set the limits of the species' distribution. Differentiation mainly modeled by genetic drift after Holocene collapse.	Aguirre-Planter et al. (2000), Jaramillo-Correa et al. (2008)

Picea chihuahuana	Pinaceae	Allozymes, cpDNA SSRs, mt DNA sequences	2	North-south differentiation along the Sierra Madre Occidental observed with mtDNA markers. Population divergence mainly modeled by genetic drift	Ledig et al. (1997), Jaramillo-Correa et al. (2006)
Picea martinezii	Pinaceae	Allozymes, cpDNA, mt DNA sequences	1	Almost no differentiation among stands	Ledig et al. (2000)
Picea mexicana	Pinaceae	Allozymes, cpDNA, mt DNA sequences	3	Virtually all populations differ from each other. Divergence mainly modeled by genetic drift after Holocene collapse	Ledig et al. (2000)
Pinus chiapensis	Pinaceae	RAPDs, mtDNA restriction sites	11	Virtually all populations differ from each other. Divergence mainly modeled by genetic drift after Holocene collapse.	Newton et al. (2002)
Pinus greggii	Pinaceae	Allozymes	3	Divergence of isolated stands along the Sierra Madre Oriental mainly modeled by genetic drift during the Holocene	Parraguirre-Lezama et al. (2002)
Pinus montezumae-P. hartwegii complex	Pinaceae	cpDNA restriction sites	4	Divergence of isolated stands along the Transverse Volcanic Belt and the Sierra Madre Oriental. Introgressive hybridization between the two taxa	Matos and Schaal (2000)
Pinus montezumae-P. pseudostrobus complex	Pinaceae	cpDNA SSRs	3	Ancient vicariance with lineage sorting. Introgressive hybridization between the two taxa. Isolation by distance among stands of the same species	Delgado et al. (2007)
Pinus nelsonii	Pinaceae	cpDNA SSRs	1	Low population differentiation probably associated with recent divergence. Isolation by distance among stands	Cuenca et al. (2003)

(continued)

Table 13.1 (continued)

Species	Family	Genetic markers	Number of groups	Pattern	References
Pinus oocarpa	Pinaceae	Allozymes, RAPDs, AFLPs	2 (?)	Low divergence among populations. Putative introgression with *Pinus patula* subsp. *tecunumannii*	Matheson et al. (1989), Diaz et al. (2001)
Pinus pinceana	Pinaceae	Allozymes	3	North-south divergence along the Sierra Madre Oriental probably related to Holocene population collapse	Ledig et al. (2001)
Pinus rzedowskii	Pinaceae	Allozymes	2	Divergence of isolated stands along the Transverse Volcanic Belt mainly modeled by genetic drift during the Holocene	Delgado et al. (1999)
Pinus strobiformis-P. ayacahuite complex	Pinaceae	cpDNA SSRs	5	Divergence between the Sierra Madre Occidental and Sierra Madre Oriental stands. Differentiation of some isolated populations. Introgressive hybridization and (or) shared ancestral polymorphism between both taxa	Moreno-Letelier and Pinero (2009)

Modified from Table 1 in Jaramillo-Correa et al. (2009)

Table 13.2 Criteria, indicators, and verifiers in the genetic monitoring of forest tree populations

Criterion	Indicator	Verifier	Metric trait	Genetic marker
Genetic monitoring system of Aravanopoulos (2011): Conservation of genetic diversity and adaptive evolutionary potential in natural populations				
	1. Selection			
		1.1. Age and size class distribution	✓	
		1.2. Reproductive fitness (percentage of filled seeds and percentage of germination)	✓	
		1.3. Regeneration abundance	✓	
	2. Genetic drift			
		2.1. Effective population size		✓
		2.2. Allelic richness		✓
		2.3. Latent genetic potential		✓
	3. Gene flow			
		3.1. Outcrossing/actual inbreeding rate		✓
German forest genetic monitoring system (Konnert et al. 2011)				
	1. Genetic variation			
		1.1. Gene/genotype frequencies		✓
		1.2. Genetic diversity: allelic richness (A/L), P		✓
		1.3. F-value		✓
		1.4. Phenological parameters	✓	
		1.5. No. of potential parents	✓	
	2. Directional change in gene/genotypic frequencies			
		2.1. Allele, genotype, phenotype frequencies		✓
		2.2. Distribution of age classes		✓
	3. Mating system			
		3.1. Rate of cross-fertilization		✓
		3.2. Rate of biparental inbreeding		✓
		3.3. No. of effective pollen donors		✓
		3.4. Proportion of empty/full seed	✓	
		3.5. Germinability		✓
	4. Gene migration			
		4.1. Dispersion of pollen and seeds	✓	
		4.2. Population differentiation, isolation	✓	
		4.3. Family structures	✓	

(continued)

Table 13.2 (continued)

Criterion	Indicator	Verifier	Metric trait	Genetic marker
Local genetic diversity indicators of evolutionary potential (operational indicator; Graudal et al. 2014)				
	1. Demographic condition of selected populations (diversity in adaptive traits/genes)			
		1.1. Age/size class distribution	✓	
		1.2 Number of reproducing trees	✓	
		1.3 Abundance of regeneration	✓	
		1.4. Environmental heterogeneity	✓	
		1.5. Number of filled seeds	✓	
		1.6. Percentage of germination	✓	
	2. Genetic condition of selected populations (population genetic structure where appropriate)			
		2.1. Effective population size		✓
		2.2. Allelic richness		✓
		2.3. Outcrossing/inbreeding rate		✓
		2.4. Spatial genetic structure		✓
		2.5. Hybridization/introgression		✓
EUFORGEN genetic monitoring system for genetic conservation units of forest trees in Europe (Aravanopoulos et al. 2015): Conservation of the processes that maintain genetic variation				
	1. Selection			
		1.1. Age/size class distribution	✓	
		1.2. Mortality	✓	
		1.3. Regeneration abundance	✓	
		1.4. Fructification	✓	
		1.5. Reproductive fitness in mast years (% of filled seeds and % germination)	✓	
	2. Genetic variation and mating system			
		2.1. Effective population size (N_E)		✓
		2.2. Allele/genotype frequencies		✓
		2.3. Genetic diversity parameters: allelic richness (A/L), N_e, P, H_e, H_0, latent genetic potential, F_{IS}, F_{ST} (+outlier tests)		✓
		2.4. Outcrossing or actual inbreeding rate		✓
		2.5. Interspecific hybridization percentage (where applicable)		✓
		2.6. Sex ratio (dioecious species)	✓	

From Aravanopoulos (2016)

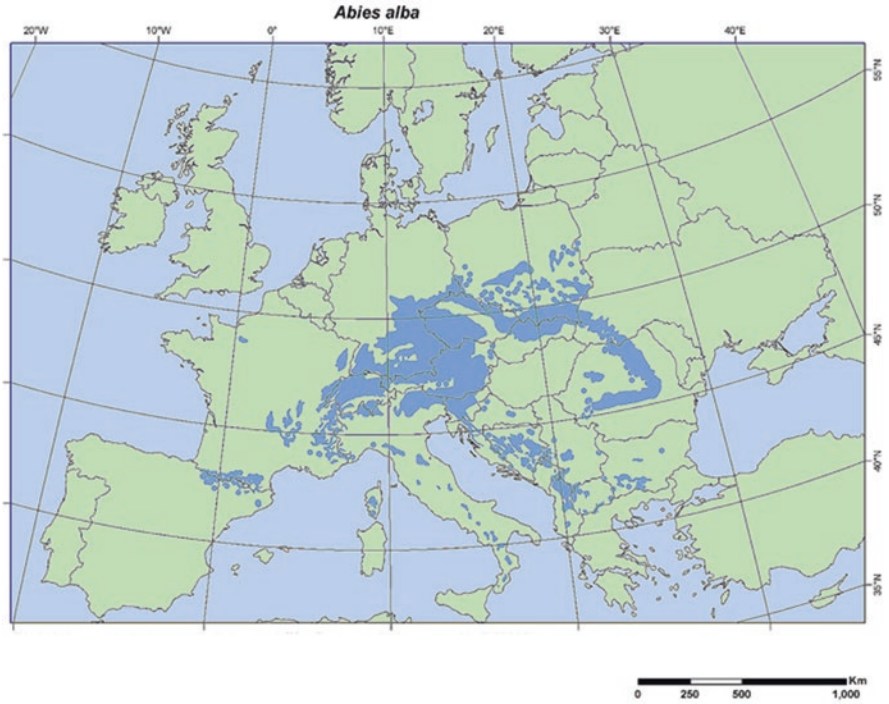


Fig. 13.1 Range map of *Abies alba*, from EUFORGEN 2009, www.euforgen.org

Liepelt et al. 2002; Gömöry et al. 2004; Liepelt et al. 2009; Gömöry et al. 2012). The demographic, postglacial migration history of *A. alba* is firmly established by this body of work. The neutral marker data support five refugia, (1) eastern Balkan, (2) western Balkan, (3) Italian Apennine, (4) southern Italy, and (5) Pyrenees, with only the first three contributing to the present-day central, continuous distribution portion of the range. The neutral marker data indicate that the isolated populations on the Italian peninsula (Parducci et al. 1996; Piovani et al. 2010) and southeastern France (Sagnard et al. 2002) are genetically distinct and merit conservation activity as they may harbor unique adaptive potential to the environments they inhabit. These conclusions, however, are reached entirely based on neutral marker data.

Two studies have been conducted more recently to determine what, if any, adaptive genetic diversity resides in the isolated populations. Sagnard et al. (2002) sampled 569 trees from 16 isolated populations in southwestern France and typed all individuals for isozyme loci, monoterpene composition, and three adaptive traits (growth, bud phenology, and drought resistance). They found considerable variability for all traits, but only monoterpene composition was correlated with population source (Fig. 13.2). The important implication of this result is that there must in fact be important adaptive genetic diversity within isolated populations but because the within-population component of quantitative trait variation is higher than the between-population component, knowing which trees (genotypes) should be

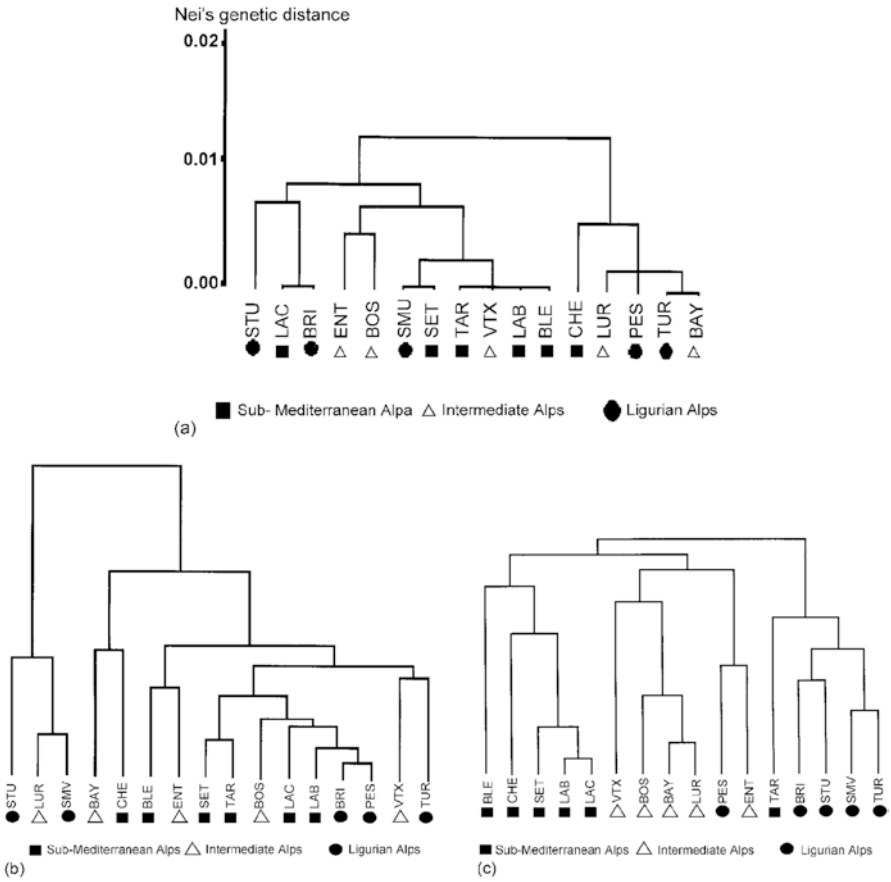


Fig. 13.2 Cluster analysis of 16 populations of *Abies alba* (**a**) isozyme data, (**b**) adaptive traits, and (**c**) mean monoterpene composition. (From Sagnard et al. 2002)

included in a conservation genetics program will require an understanding of which trees within stands would be the best to include in an in situ conservation genetics program. This information could be obtained from common garden tests where candidate trees for conservation are evaluated; however, the practicality of such an approach is limited. As an alternative, this group of researchers has turned to F_{ST} *outlier analysis* (Chap. 10) and *Environmental Association Analysis* (Chap. 12) to identify trees carrying allelic variation for adaptive genetic potential (Roschanski et al. 2016). Loci potentially harboring adaptive genetic diversity based on these analyses are shown in Fig. 13.3. This approach to conservation genetics is just beginning and is an early empirical example of the direction advocated by Aravanopoulos (2016).

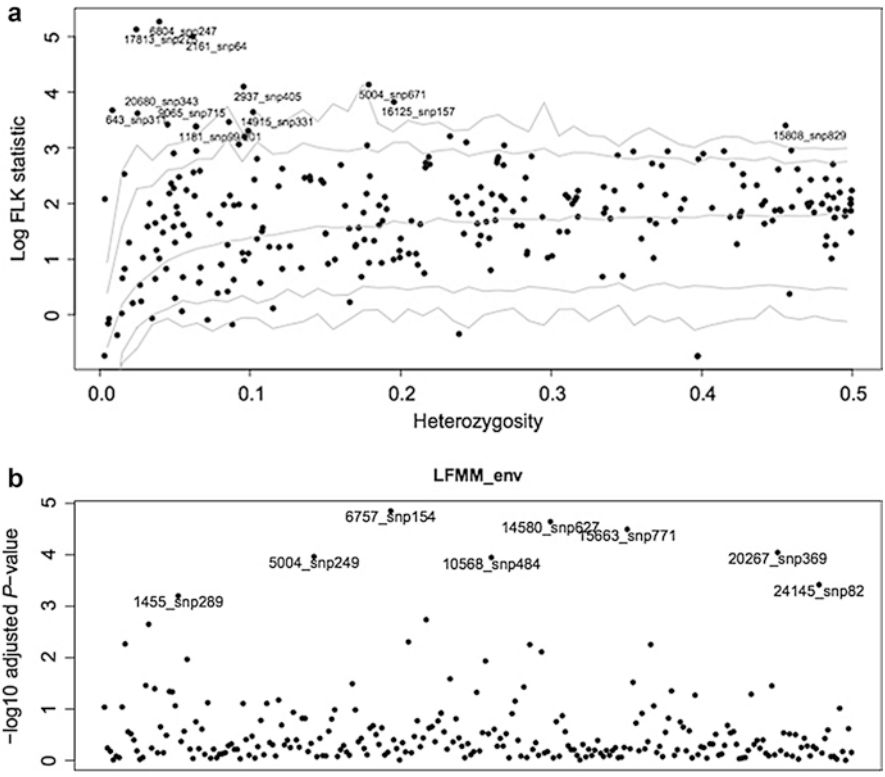


Fig. 13.3 Evidence for directional selection from three F_{ST} outlier tests. (**a**) Distribution of log FLK statistic as a function of heterozygosity. Lines depict the 0.5%, 2.5%, 50%, 97.5%, and 99.5% boundaries of the empirical neutral distribution of FLK statistic obtained by forward in time simulation. Only outlier SNPs are labeled with their SNP IDs. (**b**) Manhattan plot of LFMM_env and LFMM_PCA. The *x*-axis represents an arbitrary order of the candidate gene SNPs. Only the *P*-values of significant SNPs with FDR <0.05 are labeled. (From Roschanski et al. 2016)

Habitat Loss

Many forest tree species, including conifers, are undergoing habit loss due to human impacts. Habit loss for rare and endangered species is of greatest interest and urgency. The example we will use here is of *Pinus torreyana* (Torrey pine), a species endemic to California, which was likely widespread a few thousand years ago but is now found in just two small disjunct populations, separated by 280 km. One is on the coastal mainland near San Diego and the other is on Santa Rosa Island. Census data estimate there are 7000 trees on the mainland and 2000 trees on the island. Ledig and Conkle (1983) performed an isozyme study with a surprising result. Both populations were completely monomorphic at 59 loci, all with the same allele except for two loci that were fixed for different alleles. This result suggests that

these populations harbor no adaptive evolutionary potential and are at great risk of extinction. To further assess this inference made from 59 neutral genetic markers, Hamilton et al. (2017) conducted a common garden study, located on the mainland near Santa Barbara, including offspring from mainland, island, and F_1 individuals from mating island with mainland trees. They found genetic variation in two quantitative traits, height growth and conelet production (Fig. 13.4), suggesting that there may in fact be genetic variation within and between these populations, although both populations are likely quite narrowly adapted. The most interesting result was that the F_1 individuals had more variation than either of the parent types. This led Hamilton et al. (2017) to suggest that *genetic rescue* might be an effective conservation strategy for this species by introducing individuals into each population from the other population. The alternate risk associated with this approach is

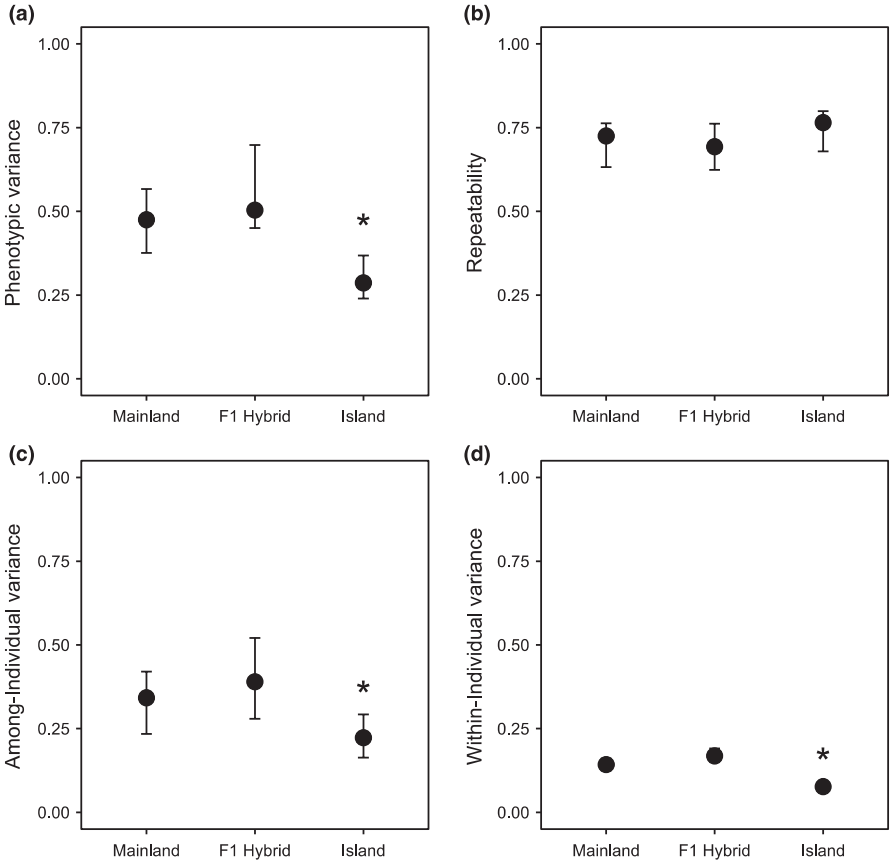


Fig. 13.4 Phenotypic variance based on a Bayesian hierarchical mixed model of height, repeatedly measured between 2008 and 2016 of mainland, island, and F_1-hybrid *Pinus torreyana* populations. Variance is partitioned between the three populations into total phenotypic variance (**a**), repeatability of multiyear measurements (**b**), and variance attributable to among-individual (**c**) and within-individual sources of variation (**d**). (From Hamilton et al. 2017)

maladaptation due to outbreeding depression. Hamilton et al. (2017) further suggested that genomics approaches are needed to understand specifically what genetic loci and alleles at those loci will be important for maintaining adaptive evolutionary potential in these two populations or in new habitats that *P. torreyana* might colonize.

Forest Practice

Harvesting of natural stands of forest trees can potentially impact the genetic diversity in the remaining unharvested stand. Silvicultural systems such as clear cutting, shelterwood, and seed tree, followed by natural regeneration, have all been evaluated in several conifer species (reviewed in Ratnam et al. 2014) (Table 13.3). It might be expected that the fewer trees left after harvesting the greater the loss of diversity might be (genetic drift). However, because of long dispersal distances of wind-dispersed seeds in conifers, the source of seed to regenerate stands can easily come from stands at great distances from the harvested stand, thus maintaining diversity.

The general approach used in these studies was to sample the stand before harvest, or an adjacent stand if the harvest had already occurred, and then sample trapped seed and natural regeneration after harvesting. In all the studies reported in Ratnam et al. (2014), neutral marker data were obtained on all samples and standard measures of genetic diversity were estimated and then compared. Results obtained in an early study of the shelterwood harvest system in *Pseudotsuga menziesii* are characteristic of most studies (Neale 1985; Neale and Adams 1985). The amount of allelic diversity between the preharvest stand and the regeneration is generally about the same (Neale 1985), suggesting no significant loss of diversity was due to harvesting. What does seem to increase is the amount of inbreeding in the regenerated stand (Neale and Adams 1985). However, the regenerated stands are generally evaluated at a young age, so as the stand matures the amount of inbreeding often decreases. Therefore, with just a few exceptions, the general result is that harvesting systems in conifers do not pose a significant threat to the maintenance of genetic diversity in populations. However, as we have noted earlier, these studies have all been done with neutral genetic markers so the maintenance of adaptive genetic diversity remains unknown.

In an interesting study by DeWald and Kolanoski (2017), the impact of different possible thinning scenarios to restore overstocked *Pinus ponderosa* (ponderosa pine) stands in Arizona back to presettlement stand structures was evaluated. This study was again done with neutral genetic markers (isozymes). They found that the somewhat "clumpy" population structure of presettlement stands had slightly less diversity than post-settlement stands (Fig. 13.5), suggestive of "genetic neighborhoods." However, under simulated conditions it was inferred that a 50% thinning of overstocked stands would not lead to a loss of diversity but a 75% thinning would (Fig. 13.6). This is just one small study but illustrates the type of study that could be done in advance of restoration efforts that include removal of large numbers of trees.

Table 13.3 Summary of key studies on genetic effects of forest management practices in different regions of the world

Region	Species	Mating system/ Pollen dispersal	Other relevant traits	Silvicultural practices	Stand types studied	Key results	References
North America	*Picea glauca*	Predominantly outcrossing, wind-pollinated	Boreal, late successional	Clear-cut harvesting, tree improvement			Rajora (1999), Fageria and Rajora (2014)
				Experimental harvesting of increased intensities (green tree retention of 75%, 50%, 20%, and 10%, and clear-cut)	Natural old-growth, natural regeneration after harvesting of five intensities	No negative impact of harvesting of five increased intensities on genetic diversity, inbreeding (F_{IS}) and population structure	Fageria and Rajora (2013)
	Picea mariana	Predominantly outcrossing, wind-pollinated	Boreal, early successional semi-serotinous cones	Clear-cut harvesting, artificial and natural regeneration, forest fires	Natural mature, postharvest and postfire natural regeneration and plantations	No adverse effects on genetic diversity (A, A_e, H_e, H_o and genotypic diversity [GA]) and inbreeding levels	Rajora and Pluhar (2003)
				Clear-cut harvesting	Natural mature, natural young, plantations	No significant allelic heterogeneity among different stand types	Knowles (1985)
	Pinus contorta	Predominantly outcrossing, wind-pollinated	Boreal, temperate, early successional semi-serotinous cones	Clear-cut harvesting	Natural mature, postharvest natural young regeneration, plantations	No significant differences in genetic diversity of unmanaged mature, postharvest naturally regenerated young and planted young stands	Thomas et al. (1999), Macdonald et al. (2001)

	Pseudotsuga menziesii	Predominantly outcrossing, wind-pollinated	Temperate, shade tolerant	Shelterwood, group selection, clear-cut	Natural old- growth, residual trees, postharvest seed, plantations	No negative impacts on genetic diversity and mating system; loss of rare alleles after shelterwood harvesting	Neale (1985), Neale and Adams (1985), Adams et al. (1998)
	Pinus strobus	Predominantly outcrossing, wind-pollinated	Temperate and boreal, early to late successional	Seed tree (removal of 75% trees), shelterwood, pine release	Natural old- growth, postharvest residual and natural regeneration	Significant reduction in allelic and genotypic diversity after seed tree cut, loss of 20–90% rare and low-frequency alleles; no significant reduction in heterozygosity after any of the three harvesting systems practiced	Buchert et al. (1997), Rajora et al. (2000b), Marquardt et al. (2007)
	Tsuga heterophylla	Predominantly outcrossing, wind-pollinated	Temperate montane, shade tolerant	Patch cut, shelterwood cut, green tree retention	Natural mature, natural young regeneration	No negative genetic impacts of patch cut and green tree retention; shelterwood system resulted in lower heterozygosity	El-Kassaby et al. (2003)
	Abies amabilis	Predominantly outcrossing, wind-pollinated	Temperate montane, shade tolerant	Patch cut, Shelterwood cut, green tree retention	Natural mature, natural young regeneration	No significant impacts of silvicultural treatments on genetic diversity	El-Kassaby et al. (2003)
	Tsuga canadensis	Predominantly outcrossing, wind-pollinated	Temperate, shade tolerant	Selection and diameter limit cuts	Natural mature, natural young regeneration	Change in allele frequencies and genetic structure	Hawley et al. (2005)

(continued)

Table 13.3 (continued)

Region	Species	Mating system/ Pollen dispersal	Other relevant traits	Silvicultural practices	Stand types studied	Key results	References
	Pseudotsuga menziesii, *Tsuga heterophylla*, *Thuja plicata*, *Pinus monticola*, *Abies amabilis*	Predominantly outcrossing to mixed-mating, wind-pollinated	Temperate	Commercial thinning	Pre- and post-commercially thinned plantations	Loss of 1 to 7 alleles	El-Kassaby and Benowicz (2000)
Northern and Central Europe	*Pinus sylvestris*	Wind-pollinated		Shelterwood cut, group selection cut	Natural	Nonsignificant increase in gene flow and outcrossing rate, reduction in consanguineous mating	Robledo-Arnuncio et al. (2004)
Mediterranean	*Pinus brutia*	Wind-pollinated		Overexploitation	Natural	No reduction in genetic diversity, slight increase in inbreeding	Lise et al. (2007)
	Taxus baccata	Wind-pollinated		Fragmentation	Natural	Demographic bottlenecks, high levels of inbreeding, strong genetic structure	Dubreuil et al. (2010)

From Ratnam et al. (2014)

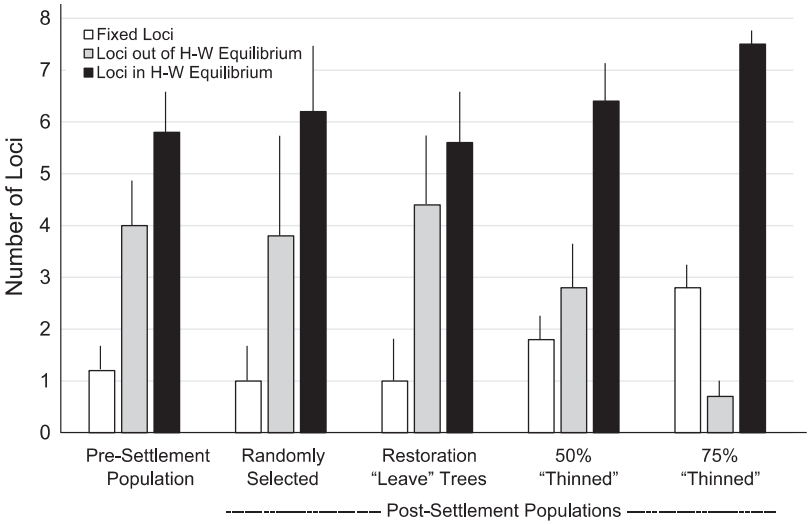


Fig. 13.5 Mean number of loci fixed, deviating, and not deviating from Hardy-Weinberg equilibrium with 95% confidence for the different types of *Pinus ponderosa* populations studied at the Fort Valley Experimental Forest, Flagstaff, AZ. (From DeWald and Kolanoski 2017)

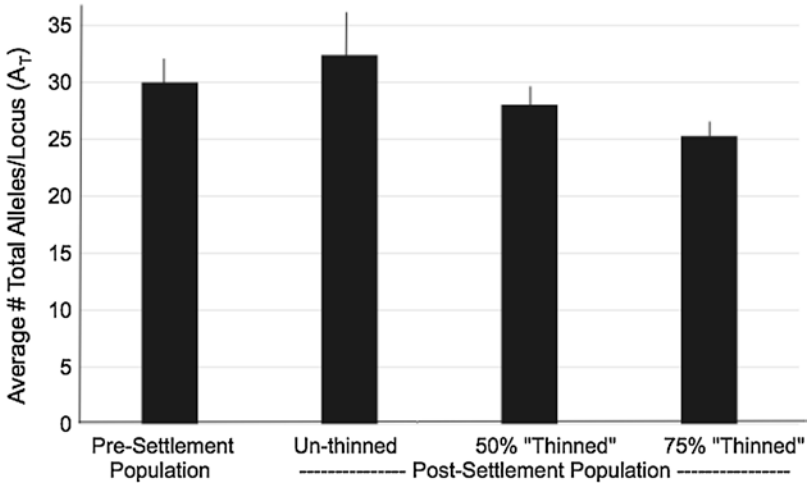


Fig. 13.6 Differences in average total number of alleles per locus for different *Pinus ponderosa* populations at the Fort Valley Experimental Forest, Flagstaff, AZ. 75% "Thinned" value is significantly different ($P < 0.0001$) from un-thinned post-settlement. (From DeWald and Kolanoski 2017)

Disease

Introduced pathogens have brought several forest trees to near extinction. In North America three very notable cases are *Castanea dentata* (American chestnut), *Ulmus americana* (American elm), and *Juglans cinerea* (butternut), all of which have been driven nearly to extinction by introduced diseases from Asia (Chap. 14). There are no examples from conifers quite so severe, although the situation with species of *Pinus* subgenus *Strobus* and white pine blister rust (WPBR) has become quite severe (Chap. 14). The one species of this group which may be at the greatest risk is *Pinus albicaulis* (whitebark pine), which we will feature here.

Pinus albicaulis is found at very high elevations in western North America (Fig. 13.7). For this reason much of its range is made up of small disjunct populations on mountain tops. It has a wingless seed that is dispersed by the passerine bird species Clark's nutcracker with which it has coevolved. This pine is listed as a threatened species largely due to damage inflicted by the introduced WPBR pathogen. Because of its critical role in subalpine ecosystems as a food source for wildlife, there is great interest in its conservation and restoration. This has been the topic of several important symposia (Tomback et al. 2001; Keane et al. 2011) and a comprehensive restoration strategy has been proposed (Keane et al. 2012). Keane et al. (2012) thoroughly review the state of knowledge on *P. albicaulis* genetic diversity, which we will summarize briefly here. Furthermore, the genetics of resistance to WPBR is covered in detail in Chap. 14.

Genetic variation in *P. albicaulis* has been studied using genetic markers and common garden studies. The early genetic marker studies used isozymes (Jorgensen

Fig. 13.7 Range map of *Pinus albicaulis.* (From Little 1971)

and Hamrick 1997; Bruederle et al. 1998) which showed below-average amounts of H_e and average estimates of F_{ST} (Tables 13.4 and 13.5). These early studies indicated that there is population variation but primarily within populations, not between. Three studies were done at a more local scale to quantify the effect of seed caching by the nutcracker on population diversity (Furnier et al. 1987; Rogers et al. 1999; Richardson et al. 2002b). These studies found ample diversity among trees originating from individual seed caches (Fig. 13.8) and differences among distant cohorts suggesting that the bird dispersal distance is not great. In two additional studies using isozymes, the outcrossing rate was estimated and found to be somewhat less than most other conifers, t_m being around 0.85 (Table 13.6) (Krakowski et al. 2003; Bower and Aitken 2007). The conservation genetics implication of this result is that ex situ collections might contain a fair bit of inbred seed.

Table 13.4 Summary of genetic diversity within mountain range and regional groupings of 30 populations of *Pinus albicaulis* based on 20 putative allozyme loci

Mountain range or regional grouping	*P* (SD)	*P*95 (SD)	*AP*	H_e (SD)	*F*
Mean Cascade Mtns.	35.0 (2.3)	21.3 (2.5)	2.09	0.086 (0.005)	0.037
Pooled Cascade Mtns.	60	25	2.58	0.09	
Mean Sierra Nevada	33.8 (1.3)	20.0 (0.0)	2.29	0.104 (0.006)	0.033
Pooled Sierra Nevada	50	25	2.5	0.109	
Mean west	34.6 (1.6)	20.8 (1.6)	2.16	0.092 (0.004)	0.036
Pooled west	65	25	2.77	0.098	
Mean Rocky Mtns.	45.6 (2.7)	24.7 (0.9)	2.24	0.101 (0.003)	0.107
Pooled Rocky Mtns.	85	25	2.82	0.105	
Mean Great Basin	32.5 (2.5)	22.5 (2.5)	2.07	0.077 (0.013)	0.135
Pooled Great Basin	35	25	2.14	0.086	
Mean east	44.2 (2.6)	24.4 (0.8)	2.22	0.098 (0.003)	0.11.0
Pooled east	85	25	2.82	0.104	

From Jorgensen and Hamrick (1997)

Table 13.5 Genetic variability in nine populations of *Pinus albicaulis* in Montana and Wyoming, USA

Site	*A* (SE)	A_p	A_e	$P_{(0.95}$ criterion)	$P_{(no}$ criterion)	H_o (SE)	H_e (SE)
Miller Creek	1.6 (0.2)	2.2	1.17	42.1	47.4	0.167 (0.052)	0.146 (0.045)
Henderson Mountain	1.7 (0.2)	2.3	1.2	36.8	52.6	0.139 (0.043)	0.164 (0.049)
Fisher Creek	1.5 (0.1)	2.1	1.18	42.1	47.4	0.139 (0.043)	0.151 (0.047)
Island Lake	1.7 (0.2)	2.2	1.18	36.8	57.9	0.153 (0.047)	0.151 (0.046)
Mount Washburn	1.6 (0.1)	2.1	1.17	36.8	52.6	0.139 (0.045)	0.145 (0.046)
Union Pass	1.5 (0.1)	2.1	1.17	36.8	42.1	0.161 (0.054)	0.146 (0.047)
Togwotee Pass	1.7 (0.2)	2.3	1.18	36.8	52.6	0.141 (0.047)	0.153 (0.046)
Sheep Pass	1.6 (0.2)	2.3	1.18	36.8	47.4	0.152 (0.045)	0.155 (0.047)
Commissary Ridge	1.6 (0.2)	2.4	1.19	42.1	42.1	0.138 (0.042)	0.161 (0.048)

From Bruederle et al. (1998)

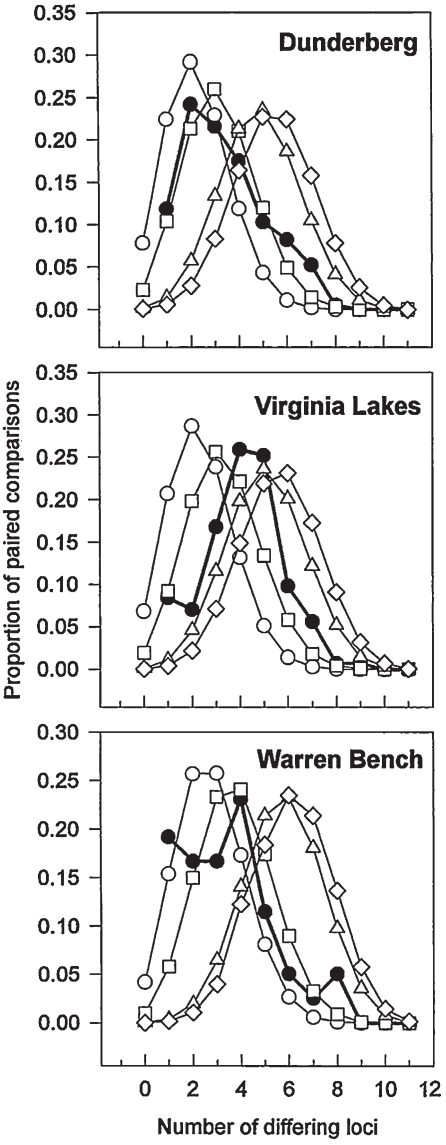


Fig. 13.8 Observed (●) and expected proportions of pairs of individuals differing by a given number of loci within krummholz thickets of *Pinus albicaulis* in the California Sierra Nevada mountains; expected distributions are based on four models of relatedness: selfs (○); full-siblings (□); half-siblings (Δ); and unrelated (◇). (From Rogers et al. 1999)

Finally, a range-wide demographic study was done using both mtDNA (maternal) and cpDNA (paternal) SSR markers (Richardson et al. 2002a). This study shows the major haplotypic groups arising from postglacial seed migration (Fig. 13.9) and pollen migration (Fig. 13.10). This study reveals the major genetic structuring across the range but also confirms the high within versus between component of variation. None of these studies provide any direct evidence for the patterning of adaptive genetic variation.

Table 13.6 *Pinus albicaulis* population-level estimates of outcrossing rates t_m and t_s (multilocus and single-locus, respectively) by locus in two populations in British Columbia, Canada; r_p is correlation of paternity among progeny; r_t is correlation of outcrossing rates among progeny arrays, averaged for all parent trees sampled; standard errors of the means given in parentheses

Locus	Mount Baldy	Manning
Pgi1	0.762 (0.109)	0.493 (0.230)
Pgi2	0.888 (0.083)	0.777 (0.150)
Pgm	0.396 (0.282)	0.123 (0.057)
Idh	0.621 (0.085)	0.709 (0.155)
Mdh2	0.646 (0.089)	0.758 (0.055)
Mdh3	1.319 (0.952)	0.614 (0.385)
Mdh4	0.897 (0.107)	0.759 (0.069)
6Pg1	0.913 (0.114)	0.684 (0.160)
6Pg2	monomorphic	0.294 (0.237)
Combined SL	0.735 (0.048)	0.650 (0.061)
Combined ML	0.736 (0.042)	0.722 (0.054)
$t_m - t_s$	0.001 (0.014)	0.068 (0.025)
r_t	0.082 (0.052)	0.074 (0.046)
r_p	0.208 (0.082)	0.148 (0.063)
No. of families	30	25
No. of observations	853	750

From Krakowski et al. (2003)

A comprehensive common garden study in *P. albicaulis* was conducted by Canadian researchers to estimate the adaptive evolutionary potential within populations (Bower and Aitken 2006, 2008). In these studies, several quantitative traits (height, biomass, root-to-shoot ratio, date of bud flush, and spring and fall cold injury) were evaluated from a sample of families from 48 populations. Q_{ST} for the growth traits was low (0.07–0.14) but fairly high for cold adaptation (0.36–0.47). This led the authors to conclude that local adaptation for cold adaptation was important and that moving sources from warmer to colder environments could risk maladaptation. This inference is made at the source level, but as we noted earlier, the greatest component of variation is always within versus between populations, so considerable variation among trees within source should be expected and accounted for in conservation genetics and seed-transfer guidelines.

Genomic approaches to understanding adaptive evolutionary potential in *P. albicaulis* have only recently been applied. Liu et al. (2016a) conducted a large study to determine patterns of variation among populations in genes potentially associated with WPBR resistance in *P. albicaulis.* They began with an RNA-seq study (Chap. 6) from which they identified 22,000 non-synonymous SNPs. Following gene annotation, 216 SNPs from candidate genes for disease resistance were used to genotype 371 individuals from a large part of the range of *P. albicaulis.* The large amount of SNP variation was classified into each of nine subgroups and the distribution of that variation displayed (Fig. 13.11). Here for the first time we begin to see individual trees within populations that may contain allelic variation at loci underlying adaptive genetic potential.

Lind et al. (2017) conducted a microgeographical study in the Lake Tahoe Basin of California that led to a discovery similar to that of Liu et al. (2016a). The added

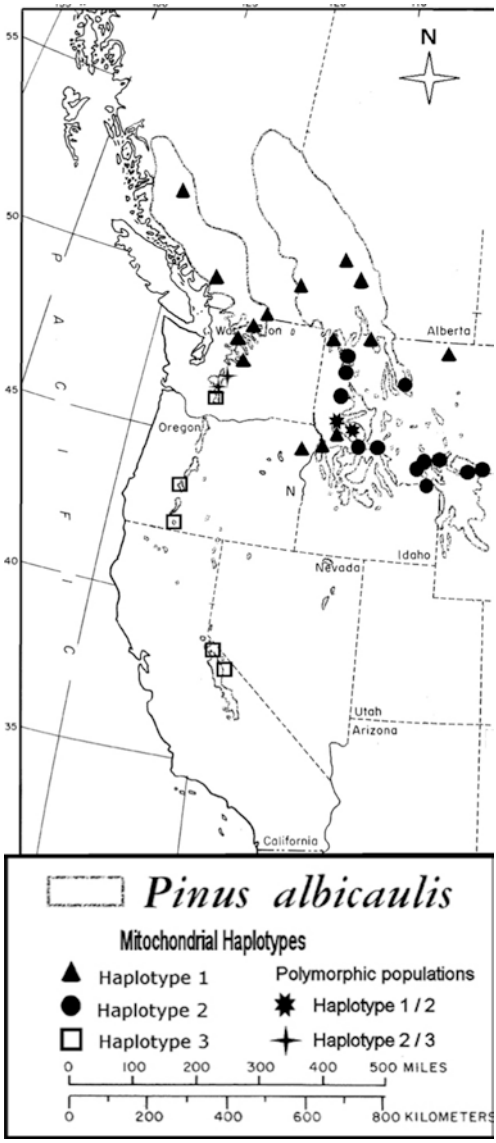


Fig. 13.9 Distribution map of *Pinus albicaulis* mtDNA haplotypes. (Modified From Richardson et al. 2002a)

dimension of the Lind et al. (2017) study was that the SNP genetic variation was associated with WPBR resistance in common gardens as well as being associated with the environments from where the trees originated. This is a rare example of where variations in the phenotype, genotype, and environment are used in parallel in a single experiment in an attempt to discover allelic variation responsible for adaptive evolutionary potential.

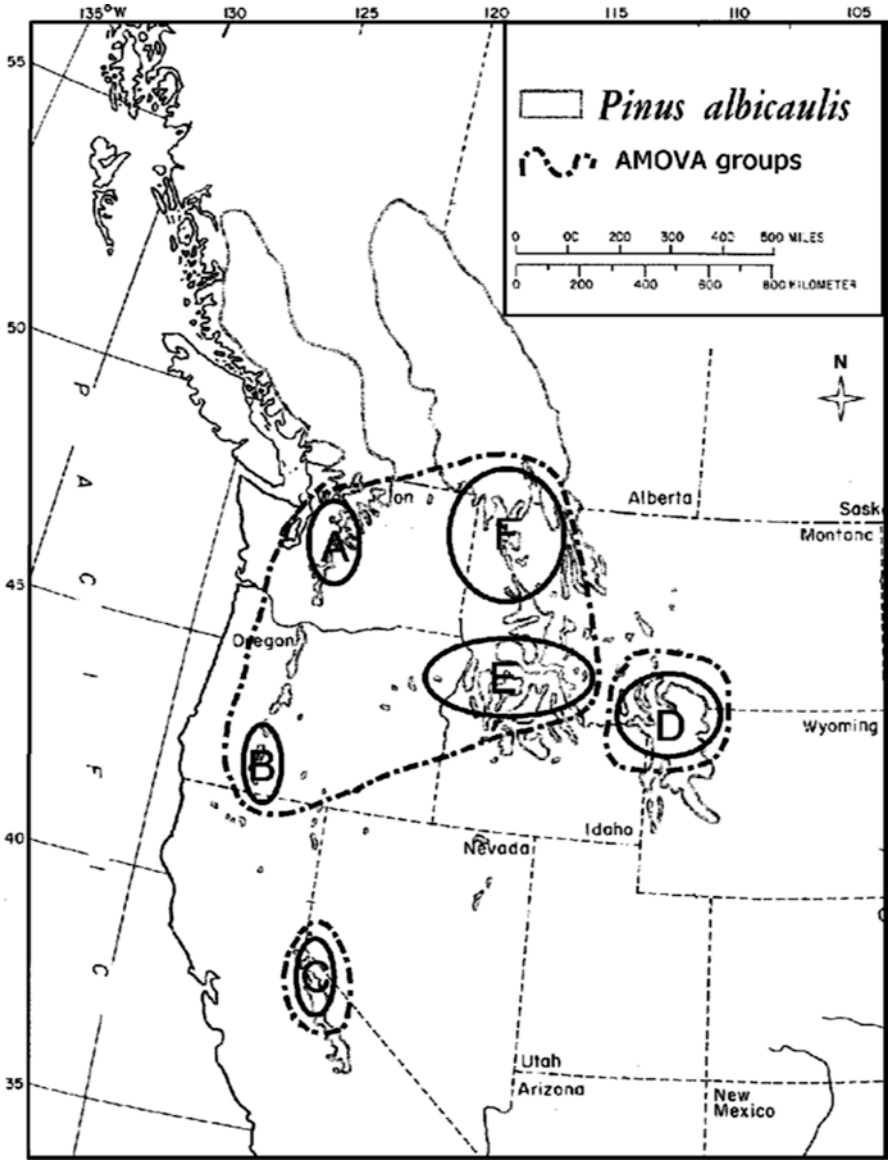


Fig. 13.10 Areas of sampling for *Pinus albicaulis* cpSSR haplotype regions (solid lines) and AMOVA haplotype groups (broken lines) based on an exact test; A, northern Cascades; B, southern Oregon; C, Sierra Nevada Mountains; D, Yellowstone; E, central Idaho; F, northern Idaho. (From Richardson et al. 2002a)

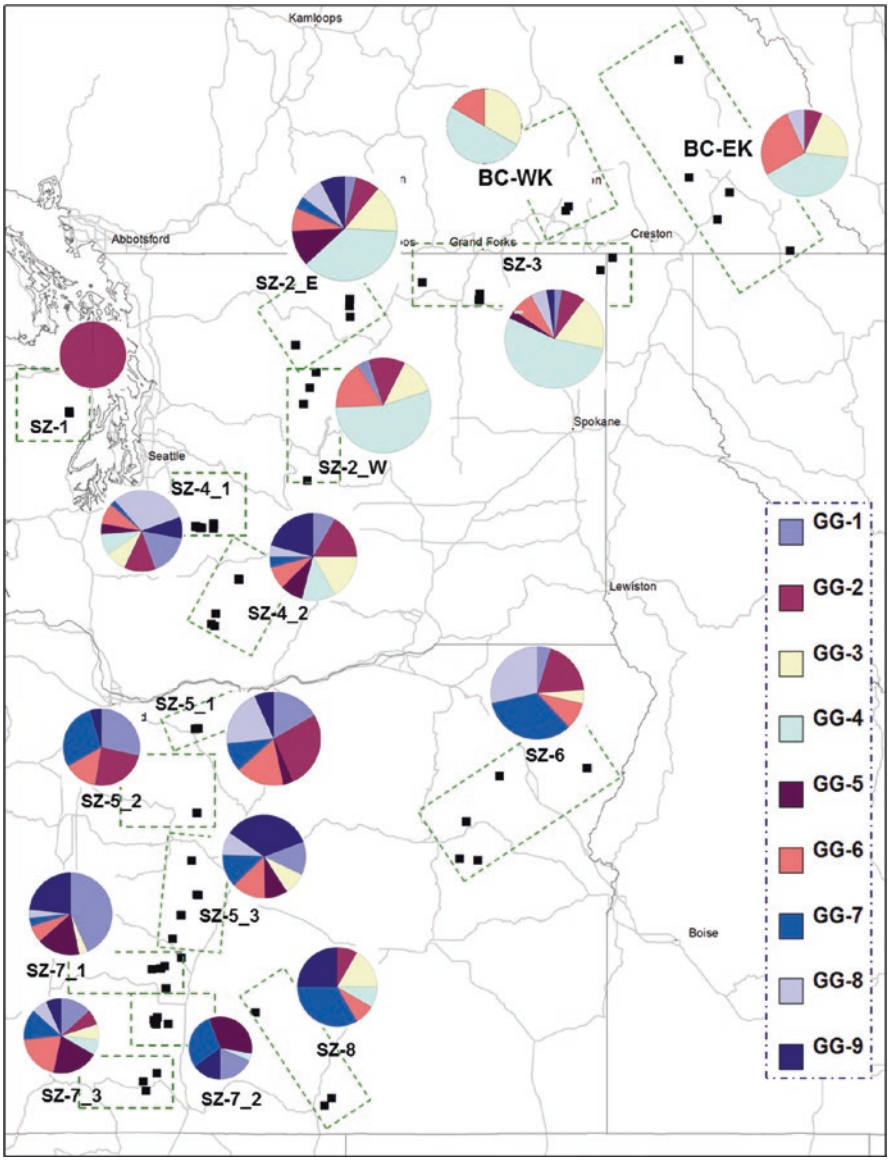


Fig. 13.11 Locations of *Pinus albicaulis* seed families and geographic distribution of genetic subgroups in the Canadian province of British Columbia and the USA states of Oregon and Washington; each pie chart represents the proportion of genetic subgroups (GG-1 to GG-9) as identified by STRUCTURE analysis in a given area. (From Liu et al. 2016a)

Insects

Four species of *Tsuga* (hemlock) are found in North America: *T. heterophylla* and *T. mertensiana* in the west and *T. canadensis* and *T. caroliniana* in the east. The two eastern species are threatened by the woolly adelgid insect (*Adelges tsugae*) that was introduced from Asia in the 1920s. A number of neutral marker studies have been conducted, primarily in *T. canadensis*, to estimate the extent of genetic variation within and between populations across the species ranges (Zabinski 1992; Wang et al. 1997; Camcore 2006; Potter et al. 2008; Lemieux et al. 2011; Potter et al. 2012). These studies included isozymes, cpDNA, AFLP, and nDNA SSRs (Chap. 9). The most recent study by Potter et al. (2012) is the most comprehensive and will be described in more detail here.

Potter et al. (2012) analyzed 13 highly polymorphic nDNA SSR markers in 1180 trees across 60 populations of *T. canadensis*. The results were, in some cases, consistent with population genetic expectations and in other cases, quite unexpected. Averaged over the 60 populations, measures of genetic diversity were moderate, but variation among populations was notable (A = 4.90; P = 97.2%; H_o = 0.526; H_e = 0.566; Fig. 13.12). The inbreeding coefficient (F_{IS}) averaged 7.3% but ranged from a negative 15.2% (excess of heterozygotes) to a positive 26.6%, the latter being rather high. The mean F_{ST} across loci was 0.077 but the among-population differentiation estimated using Jost's D (Jost 2008) was considerably higher (0.134) suggesting considerable variation exists among populations throughout the range.

As expected, peripheral disjunct populations were less genetically diverse than populations located in the central portions of the species range, but some were highly genetically differentiated. Unexpectedly, populations in formerly glaciated regions were no less genetically diverse than those in the putative southern refugial regions. In fact, allelic richness showed centers of diversity both north of the last glacial front and near the hypothesized primary refugial region, based on genetic and fossil evidence, in the southern Appalachian Mountains. These results led the authors to speculate that a glacial refuge may also have existed on the currently submerged continental shelf south of New England.

To better understand the population structure and inferred ancestry of eastern hemlock, the authors subjected population allelic information to two spatially explicit Bayesian clustering analyses (TESS 2.3.1, Chen et al. 2007; Structure 2.3.3, Pritchard et al. 2000). Though the results of the two analyses differed slightly, they both identified distinct genetic clusters that support the theory that the species expanded rapidly into its current range from multiple Pleistocene glacial refugia (Fig. 13.13). Overall, the authors suggest the most parsimonious explanation for the observed patterns of variation in *T. canadensis*, including the high levels of inbreeding observed across the species range, incorporates a well-documented sudden and drastic decline in abundance throughout the range about 5000 years ago, likely precipitated by a large-scale pathogen or insect outbreak. Such a decline would have resulted in an extreme genetic bottleneck with trees surviving in rare but widespread populations. The story in nuanced and fascinating, and the results obtained are helpful in making decisions regarding conservation efforts (Jetton et al. 2013; Hastings

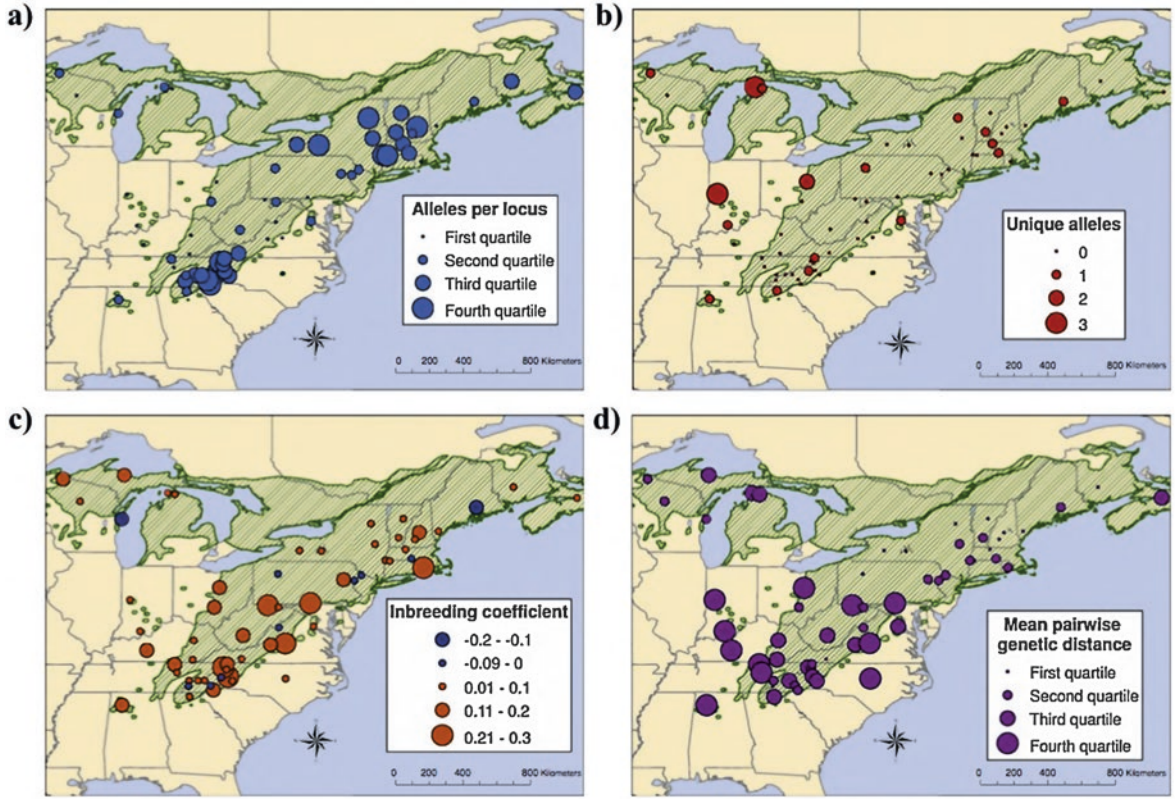


Fig. 13.12 *Tsuga canadensis* population classifications of (**a**) alleles per locus (A), (**b**) unique alleles (A*U*), (**c**) inbreeding coefficient (F_{IS}), and (**d**) mean pairwise chord distance (D_C), based on 13 polymorphic nuclear microsatellite loci. (From Potter et al. 2012)

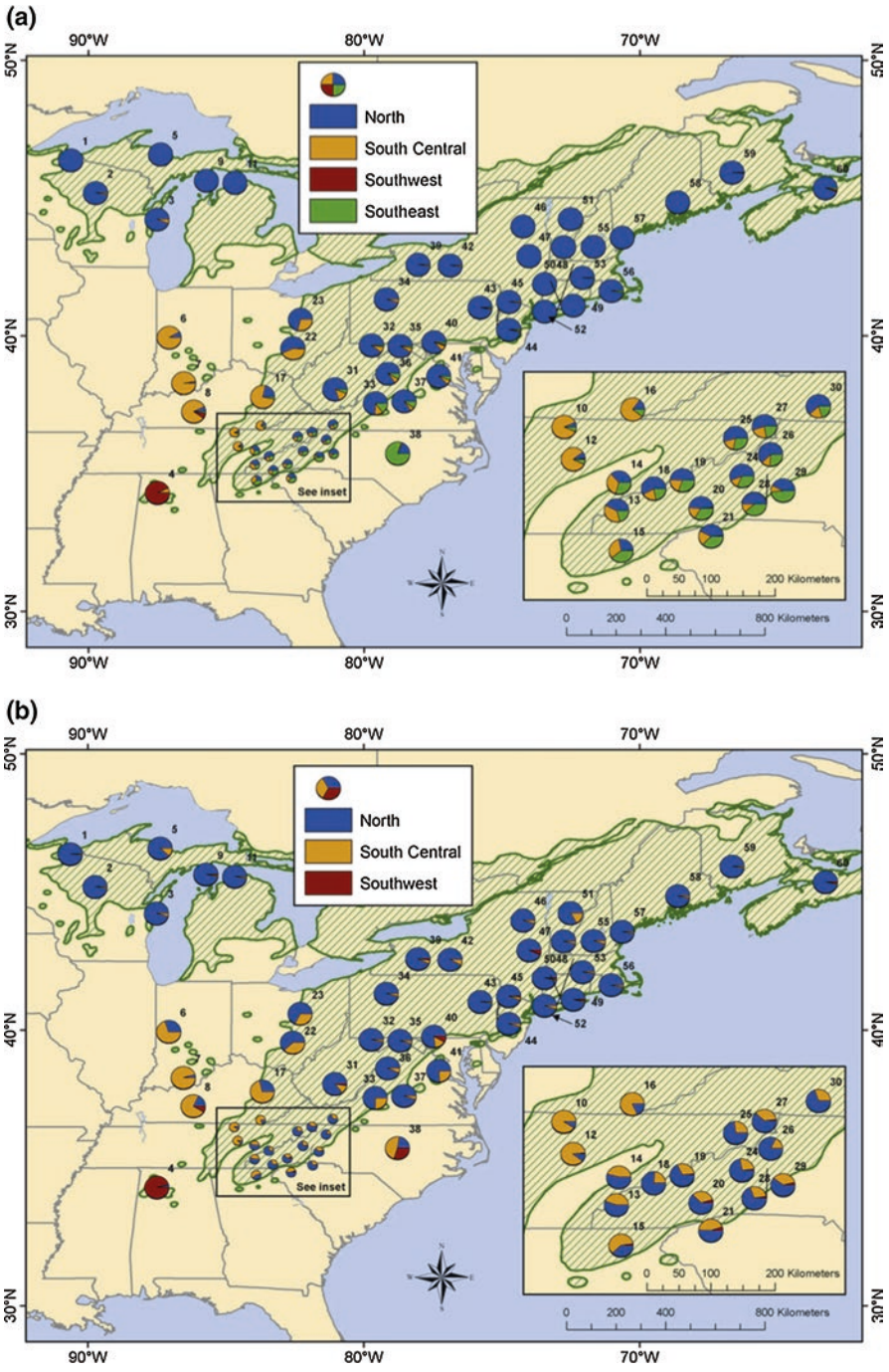


Fig. 13.13 The proportion, within each *Tsuga canadensis* population, of inferred ancestry from the genetic clusters inferred using (**a**) TESS2.3.1, Chen et al. (2007) and (**b**) Structure 2.3.3, Pritchard et al. (2000). (From Potter et al. 2012)

et al. 2017). Notably, the authors recommended that efforts should emphasize the capture of broad adaptability that occurs across the species range, as well as variability within regions with the highest allelic richness and heterozygosity. Sampling within disjunct populations that are genetically distinct was also warranted. However, as we have noted several times in this chapter, all conservation genetics recommendations are based on neutral genetic marker studies, thus information is still lacking as to which genetic material is best to capture toward maintaining adaptive evolutionary potential.

Climate Change

Climate change will affect nearly all living things on Earth including forest trees. Widespread conifer species in temperate regions will undergo range shifts and local extirpation but will avoid extinction (Aitken et al. 2008). However, species with highly fragmented ranges, and especially those found at higher elevations, may be at risk of extinction. The spruces found in Mexico are exemplary in this regard and are featured here.

Three species of genus *Picea* are found only in Mexico, *P. chihuahuana* (Chihuahua spruce), *P. mexicana*, and *P. martinezii* (Martinez spruce). Only recently have these three species been confirmed to be separate species (Ledig et al. 2004). Two (*P. mexicana* and *P. martinezii*) are extremely rare and found in only a couple populations each, whereas *P. chihuahuana* is found in 40 or more populations. Diversity studies using neutral markers have been done in all three species and have been summarized in several review papers (Ledig 2012; Quiñones-Pérez et al. 2014a; Wehenkel et al. 2017) (Table 13.7).

Results obtained from isozyme studies in *P. mexicana* and *P. martinezii* (Ledig et al. 2000) were quite similar; H_e was low ~0.10 in both cases, but not extremely low, and estimates of F_{IS} ranged from −0.107 to 0.121, indicating that inbreeding was not a problem in either of these small populations. This was unexpected as both species had very high selfing rates, $t_m < 0.50$ in some cases. This indicates that selection against self-pollinated offspring must be very strong to limit inbreeding, and thus genetic risk, to these species. The conclusion reached for both species is that they are likely at greater risk due to factors such as fire or deforestation than they are to genetic risk under static climate conditions but may be at great risk under changing climate. Even though genetic diversity estimated from neutral markers was not extremely low, the magnitude of adaptive evolutionary potential in these species is unknown and should be a research priority.

There have been several genetic diversity studies in *P. chihuahuana.* Ledig et al. (1997) studied 10 populations using 24 isozyme loci and found quite low estimates of H_e that generally tracked population size (Table 13.8). Inbreeding (F_{IS}), however, was not high. Pairwise F_{ST} values among populations were quite high but were not correlated with geographic distance suggesting that populations are not exchanging genes. Selfing rates in some populations were extremely high. The *P. chihuahuana* populations are restricted to north-facing slopes of steep-walled arroyos, so this

Table 13.7 Genetic diversity in two endangered (IUCN status EN, IUCN 2013) Mexican species of *Picea*

Species	MM	PP	P	I	L	Genetic diversity					References
						H_e	% *P*	*H*	δT	$v_{mean,2}$	
Picea chihuahuana	mtDNA	16	16	156	16	–	–	0	–	–	Jaramillo-Correa et al. (2006)
	cpSSR	6	16	156	6	–	–	0.41	–	–	Jaramillo-Correa et al. (2006)
	AFLP	1	14	669	243	–	–	–	–	1.5	Simental-Rodriguez et al. (2014)
	AFLP	1	5	254	319	–	–	–	0.3	–	Wehenkel and Sáenz-Romero (2012)
	Iso	–	10	164	24	0.09	27	–	–	–	Ledig et al. (1997)
Picea martínezii	Iso	–	2	54	22	0.11	32	–	–	–	Ledig et al. (2000)
	Iso	–	3	82	18	0.12	35	–	–	–	Ledig et al. (2002)

From Wehenkel et al. (2017)

MM molecular marker, *PP* number of primer-pair combinations, *P* number of populations, *I* number of individuals included per populations, *L* number of loci, H_e expected heterozygosity, *H* total haplotypic diversity, % *P* percent polymorphic loci, δT Gregorius' total differentiation, $v_{mean,2}$ mean genetic diversity, *mtDNA* mitochondrial DNA, *cpSSR* chloroplast microsatellite, *AFLP* amplified fragment length polymorphism, *Iso* isozymes

Table 13.8 Genetic diversity in *Picea chihuahuana* in northern Mexico: average sample size (n), mean expected heterozygosity (H_e, unbiased estimate), observed heterozygosity (H_o), percent polymorphic loci (P; absolute and 95% criteria), number of alleles per locus (A), fixation index or mean deviation from Hardy-Weinberg equilibrium (F), population census (N), and effective population size (N_e) calculated as explained in the text; standard errors appear in parentheses

Population	n	H_e	H_o	P		A	F	N	N_e
				100	95				
Arroyo de la Pista	21.6	0.114	0.09	33	21	1.5	0.091	–	–
	−0.7	−0.048	−0.038						
Arroyo del Infierno	14.6	0.055	0.054	21	17	1.2	−0.042	36	23
	−0.2	−0.027	−0.029						
Cebollitas	23.5	0.076	0.045	25	25	1.3	0.343	–	–
	−0.8	−0.032	−0.025						
Arroyo de al Quebrada	17.9	0.08	0.063	33	29	1.4	0.151	–	–
	−0.5	−0.03	−0.029						
Rio Vinihueachi	18.3	0.131	0.103	42	33	1.6	0.179	2441	–
	−0.4	−0.044	−0.037						
Talayotes	7.8	0.094	0.094	25	25	1.4	−0.077	377	59
	−0.2	−0.039	−0.039						
Cerro de la Cruz	11.4	0.066	0.028	25	17	1.3	0.414	17	29
	−0.3	−0.029	−0.029						
El Realito	14.2	0.127	0.09	29	29	1.5	0.197	587	518
	−0.2	−0.048	−0.034						
La Tinaja	14	0.071	0.071	17	17	1.2	−0.059	120	33
	0	−0.034	−0.035						
Arroyo Ancho	20.7	0.115	0.093	21	21	1.3	0.164	683	154
	−0.4	−0.047	−0.038						
Mean		0.093	0.073	27	23	1.37			

From Ledig et al. (1997)

specialized local adaptation and low amounts of genetic diversity estimated from neutral markers suggests that these populations maybe at great risk to climate change. As always, better data on adaptive evolutionary potential are needed.

Jaramillo-Correa et al. (2006) conducted a large demographic study (16 populations) using mtDNA and cpDNA markers. The mtDNA markers divided the populations into two major groups (north and south), whereas the cpDNA markers revealed six haplotypes that were mostly clinally distributed (Fig. 13.14). These patterns are consistent with postglacial migration of seed and pollen, but do not provide much information related to adaptive genetic potential under changing climate.

Wehenkel and coworkers used AFLP markers in a series of studies to search for a relationship between neighboring tree diversity and AFLP diversity in *P. chihuahuana* (Quiñones-Pérez 2014b; Simental-Rodríguez et al. 2014) and also for evidence of genetic erosion (Wehenkel and Sáenz-Romero 2012; Quiñones-Pérez et al. 2017). The neighboring tree diversity with *P. chihuahuana* AFLP diversity was not strong and purely correlative. Likewise, there was very little evidence for genetic erosion between mature trees and regeneration. AFLP markers, like RAPD markers before, have fallen out of use due to issues of repeatability, so it would seem that

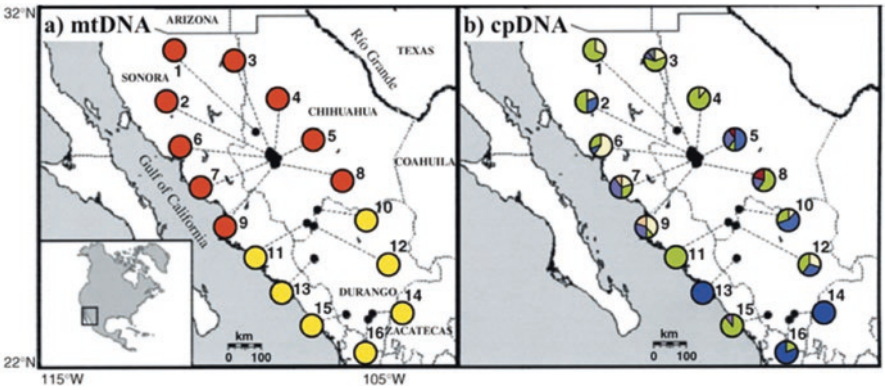


Fig. 13.14 Geographic distribution of mitochondrial DNA (**a**) and chloroplast DNA haplotypes (**b**) in populations of *Picea chihuahuana* in northern Mexico. (From Jaramillo-Correa et al. 2006)

these results are quite preliminary and in need of replication before any conservation genetics recommendations could be made.

The overall summary from this body of work in *P. chihuahuana* using neutral genetic markers is that there remains genetic variation in isolated populations and even though selfing can be very high, selection acts against selfed progeny to control inbreeding. Therefore, under static environments these populations may not be at great risk, but under changing environments the ability to migrate naturally is very limited and the adaptive genetic potential in place is unknown.

Summary

Conservation genetics has been a long-standing topic of interest for conifers that is growing in importance considering the many anthropomorphically induced causes impacting adaptive evolutionary potential of conifer populations (deforestation and overharvesting, fragmentation, habitat loss, forest practice, disease, insects, pollution, invasive species, hybridization, and climate change). Two types of data have routinely been obtained from populations to assess these impacts: (1) neutral marker data and (2) quantitative trait variation from common gardens. Neutral marker data have been very informative for revealing patterns of demographic processes (genetic drift, migration, mating system) and has also been used to infer adaptive potential. Correlations with adaptive genetic variation and neutral variation may exist but this relationship may turn out to be weak or even incorrect. Genomic approaches to understanding patterns of adaptive genetic variation are just beginning and should ultimately provide useful information to guide conservation genetics programs. Information from common garden tests is highly useful but these data are generally found from only a small number of widespread and commercially important conifers as part of tree-breeding activities and are less frequently available from rare and noncommercial species. In addition, common garden data take a long time to obtain and are increasingly very expensive to acquire.

Forest Health

14

Introduction

The conifers, in general, retain a great deal of genetic variation for all manner of traits, from the molecular level to whole-plant phenotypes, and everything in between. It is also true that species vary widely in the amount and distribution of variation they possess for specific traits, and some, like those with very restricted ranges or those that have been forced through genetic bottlenecks, may be severely limited for variation in some or all traits. These conditions apply equally to traits associated with forest health, the consequences of which are becoming increasingly important to the world's forests, and our reliance on the resources they provide.

In this chapter, we have attempted to capture and cite a significant cross-section of the scientific literature that identifies genetic variation within and among species with respect to resistance or tolerance to biotic and abiotic agents. This includes challenges from insect and disease organisms, browsing animals like voles or deer, and physiological stresses imposed by drought, cold, saline soils, or air pollution. Our effort will be organized around the types and approaches of studies conducted, the causal organisms or abiotic agents involved, and the mechanisms conifers employ to resist or tolerate the stressors.

The Growing Relevance of Forest Health

Dedicating an entire chapter to forest health is driven, in large part, by the measurable and often alarming declines in forest productivity noted in recent decades, particularly, though not exclusively, in North America (FAO 2009; Shifley and Moser 2016; United States Dept. of Agriculture, Forest Service, Forest Health Protection 2017). In 2006, almost 8% of forests in the United States (approximately 58 million acres) were at significant risk from insect and disease mortality (Alvarez 2007). A more recent illustration of the extent to which forest watersheds in the United States are threatened is striking (Fig. 14.1). Conditions in Canada are substantially similar, driven primarily by massive bark-beetle outbreaks (Natural Resources Canada 2017).

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_14

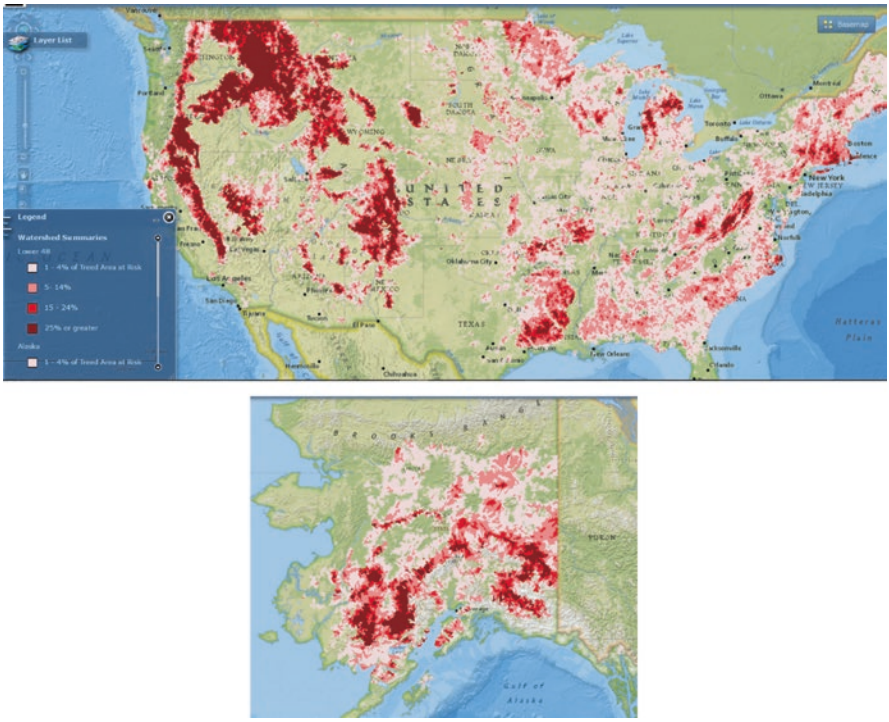


Fig. 14.1 A projected map view, by drainage, of the national insect and disease risk in the United States for the years 2013–2027. (United States Department of Agriculture, Forest Service, Forest Health Protection 2017)

Though forests have always experienced damage or mortality due to biotic and abiotic factors, the level of damage appears to have increased measurably over the last 30 years (Hayden et al. 2011). Increased forest health issues are often attributed to a series of factors, most notably changing global climates and the introduction of alien pests and diseases. Less serious are the introduction of alien weedy species, forest fragmentation caused by development, and long-term forest management treatments like fire control. The dual threats of climate change and introduced forest pests and diseases are already changing structure and function of forest landscapes. Changing climates, exemplified by dramatic shifts in temperature and precipitation patterns, and increased frequency of extreme events, can profoundly affect both abiotic (i.e., drought indices) and biotic agents. For instance, massive outbreaks of bark beetles in conifer forests throughout North America are attributed in part to warmer winter temperatures that do not kill insect broods, increased drought that weakens trees' existing defenses, and decades of fire protection management that has resulted in overstocked and less thrifty stands. Climate change affects all forest tree species and their ecosystems to some degree and fundamentally alters host-pathogen interactions.

Introduced pests that have not coevolved with host species find native species often lack natural genetic resistance or sources of resistance are rare and unevenly distributed across the landscape. The results of such introductions are often catastrophic and have been well chronicled over time (Campbell and Schlarbaum 1994, 2002, 2014). Within the last 100 years, the United States alone has witnessed near extirpation of several important hardwood tree species including *Castanea dentata* (American chestnut), *Ulmus americana* (American elm) and related elm species, and *Juglans cinerea* (butternut), all of which have been victimized by introduced diseases from Asia. Currently, most species of *Fraxinus* (ash) in North America are being devastated by the introduced emerald ash borer (*Agrilus planipennis*). While the situation with conifers is less grim, it remains serious. Historically speaking, the greatest impact of an introduced pest on conifers in North America has been on the native white pines by *Cronartium ribicola,* the white pine blister rust. Once restricted primarily to *Pinus monticola* (western white pine), *P. lambertiana* (sugar pine), and *P. strobus* (eastern white pine), the disease has spread to all but one of the nine native white pine species in the United States and Canada in recent decades, aided by climate warming trends (Sniezko and Zambino 2005; Fig. 14.2). Fortunately, as of this writing, the disease has not appeared on Mexican white pine species or the long-lived *P. longaeva* (Great Basin bristlecone pine), though evidence from inoculation trials suggest all are susceptible to the disease.

Equally as disturbing are the results of foliar feeding by the introduced hemlock wooly adelgid (*Adelges tsugae*) on *Tsuga canadensis* (eastern hemlock) and *T. caroliniana* (Carolina hemlock). Elsewhere around the globe, host-pathogen

Fig. 14.2 Dead and dying *Pinus albicaulis* (whitebark pine) in the Wind River Range of the Rocky Mountains, Wyoming, USA. Trees are succumbing to the combined assaults of white pine blister rust (*Cronartium ribicola*), the mountain pine beetle (*Dendroctonus ponderosae*), and drought. (Photo courtesy of Dr. Bruce Bongarten, Emeritus Professor, SUNY-ESF))

interactions have been altered by the large-scale introduction of nonnative conifer species for commercial purposes. These plantations have been under assault from both native and nonnative pathogens and pests (Wilcox 1982; Kobayashi and Muramoto 1989; Gordon et al. 1998; Wingfield et al. 2002, 2008; Aegerter and Gordon 2006; Donoso et al. 2015).

Genetic Variation in Forest Health Traits

Insects and Disease

Studies identifying heritable variation in forest health traits are numerous and the literature describing their results robust. A survey focused primarily on resistance to insect and disease pests conducted in conjunction with the UN FAO (United Nations Food and Agriculture Organization) identified 275 references, roughly half of which were specific to conifers (Yanchuk and Wheeler 2008). That document, available online, was used as the foundation for the significantly updated list of conifer citations noted here (Table 14.1). All citations listed are identified by host and pathogen name and are categorized by the approach taken to gather results and the status of the research program that conducted the study. Approaches included (1) traditional plant breeding methods like clonal, provenance, or progeny testing; (2) molecular biology approaches like QTL and association mapping, sequencing and gene annotation, and gene expression studies; and (3) genetic engineering. Status is a bit of a hybrid term that refers both to the methods used to detect variation and the applied purpose of the study. Status levels were defined as (1) genetic variation in resistance detected in seedling or clonal screens; (2) genetic variation in resistance detected in genetic or provenance trials; (3) genetic variation in resistance detected in programs breeding for resistance, but no materials have yet to be deployed; and (4) genetic variation in resistance detected in breeding programs with resistant materials deployed in the field operationally. Pest types are denoted as disease (D), insect (I), mammal (M), or nematode (N).

Since much of the information found in Table 14.1 was gathered through voluntary submittal of survey data, it is difficult to be current with status level for programs. Therefore, data noted here should already be considered dated. Still, the information reveals strong tendencies. Most studies identifying resistance rely on traditional approaches and low intensity screens. Use of molecular approaches to identify and characterize variation in resistance is less common than other methods, overall, but constitutes an increasingly significant proportion of studies being conducted today. Notably, such studies can rapidly and accurately identify the genetic basis of resistance leading to more efficient screening and deployment strategies. These methods are amenable to all species, not simply those few considered to be of greatest commercial value and have found strong support for future funding by private and government agencies (Wheeler et al. 2015). Genetic modification has received very little consideration in conifers, despite its promise as an approach to deliver rapid results, due in part to both biological constraints and

Table 14.1 An enumeration of studies characterizing genetic diversity in disease and insect resistance in conifer taxa, listed by family, genus, and species. All citations listed are identified by host and pathogen name and are categorized by the approach taken to gather results and the status of the research program conducting the study. Codes for pest type, status, and approach are noted in table footnotes and are described in the chapter text

Genus	Species	Pest type[a]	Common name	Scientific name	Status[b]	Approach[c]	Literature cited
Cupressaceae							
Chamaecyparis	*lawsoniana*	D	Root rot	*Phytophthora lateralis* (root)	4	1	Kitzmiller and Sniezko (2000), Sniezko (2003, 2004), Sniezko and Hansen (2001, 2003), Sniezko et al. (2000b, 2003a, b, 2005, 2006a, 2012a, b)
Chamaecyparis	*obtusa*	D	Seiridium canker	*Seiridium unicorne*	1	1	Kato (1996)
Cryptomeria	*japonica*	I	Sugi bark borer	*Semanotus japonicus* (Coleoptera: Cerambycidae)	1	1	Kato and Taniguchi (2003)
Cryptomeria	*japonica*	I	Sugi bark midges	*Resseliella odai*	1	1	Fujimoto et al. (1983), Sasaki et al. (2003)
Cryptomeria	*japonica*	I	Sugi bark borer	*Semanotus japonicus*	1	1	Kato and Kawamura (1993), Ueki (2004)
Cryptomeria	*japonica*	I	Sugi bark borer	*Semanotus japonicus*	1	2	Nishiyama et al. (2002)
Cunninghamia	*lanceolata*	I & D			2	1	Yu (2000), Zhou and Ma (2000)
Cupressus	*lusitanica*	I	Aphid	*Cinara cupressi*	2	1	Kamunya et al. (1999)
Cupressus	*macrocarpa*	D	Canker	*Seiridium unicorne*	2	1	Gea and Low (1997)
Cupressus	*sempervirens*	D	Canker	*Seridium cardinale*	1	1	Santini et al. (1997)

(continued)

Table 14.1 (continued)

Genus	Species	Pest type[a]	Common name	Scientific name	Status[b]	Approach[c]	Literature cited
Thuja	*plicata*	D	Needle blight	*Didymascella thujina*	2	1	Russell et al. (2007)
Thuja	*plicata*	M	Deer		3	1	Russell (2004)
Pinaceae							
Abies	*alba*	I	Silver fir woolly aphid	*Dreyfusia (Adelges) nordmannianae Eckst.*	1	1	Beier-Petersen et al. (1974)
Abies	*balsamea*	I	Balsan twig aphid	*Mindarus abietinus*	1	1	Edwards et al. (2016)
Abies	*fraseri*	D	Root rot	*Phytophthora cinnamomi*	1	1	Potter and Frampton (2003), Frampton and Benson (2004)
Abies	*nordmanniana*	I	Silver fir woolly aphid	*Dreyfusia nordmannianae*	2	1	Nielsen et al. (2002)
Abies	*spp.*	I	Balsan woolly adelgid	*Adelges piceae*	1	1	In: http://www.fs.fed.us/psw/publications/documents/psw_gtr240/
Larix	*decidua*	D	Larch canker	*Lachnellula willkommii*	1	1	Pâques et al. (1999)
Larix	*decidua var. polonica*	D	Larch canker	*Lachnelluna willkommii*	2	1	Matras (2001)
Larix	*leptolepis* & hybrid, *gmelini* var. *japnica* & hybrid	M	Vole	*Clethrionomys rufocanus bedfordiae*	1	1	Chiba (1963), Chiba et al. (1982, 1989), Nagata et al. (1989)
Larix	*occidentalis*	D		*Meria laricis, Hypodermella laricis*	4	1	Rehfeldt (1992)

Larix	*occidentalis*	D	Honey mushroom	*Armillaria ostoyae*	2	1	Hoff and McDonald (1994)
Larix	spp.	I			1	3	Shin et al. (1994)
Larix	spp.	D	Needle cast, Shoot blight	*Mycosphaerella larici-leptolepis, Pysalospora laricina*	1	1	Chiba and Nagata (1963, 1972, 1978), Kohda et al. (1981)
Picea	*abies*	D	Root rot	*Heterobasidion annosum*	2	1	Wellendorf and Thomsen (2007)
Picea	*abies*	D	Root rot	*Heterobasidion annosum*	2	1	Karlsson and Swedjemark (2006)
Picea	*abies*	D	Root rot	*Heterobasidion annosum*	1	2	Hietala et al. (2003)
Picea	*abies*	D	Root rot	*Heterobasidium annosum*	1	1	Swedjemark et al. (1997)
Picea	*abies*	I	White pine weevil	*Pissodes strobi*	2	2	Larochelle et al. (1993)
Picea	*glauca*	I	White pine weevil	*Pissodes strobi*	4	1, 2	King et al. (1997), King and Alfaro (2004), Porth et al. (2011)
Picea	*glauca*	I	Spruce budworm	*Choristoneura fumiferana*	2	1, 2	Mageroy et al. (2015), Parent et al. (2017)
Picea	*glauca*	I	Spruce budworm	*Choristoneura fumiferana*	2	3	Lachance et al. (2007)
Picea	*glauca*	D	Wood decay	*Trametes versicolor/ Fomitopsis pinicola/ Gloeophyllum trabeum*	1	1	Yu and Yang (2003)

(continued)

Table 14.1 (continued)

Genus	Species	Pest type[a]	Common name	Scientific name	Status[b]	Approach[c]	Literature cited
Picea	*sitchensis*	I	Green spruce aphid	*Elatobium abietinum*	4	1	Harding et al. (2003)
Picea	*sitchensis*	I	Spruce aphid	*Elatobium abietinum*	1	1, 2	Skov and Wellendorf (2000)
Picea	*sitchensis*	I	White pine weevil	*Pissodes strobi*	4	1	Alfaro et al. (2000)
Picea	*sitchensis*	I	White pine weevil	*Pissodes strobi*	2	1, 2	Roach et al. (2014)
Pinus	*albicaulis*	D	White pine blister rust	*Cronartium ribicola*	1	1	Mahalovich (2004), Sniezko et al. (2007)
Pinus	*aristata*	D	White pine blister rust	*Cronartium ribicola*	2	1	
Pinus	*aristata*	D	White pine blister rust	*Cronartium ribicola*	1	1	In: http://www.fs.fed.us/psw/publications/documents/psw_gtr240/
Pinus	*armandii*	I			1	3	Liu et al. (2010)
Pinus	*contorta*	D	Western gall rust	*Endocronartium harknessii*	4	1	Mahalovich (1995)
Pinus	*contorta*	D	Tip blight	*Gremmeniella abietina*	1	1	Laflamme et al. (2006)
Pinus	*contorta*	I	Mountain pine beetle	*Dendroctonus ponderosae*	2	1	Yanchuk et al. (2008)
Pinus	*contorta*	D	Needle cast	*Lophodermium sp.*	2	1	Hunt (1981), Hunt et al. (1987), Rehfeldt (1987)
Pinus	*contorta*	D	Western gall rust, Stalactiform blister rust, Needle cast, Sequoia pitch moth	*Endocronartium, Cronartium, Lophodermella*	2	1	Hoff (1985), Wu and Ying (1997)
Pinus	*contorta*	I	Terminal weevil	*Pissodes terminalis*	2	2	

Pinus	*contorta var. bolanderi*	M	Possum		2	1	Miller and Ecroyd (1987)
Pinus	*densiflora*	N	Pine wilt/nematode	*Bursaphelenchus xylophilus*	1	1	Toda et al. (2002)
Pinus	*elliotti*	D	Fusiform rust	*Cronartium quercuum*	4	1	Lopez-Upton et al. (2000)
Pinus	*elliotti*	D	Tip blight	*Diplodia pinea, Lecanosticta acicola*	2	2	Yi et al. (2000), Wang et al. (1998)
Pinus	*elliotti*	D	Brown-spot needle blight	*Lecanosticta acicola*	1	1	In: http://www.fs.fed.us/psw/publications/documents/psw_gtr240/
Pinus	*elliotti*	D	Fusiform rust	*Cronartium quercuum*	1	1	Bridgwater et al. (2003)
Pinus	*flexilis*	D	White pine blister rust	*Cronartium ribicola*	2	1	Vogan and Schoettle (2015)
Pinus	*hybrids*	I	Pine needle gall midge	*Thecodipdiplosis japonicus*	1	1	Son et al. (1999)
Pinus	*lambertiana*	D	White pine blister rust	*Cronartium ribicola*	3	1	Harkins et al. (1998), Kitzmiller (1976), Kinloch (1982, 2003), Kinloch and Dupper (2002), Kinloch et al. (2004)
Pinus	*lambertiana*	D	White pine blister rust	*Cronartium ribicola*	1	2	Devey et al. (1995)
Pinus	*lambertiana*	D	White pine blister rust	*Cronartium ribicola*	1	2	Harkins et al. (1998)
Pinus	*lambertiana*	D	White pine blister rust	*Cronartium ribicola*	1	2	Jermstad et al. (2006)
Pinus	*leiophylla*	D	Pitch canker	*Fusarium circinatum*	2	1	Dvorak et al. (2007)

(continued)

Table 14.1 (continued)

Genus	Species	Pest type[a]	Common name	Scientific name	Status[b]	Approach[c]	Literature cited
Pinus	*massoniana*	I	Pine wilt/nematode	*Bursaphelenchus xylophilus*	2	1	In: http://www.fs.fed.us/psw/publications/documents/psw_gtr240/
Pinus	*monticola*	D	White pine blister rust	*Cronartium ribicola*	4	1	Hunt (2004), Hunt and Jensen (2001)
Pinus	*monticola*	D	White pine blister rust	*Cronartium ribicola*	4	1	Bingham et al. (1969), Bingham (1983), Goddard et al. (1985), Hoff (1986), Hoff and McDonald (1971, 1993)
Pinus	*monticola*	D	White pine blister rust	*Cronartium ribicola*	4	1	Fins et al. (2002)
Pinus	*monticola*	D	White pine blister rust	*Cronartium ribicola*	4	1	Kinloch et al. (1999, 2003), Sniezko (2003, 2004), Sniezko et al. (2000a, 2004a, b, 2006b, 2008, 2011, 2014), Sniezko and Kegley (2003a, b, 2005), Sniezko and Zambino (2005)
Pinus	*monticola*	D	White pine blister rust	*Cronartium ribicola*	4	2	Liu et al. (2011, 2013)
Pinus	*monticola*	D	Needle blight	*Lecanosticta*	2	1	Hoff and McDonald (1978)
Pinus	*nigra*	D	Needle blight	*Dothistroma pini*	2	1	Wilcox and Miller (1975), Wheeler et al. (1976)
Pinus	*nigra*	D	Canker	*Cenangium ferruginosum*	2	1	Wheeler et al. (1976)
Pinus	*nigra*	I	Zimmerman pitch moth	*Dioryctria zimmermani*	2	1	Wheeler et al. (1976)
Pinus	*nigra*	I	Black-headed pine sawfly	*Neodiprion sertifer*	2	1	Wheeler et al. (1976)

Pinus	*nigra*	D	Tip blight	*Sphaeropsis sapinea*	2	1	
Pinus	*palustris*	D	Brown spot	*Scirrhia acicola*	1	1	Gwaze et al. (2003)
Pinus	*pinaster*	D	Pitch canker	*Fusarium circinatum*	2	1	Elvira-Recuenco et al. (2014)
Pinus	*pinaster*	D	Root rot	*Armillaria ostoyae*	1	1	Zas et al. (2007)
Pinus	*pinaster*	I	Moth	*Dioryctria sylvestrella*	1	1	Kleinhentz et al. (1998)
Pinus	*pinaster*	I & D		*Fusarium, Armilaria, Thaumetopoea, Hylobius, Dioryctria*	2	1	In: http://www.fs.fed.us/psw/publications/documents/psw_gtr240/
Pinus	*ponderosa var. ponderosa*	D	Western gall rust	*Endocronartium harknessii*	4	1	Mahalovich (2003)
Pinus	*ponderosa* var. *ponderosa*	I	Tip moth	*Rhyacionia sp.*	2	1	Kegley et al. (1994)
Pinus	*ponderosa* var. *ponderosa*	D	Mistletoe	*Arceuthobium sp.*	2	1	Roth (1974), Scharpf and Roth (1992)
Pinus	*ponderosa* var. *ponderosa*	D	Needle & tip blight	*Dothistroma/ Diplodia*	2	1	Burdon and Low (1991)
Pinus	*ponderosa* var. *ponderosa*	D	Needle cast	*Lophodermium sp.*	2	1	Hoff (1988b)
Pinus	*ponderosa* var. *ponderosa*	I	Gouty pitch midge	*Cecidomyia piniinopis*	2	1	Hoff (1989), Hoff (1988a)

(continued)

Table 14.1 (continued)

Genus	Species	Pest type[a]	Common name	Scientific name	Status[b]	Approach[c]	Literature cited
Pinus	*radiata*	D	Needle blight	*Dothistroma pini*	4	1	Carson (1990), Carson and Carson (1991)
Pinus	*radiata*	D	Tip blight	*Diplodia pinea, (Sphaeropsis sapinea)*	3	1	Burdon et al. (1982)
Pinus	*radiata*	D	Pitch canker	*Fusarium circinatum*	2	1	Devey et al. (2001), Bonello et al. (2001), Matheson et al. (2006)
Pinus	*radiata*	D	Tip blight	*Diplodea pinea*	2	1	Arregui et al. (1999)
Pinus	*radiata*	D	Tip blight	*Sphaeropsis sapinea*	2	1	
Pinus	*radiata*	D	Needle blight	*Dothistroma septospora*	1	2	Devey et al. (2004b)
Pinus	*radiata*	D	Needle blight	*Dothistroma septospora*	1	1	Burdon et al. (1997)
Pinus	*radiata*	D	Needle cast	*Cyclaneusma minus*	1	1	Burdon et al. (1997)
Pinus	*radiata*	D	Rust	*Endocronartium harkinessii*	1	1	Old et al. (1986)
Pinus	*radiata*	D	Tip blight	*Diplodia pinea*	1	1	Burdon and Bannister (1973)
Pinus	*radiata*	D	Pitch canker	*Fusarium circinatum*	1, 2	1	Aegerter and Gordon (2006), Vivas et al. (2012), Donoso et al. (2015)
Pinus	*radiata*	I	Pine aphid	*Esigella californica*	2	1	Sasse et al. (2009)
Pinus	*radiata*	I	Large pine weevil	*Hylobius abietis*	1	1	Zas et al. (2007)
Pinus	*radiata*	I	Pine woolly aphid	*Pineus pini*	1	1	Burdon et al. (1998)
Pinus	*radiata*	I			1	3	Grace et al. (2005)

Pinus	*radiata*	I & D			1	1	Simpson and Ades (1990)
Pinus	*radiata*	I	Painted apple moth	*Teia anartoides*	1	3	Grace et al. (2005)
Pinus	*radiata*	I	Pine aphid	*Essigella*	1	1	
Pinus	spp	D	Pitch canker	*Fusarium circinatum*	1, 2	1, 2	Hodge and Dvorak (2000), Smith et al. (2003)
Pinus	spp	I	Cone worm	*Dioryctria amatella*	1	1	Sartor and Neel (1971)
Pinus	spp	I	Tip moth	*Rhyacionia frustrana*	1	1	Highsmith et al. (2003)
Pinus	*strobiformis*	D	White pine blister rust	*Cronartium ribicola*	3	1	Sniezko et al. (2006)
Pinus	*strobus*	D	White pine blister rust	*Cronartium ribicola*	2, 3	1	Lu et al. (2005), Lu and Derbowka (2009)
Pinus	*strobus*	I	White pine weevil	*Pissodes strobi*	2	1	Wilkinson (1983)
Pinus	*strobus*	D	White pine blister rust	*Cronartium ribicola*	2	1	Jurgens et al. (2003), Smith et al. (2005)
Pinus	*strobus x peuce*	D	White pine blister rust	*Cronartium ribicola*	1	1	Blada (2000)
Pinus	*sylvestris*	D	Pinetwist rust	*Melampsora pinitorqua*	1	1	Quencez and Bastien (2001)
Pinus	*sylvestris*	I	Shoot beetle, Weevil	*Tomicus, Pissodes*	1	1	Schneck (1992)
Pinus	*sylvestris*	D	Brown rot	*Coniophora puteana*	1	1	Harju et al. (2001)
Pinus	*tabulaeformis*	I	Pine bast scale	*Matsucoccus matsumurae*	1	1	Zhao and Li (1989)
Pinus	*taeda*	D	Fusiform rust	*Cronartium quercuum*	3	1	Huber and Amerson (2011)

(continued)

Table 14.1 (continued)

Genus	Species	Pest type[a]	Common name	Scientific name	Status[b]	Approach[c]	Literature cited
Pinus	*taeda*	D			2	1	Zhao (2000)
Pinus	*taeda*	D	Fusiform rust	*Cronartium quercuum*	2	1, 2	Isik et al. (2012)
Pinus	*taeda*	D	Pitch canker	*Fusarium circinatum*	2	1, 2	Quesada et al. (2010)
Pinus	*taeda*	D	Fusiform rust	*Cronartium quercuum*	1	1	Kraus (1986)
Pinus	*taeda*	D	Pitch canker, Fusiform rust	*Fusarium circinatum, Cronartium quercuum.*	1	2	Kayihan et al. (2005)
Pinus	*taeda*	I	Pine caterpillar	*Dendrolimus punctatus, Crypyothelea formosicola*	1	3	Tang and Tian (2003)
Pinus	*taeda*	D	Fusiform rust	*Cronartium quercuum*	4	1	Wilcox et al. (1996), Young (2003), Kuhlman (1997), Isik et al. (2003), Lambeth (2000), Sierra-Lucero et al. (2002), Amerson et al. (2015)
Pinus	*taeda*	D	Fusiform rust	*Cronartium quercuum*	1	2	Morse et al. (2004)
Pinus	*taeda*	D	Fusiform rust	*Cronartium quercuum*	1	2	Zeng et al. (2004)
Pinus	*taeda*	D			1	2, 3	Juang et al. (2003)
Pinus	*taeda*	I	Bark beetles	*Dendroctonus*, *Ips*, etc.	2	1, 2	Westbrook et al. (2013)
Pinus	*thunbergii*	I	Pine needle gall midge	*Thecodiplosis japonensis*	4	2	Kondo et al. (2000), Hayashi et al. (2004)

Pinus	*thunbergii*	I	Pine needle gall midge	*Thecodiplosis japonensis*	4	1	Terada (1992)
Pinus	*thunbergii*	I	Pine needle gall midge	*Thecodiplosis japonensis*	4	1	Kim et al. (1978)
Pseudotsuga	*menziesii*	D	Needle cast	*Phaeocryptopus gaeumannii*	3	1	Temel et al. (2004, 2005), Johnson (2002)
Pseudotsuga	menziesii	D	Swiss Needle Cast		3	1	Dungey et al. (2012)
Pseudotsuga	menziesii	D	Swiss Needle Cast		3	1	In: http://www.fs.fed.us/psw/publications/documents/psw_gtr240/
Pseudotsuga	*menziesii*	D	Needle cast	*Rhabdocline sp.*	2	1	Hoff (1987), Stephan (1973)
Pseudotsuga	*menziesii*	D	Needle cast	*Phaeocryptopus gaeumannii*	1	1	Kimberley and Hood (2005)
Pseudotsuga	*menziesii*	D	Root rot	*Armillaria ostoyae*	1	1	Cruickshank et al. (2006), Robinson and Morrison (1994)
Pseudotsuga	*menziesii*	D	Root rot	*Phellinus sulphurascens, P. weirii*	1	12	Ekramoddoullah et al. (2000), Sturrock et al. (2007), Sturrock (2005)
Pseudotsuga	*menziesii*	I	Western spruce budworm	*Choristoneura occidentalis*	2	1	McDonald (1985)
Tsuga	*canadensis*	I	Hemlock woolly adelgid	*Adelges tsugae*	1, 2	1	McKenzie et al. (2014), Oten et al. (2014)

[a] *I* insect, *D* disease, *M* mammal, *N* nematode

[b] 1 = Genetic variation in resistance detected in seedling or clonal screens; 2 = Genetic variation in resistance detected in genetic or provenance trials; 3 = Genetic variation in resistance detected in programs breeding for resistance, but no materials have yet to be deployed, and 4 = Genetic variation in resistance detected in breeding programs with resistant materials deployed in the field operationally

[c] 1 = traditional plant breeding methods like clonal, provenance, or progeny testing; 2 = molecular biology approaches like QTL and association mapping, sequencing and gene annotation, and gene expression studies; 3 = genetic engineering

lack of social acceptance. New technologies such as CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) may find more acceptance, but deployment of the technology for forest health issues has only just begun. In coming decades, genomic resources and tools developed from them have the potential to greatly increase the efficiency of finding and exploiting genetic resistance to insects and diseases, likely increasing the number of species used in restoration and reforestation efforts (Chap. 13).

Despite decades of research in this area, genetic resistance has been deployed relatively infrequently in conifers (12 programs for disease organisms and 6 for insects) and includes very few tree species and pathogens, evidence of the difficulty and costs associated with breeding of conifers. Most studies looked at resistance to diseases (ca. 60%), very few to mammals, nematodes, and the like, and the remainder to insect pests. Some studies included both insects and disease organisms. The most commonly studied conifer genus was *Pinus* and the most studied species were *Pinus radiata*, *P. taeda*, *P. monticola*, and *P. ponderosa*.

For those seeking current operationally oriented reports on host tree-pest interaction researches, the periodic proceedings of symposia on the topic provide a wealth of information (Sniezko et al. 2012c; Jorge et al. 2015).

Abiotic Stress

Forest geneticists have long been aware that most conifer species retain large reservoirs of genetic variation in tolerance to environmental stress factors such as drought or cold, especially the latter. Cold temperatures are often cited as the limiting factor controlling northward expansion of conifer natural ranges (in the Northern Hemisphere; southward expansion in the Southern Hemisphere), and drought is the leading cause of seedling mortality in most temperate forests of the southeastern and southwestern United States (Bigras et al. 2001; Hänninen et al. 2001; Lorenz et al. 2006; Kolb et al. 2016). For decades, variation in cold and drought tolerance has been evaluated in nursery and field trials, as chronicled in Chap. 8 of this volume (see Table 8.2). These genecological studies often demonstrated the existence of steep genetic clines, especially for cold adaptation traits, in relation to environmental gradients (Howe et al. 2003). Heritability of cold adaptation traits varies from weak to strong with phenological traits, such as budburst or budset, being under the strongest genetic control (Anekonda et al. 2000; Howe et al. 2003). Consequently, phenological traits have often been used as surrogates for estimating cold tolerances, but they may be equally as useful for estimating drought tolerance. Trees with early budburst are typically susceptible to early season frosts, but are favored in areas prone to drought (deep, well-drained soils) because they complete their summer growth cycle before drought sets in. Trees leafing out late avoid spring frosts, but may suffer from early fall frosts, or late summer drought, both of which are common in Mediterranean-type climates around the world.

Increasingly, investigators have eschewed reliance on surrogate traits and relied directly on freeze-testing methods, whereby fresh needles, buds, and stems are subjected to very carefully controlled freezing conditions and then visually observed or measured for electrolyte leakage in solution to gauge level of cellular damage (Aitken and Adams 1997; Hurme et al. 1997; Wheeler et al. 2005; Prada et al. 2014). In recent years, drought has increasingly become of concern as changing climates have often led to warmer temperatures and decreased precipitation in some environments. Drought tolerance has typically been evaluated in growth chambers or nursery beds where irrigation is withheld, and tolerance is estimated using an array of physiological measurements, such as pre-dawn water potential, transpiration, photosynthesis, water-use efficiency, osmotic adjustment (turgor), xylem cavitation, or simply scoring for survival, wilting, growth, and phenological traits (Tan et al. 1992; Sonesson and Eriksson 2003; Atzmon et al. 2004).

The current thrust of research interest in drought and cold tolerance, and the genetic foundations thereof, is driven by the development of molecular technologies and large genetic datasets. Initially, molecular markers were used to map quantitative trait loci (QTLs) controlling adaptive traits, such as spring and fall cold hardiness and phenological traits (Jermstad et al. 2001a, b, 2003; Wheeler et al. 2005). Studies such as these generally defined the number and location of loci controlling traits, and helped explain the genetic correlation, or lack thereof, between traits (Howe et al. 2003). The availability of a virtually limitless number of single nucleotide polymorphism markers (SNPs) generated by next-generation sequencing technologies has opened the door to genetic association and genomic selection studies.

Association studies have been used to identify candidate genes likely controlling specific cold (Eckert et al. 2009b) and drought (González-Martínez et al. 2006a; Eckert et al. 2010b) traits in conifers. QTL and genetic association studies typically identify a handful of genes, each of which controls a very small proportion of the genetic variation in each trait (1–5% each; see Chap. 11). The rest of the variation remains unexplained. The true complexity of the genetic basis of most traits, including cold and drought tolerances, may best be illustrated by gene expression studies. Expanding genome sequence and EST (expressed sequence tag) databases have allowed researchers to develop cDNA microarrays with thousands of unique genetic elements representing genes, many of which have known functions in one or more plant species. These arrays can be queried using DNA isolated from trees that have been exposed to abiotic stress. Studies such as these have revealed gene expression patterns during fall cold acclimation in *Picea sitchensis* (Holliday et al. 2008) and following water stress in *Pinus taeda* (Lorenz et al. 2006, 2011). Literally hundreds of genes may be simultaneously up- or downregulated in such studies, highlighting the complex nature of trait regulation. In addition to the many studies of genetic variation in abiotic traits noted in Table 8.2, studies directed specifically to the topics of cold, drought, and salinity tolerance are itemized here in a manner like that for insects and disease resistance traits (Table 14.2).

Table 14.2 An enumeration of studies characterizing genetic diversity in tolerance to abiotic challenges in conifer taxa, listed by family, genus, and species. All citations listed are identified by host and pathogen name and are categorized by the approach taken to gather results and the status of the research program conducting the study. Codes for status and approach are noted in table footnotes and are described in the chapter text

Genus	Species	Abiotic Stress	Status[a]	Approach[b]	Literature Cited
Cupressaceae					
Cupressus	*nootkatensis*	Cold	4	1	Silim and Lavender (1994)
Taxodium	*distichum*	Salinity	4	1	Zhou et al. (2010b)
Thuja	*plicata*	Cold	4	1	Silim and Lavender (1994)
Pinaceae					
Abies	*many*	Cold	4	1	Jones and Cregg (2006)
Picea	*abies*	Drought	3	1	Sonesson and Eriksson (2003)
Picea	*abies*	Cold	3	1	Gomory et al. (2010)
Picea	*glauca*	Cold	3	1	Nienstaedt (1985)
Picea	*mariana*	Drought	3	1	Tan et al. (1992, 1995)
Picea	*rubens*	Pollution	4	2	Bashalkhanov et al. (2013)
Picea	*sitchensis*	Cold	3	1	Cannel and Sheppard (1982)
Picea	*sitchensis*	Cold	3	1,2	Holliday et al. (2008), Dauwe et al. (2012)
Pinus	*banksiana*	Drought	4	1	Mayne et al. (1994)
Pinus	*brutia*	Cold	3	1	Kandemir et al. (2010)
Pinus	*halepensis*	Drought	3	1	Atzmon et al. (2004)
Pinus	*many*	Cold	3	1	Hodge et al. (2012)
Pinus	*nigra*	Drought	3	1	Thiel et al. (2012)
Pinus	*nigra*	Cold	3	1	Kreyling et al. (2012)
Pinus	*pinaster*	Drought	2,3	1,2	Velasco-Conde et al. (2012)
Pinus	*pinaster*	Drought	3	2	Eveno et al. (2008)
Pinus	*pinaster*	Drought	3	1	Lamy et al. (2014)
Pinus	*pinaster*	Cold	3	1	Prada et al. (2014)
Pinus	*ponderosa*	Drought	3	1	Cregg (1994)
Pinus	*radiata*	Drought	2,3	1	De Diego et al. (2012)
Pinus	*strobus / wallichiana*	Cold	3	1	Lu et al. (2007)
Pinus	*sylvestris*	Cold	3	1	Hurme et al. (1997), Persson et al. (2010), Abrahamsson et al. (2012)
Pinus	*taeda*	Salinity	4	3	Tang et al. (2005)
Pinus	*taeda*	Drought	4	2	Lorenz et al. (2011)
Pinus	*taeda*	Drought	2,3	2	Gonzalez-Martinez et al. (2006a), Eckert et al. (2010b)
Pinus	*taeda*	Drought	4	1,2	Lorenz et al. (2006)
Pseudotsuga	*menziesii*	Drought	2,3	1	Anekonda et al. (2002)
Pseudotsuga	*menziesii*	Drought	3	1	Kavanagh et al. (1999)

(continued)

Table 14.2 (continued)

Genus	Species	Abiotic Stress	Status[a]	Approach[b]	Literature Cited
Pseudotsuga	*menziesii*	Cold	3	1	Campbell and Sugano (1979), Aitken and Adams (1997), Anekonda et al. (2000), Hawkins and Stoehr (2009), Darychuk et al. (2012)
Pseudotsuga	*menziesii*	Cold	3	2	Jermstad et al. (2001a, b)
Pseudotsuga	*menziesii*	Cold	3	1,2	Wheeler et al. (2005), Eckert et al. (2009b)

[a]1 = genetic variation in resistance detected in seedling or clonal screens; 2 = genetic variation in resistance detected in genetic or provenance trials; 3 = genetic variation in resistance detected in programs breeding for resistance, but no materials have yet to be deployed; and 4 = genetic variation in resistance detected in breeding programs with resistant materials deployed in the field operationally

[b]1 = traditional plant breeding methods like clonal, provenance or progeny testing; 2 = molecular biology approaches like QTL and association mapping, sequencing and gene annotation, and gene expression studies; 3 = genetic engineering

Mechanisms of Resistance and Tolerance

It should probably not be surprising that gene expression studies suggest hundreds of genes are implicated in biotic and abiotic challenges, either as causative or reactive effects (Chap. 6). Conifers have been coevolving with pests for millions of years, and have developed a broad array of defensive mechanisms, virtually all of which are under genetic control. Specific mechanisms vary widely by category of pest (i.e., needle, shoot, or root pathogens, bark beetles, insect herbivory, etc.), tree species, and conifer families and are difficult to classify or categorize unambiguously. Defenses have been variously described as preformed or inducible, mechanical or chemical, local or systemic, and specific or general (Lieutier 2007; Franceschi et al. 2005; Krokene 2015), but all bifurcated systems are simplifications of complex interactions. Krokene (2015) reflects on this complexity by using the example of resin ducts in sapwood. They may be preformed, or induced, and their mode of action can be both chemical (toxic or inhibitory) and mechanical (pitching out insects). In comprehensive reviews of conifer defense and resistance to bark beetles (Franceschi et al. 2005; Krokene 2015), the authors develop a classification centered around preformed or inducible defenses which we briefly describe here. This treatment will be followed by a review of defensive mechanisms mustered by conifers to deal with pathogenic pests.

Bark beetles are among the most destructive pests of conifer forests. Severe outbreaks can result in massive mortality over expansive areas (Fig. 14.3). Natural Resources Canada reports that over 50% of the standing volume of *Pinus contorta* in the province of British Columbia has been killed by an infestation that has raged since the mid-1990s, and bark beetle mortality, generally associated

Fig. 14.3 Extensive mortality of *Pinus contorta* ssp. *latifolia* (lodgepole pine) stands near Bonaparte Lake, British Columbia, Canada, caused by *Dendroctonus ponderosae* (mountain pine beetle) infestations coupled with climate change effects**.** (Photo from K. Buxton, BC Ministry of Forests, Lands and Natural Resource Operations)

with long-term drought, has been rampant throughout the western United States over the last 20 years. Such episodic outbreaks are exceptional, though seemingly increasingly common as climates change. Under more typical and historical conditions, insect populations remain rather modest, and affect only weakened or dying trees, constrained by an elaborate hierarchy of tree defense mechanisms (Krokene 2015). These mechanisms must deal not only with the beetles, but also with a host of symbiont organisms such as fungi, bacteria, and mites that associate with the beetles. Conifer defenses against bark beetles are designed primarily to protect the cambial tissues and secondarily the bark and sapwood that transport water and nutrients between roots and shoots. The first line of defense consists of preformed (constitutive) mechanical and chemical mechanisms. Mechanical (anatomical) defenses include structural elements in the wood and bark that deter "penetration, degradation and ingestion" (Franceschi et al. 2005; Krokene 2015). Chemical defenses include metabolites with toxic or inhibitory effects on the pest. Inducible defenses provide a second line of attack on invading pest and include some to many of the same characteristics as noted for constitutive defenses. Combined, preformed, and inducible defense mechanisms resist attacks in most healthy, thrifty trees.

Anatomical/mechanical defenses A conifer stem is comprised of three main tissue regions: phloem or bark, sapwood, and heartwood. The phloem and sapwood are separated by the vascular cambium, which produces new cellular tissue for each of these tissue types. Most defenses occur in the bark/phloem and include:

- The periderm – the outer surface of the bark. A complex of tissues that protect against penetration by insects and abiotic challenges like desiccation and fire. This is primarily comprised of cork tissue, which may be lignified or suberized, usually encrusted with calcium oxalate crystals and infused with phenolic compounds.
- The cortex – the inner layer of the bark with phenolic inclusions and, in some pine species, axial resin ducts.
- Secondary phloem – the primary site of both preformed and inducible stem defenses in conifers. Three concentrically arranged cell types provide key roles in defense (Franceschi et al. 2005; Whitehill et al. 2016):
  - Lignified sclerenchyma cells
  - Cells with calcium oxalate crystals
  - Polyphenolic parenchyma cells (PP cells)
- Radial rays with associated resin ducts – these hold resins under pressure.

The cambial tissue is virtually defenseless and serves only to produce new cells for the phloem (defenses and transportation of nutrients from shoot to roots) and xylem (wood, water transport). Protecting the cambial tissue is key to the tree's survival. Sapwood defenses are restricted largely to resin ducts and radial rays. The former can be induced to produce abundant traumatic resin ducts under attack or wounding.

Chemical defenses Chemical defenses are comprised primarily of two classes of compounds, terpenes and phenolics, which are the primary constituents of resins stored in resin ducts and PP cells. They are both toxic and inhibitory, able to flush out or trap insects when released under pressure. Terpenes constitute an immense array of known compounds (>30,000; Keeling and Bohlmann 2006b; see Chap. 7). This great diversity is created from three simple building blocks that are subsequently modified into many forms. Terpenes are relatively simply inherited, but the multifunctionality of the enzymes produced by these genes contributes to the biochemical diversity of the conifer resins. Phenolics are also diverse (ca. 4000 structurally diverse compounds). The major biosynthetic pathway for phenolics in conifers is the shikimic acid pathway which links the carbohydrate metabolism to the biosynthesis of aromatic amino acids (Ralph et al. 2006; Krokene 2015). While biosynthesis of phenolics occurs via a limited number of metabolic pathways, many of the biosynthetic enzymes in those pathways belong to multimember gene families, which contribute to the extreme structural diversity observed (Porth et al. 2011). Secondary phenolic compounds such as flavonoids, stilbenes, and condensed tannins derive from the aromatic amino acids and have been associated with plant defense mechanisms.

Phylogenetic differences in bark beetle defense strategy Fundamentally different defense strategies against bark beetles have been developed by the six extant

conifer families, splitting the pines from the non-pines (Franceschi et al. 2005; Krokene 2015). The non-pines, which include the Araucariaceae, Podocarpaceae, Taxaceae, Cupressaceae, and Sciadopityaceae, rely heavily on massive preformed bark defenses and extracellular calcium oxalate crystals, while the genera of the Pinaceae exhibit the combinations noted in previous paragraphs: preformed and inducible resin-producing structures, stone cell sclerenchyma in the bark, and scattered intracellular calcium oxalate crystals in the bark (Fig. 14.4).

Within the Pinaceae, considerable variation exists among species in defensive strategies, most notably the allocation of resources to constitutive or inducible mechanisms (Moreira et al. 2014). Furthermore, the patterns of variation appear to follow geographic and climatic clines and are related to species growth parameters. Slow-growing species invest more in constitutive defenses while fast-growing species invest more in inducible defenses. Moving poleward and to higher elevations, growth rate and inducible defenses decreased, while preformed defenses increased (Moreira et al. 2014). The patterns appear to be most sensitive to temperature.

Resistance mechanisms against needle and shoot pathogens Conifers deploy a wide array of mechanisms to resist or mitigate infections caused by disease organisms that attack needle, shoot, and, less frequently, main stems (Table 14.3). The great majority of these diseases are caused by fungi, like the rusts and blights. Disease-resistance mechanisms can be categorized much like those noted previously for bark-beetle infestations (e.g., preformed or induced, chemical or mechanical). The actual mechanisms employed are generally quite different, however, in part due to the types of tissues under attack and in part a function of the genetic basis of the mechanisms. Defense mechanisms may also be categorized based on whether they produce qualitative or quantitative resistance (Fraser et al. 2015). Single, major genes conferring resistance to rusts have been observed for *Pinus monticola*, *P. lambertiana*, and *P. taeda*, for instance, though broader-based resistance, dependent on multiple mechanisms, have also been noted for these species (Kinloch et al. 1970, 1999; Isik et al. 2012). In a recent review of defense mechanisms involved in host-pathogen interactions between pine species and needle/shoot pathogens, mechanisms were categorized using yet another division: defenses acting pre-penetration versus those acting post-penetration (Fraser et al. 2015; Table 14.3). Pre-penetration mechanisms were primarily anatomical or morphological in nature, while post-penetration mechanisms included chemical (phenolics, terpenoid resins), metabolic (antimicrobial peptides, pathogenesis-related proteins), and mechanical responses (hypersensitive response – HR is characterized by the rapid death of cells in the local region surrounding an infection, effectively sealing off the infection).

We have presented material on insect and disease resistance mechanisms to give a sense of the genetic basis of resistance. Even at this cursory level of discussion, it is easy to see how genetic variation in resistance or tolerance to challenges from

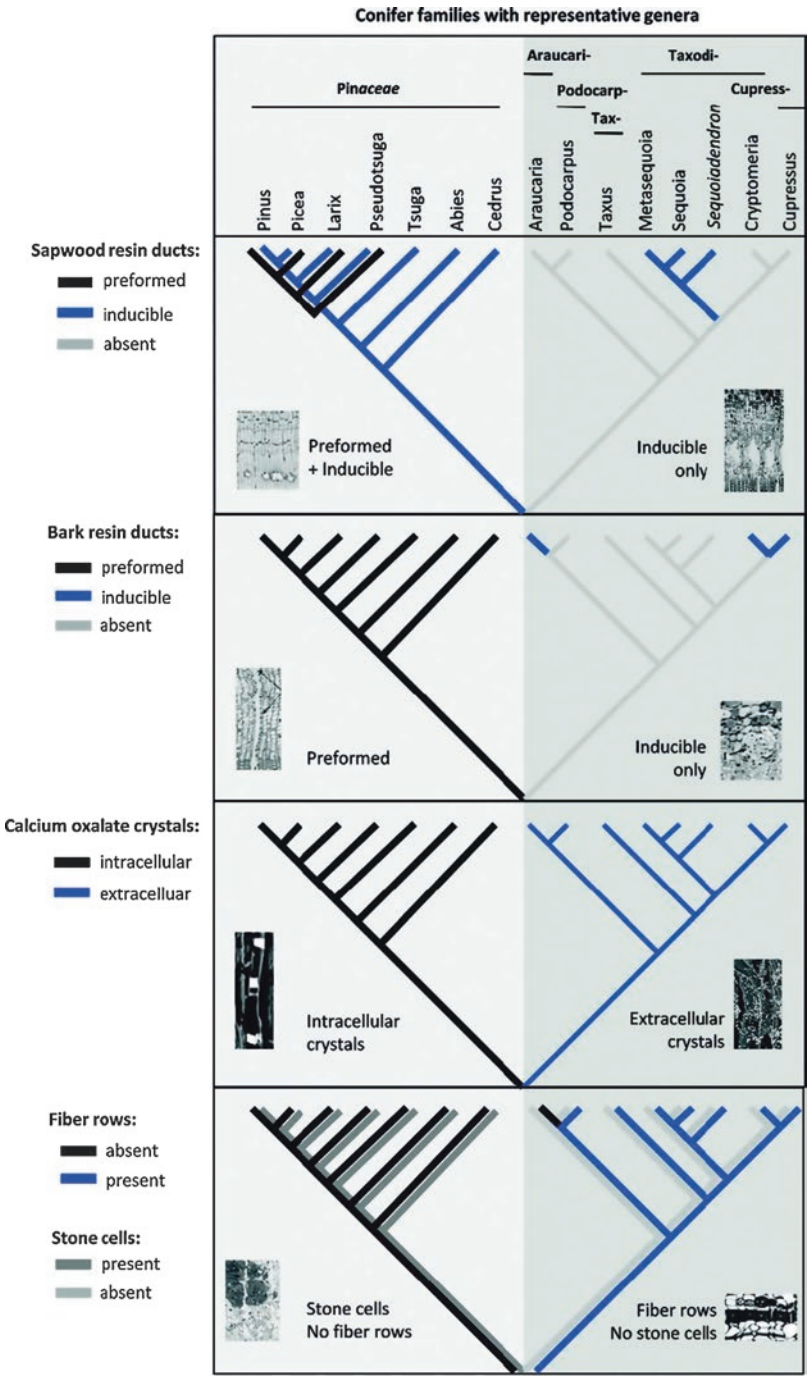


Fig. 14.4 Conifers form two phylogenetic groups with fundamentally different defense strategies. Species in the pine family (left) generally have elaborate resin-based defenses in the bark and sapwood, large intracellular calcium oxalate crystals, and large stone cells, but no fiber grows in the bark. Species in all other conifer families (right) have no preformed (but sometimes inducible) resin ducts in the bark and sapwood, small abundant extracellular calcium oxalate crystals, and closely spaced fiber rows in the bark. (From Franceschi et al. 2005, as cited in Krokene 2015)

Table 14.3 Defense mechanisms involved in interactions between Pinaceae and needle and shoot pathogens. Mechanisms are categorized by whether they act pre- or post-penetration and are listed by pest or pathogen and host species. (Extracted from Table 14.1 from Fraser et al. 2015, page 6)

Defence mechanism	Pathogen/pest	Disease	Host	References
Defences acting pre penetration		*Dothistroma* needle blight	*Pinus radiata*	Franich et al. (1977)
Needle surface topography	*Dothistroma septosporum*	White pine blister rust	*P. strobus*	Smith et al. (2006a)
Epicuticular wax	*Cronartium ribicola*	*Dothistroma* needle blight	*P. radiata*	Franich et al. (1983)
	D. septosporum	*Dothistroma* needle blight	*P. radiata*	Franich et al. (1983)
Epicuticular resinacids (diterpenes)	*D. septosporum*	White pine blister rust	*P. strobus*	Patton (1972), Smith et al. (2006a)
Epistomatalwax occlusion	*C. ribicola*	*Dothistroma* needle blight	*P. radiata*	Franich et al. (1977, 1983)
	D. septosporum		*P. muricata*	Muir and Cobb (2005)
Stomatal morphology	*C. ribicola*	White pine blister rust	*P. strobus*	Woo et al. (2001)
	D. septosporum	*Dothistroma* needle blight	*P. radiata*	Franich et al. (1977)
Defences acting post-penetration Phenolic compounds	*C. ridicola*	White pine blister rust	*P. strobus*	Boyer (1964), Boyer and Isaac (1964), Jurgens et al. (2003), Jacobs et al. (2009)
	Cronartium fusiforme	Fusiform rust	*P. elliotti*	Lesney (1989)
	Diplodia sapinea	Diplodia blight of pines	*P. nigra*	Wallis et al. (2008)
			P. radiata	Regliński et al. (1998)
			P. resinosa	Blodgett et al. (2007)
	D. septosporum	*Dothistroma* needle blight	*P. radiata*	Franich et al. (1986), Hotter (1997)
	Elytroderma deformans (Weir) Darker	Elytroderma needle cast	*P. contorta*	Wallis et al. (2010)
	Lophodermella concolor; L. montivaga Petrák	Lophodermella needle cast	*P. contorta*	Wallis et al. (2010)
	Sirococcus conigenus	Sirococcus blight	*Picea abies*	Bahnweg et al. (2000)

(continued)

Table 14.3 (continued)

Defence mechanism	Pathogen/pest	Disease	Host	References
Resin	*D. septosporum*	*Dothistroma* needle blight	*P. radiata; P. muricata*	Cobb and Libby (1968)
	Lecanosticta acicola	Brown spot needle blight	*P. palustris*	Verral (1934) (cited in Cobb and Libby 1968)
Terpenoids	*C. fusiforme*	Fusiform rust	*P. elliottii; P. taeda*	Michelozzi et al. (1995)
	D. sapinea	Diplodia blight of pine	*P. nigra*	Wallis et al. (2008)
			P. resinosa	Blodgett and Stanosz (1997)
	E. deformans	Elytroderma needle cast	*P. contorta*	Wallis et al. (2010)
	L. concolor; L. montivaga	Lophodermella needle cast	*P. contorta*	Wallis et al. (2010)
Hypersensitive response	*C. ribicola*	White pine blister rust	*P. flexilis*	Kinloch and Dupper (2002)
			P. lambertiana	Kinloch et al. (1970), Kinloch and Littlefield (1977)
			P. monticola	Kinloch et al. (1999)
			P. strobiformis	Kinloch and Dupper (2002)
	C. fusiforme	Fusiform rust	*P. taeda*	Wilcox et al. (1996)
Hypersensitive-like response	*C. fusiforme*	Fusiform rust	*P. taeda*	Gray and Amerson (1983)
	C. ribicola	White pine blister rust	*P. armandii*	Hoff and McDonald (1975)
			P. monticola	Hoff and McDonald (1975)
			P. strobus	Jurgens et al. (2003), Jacobs et al. (2009)
	D. septosporum	*Dothistroma* needle blight	*P. radiata*	Hotter (1997)
Antimicrobial peptides (AMP)	*C. ribicola*	White pine blister rust	*P. monticola*	Ekramoddoullah et al. (2006), Liu et al. (2013b)
Pathogenesis-related (PR) proteins	*C. ribicola*	White pine blister rust	*P. monticola*	Liu et al. (2003b, 2005, 2010a, b)
			P. strobus	Smith et al. (2006b)

(continued)

Table 14.3 (continued)

Defence mechanism	Pathogen/pest	Disease	Host	References
Needle shedding	*C. ribicola*	White pine blister rust	*P. monticola*	McDonald and Hoff (1971), Liu and Ekramoddoullah (2011)
			P. wallichiana	Heimburger, 1962 (in McDonald and Hoff 1971)
Stationary interface	*L. sulcigena*	Lophodermella needle cast	*P. nigra* ssp. *laricio*	Williamson et al. (1976)
	Ploioderma hedgcockii	Ploioderma needle cast	*P. palustris*	Jewell (1990)

biotic agents can be expressed in complex ways. In fact, the genetic basis of inducible resistance mechanisms is much more complex than can be covered in this review and is an active area of research in plants, including conifers (Hudgins et al. 2004; Porth et al. 2011). Inducible mechanisms of resistance generate responses to an attack (insect feeding, pathogen-host interaction) through a cascading series of regulatory actions or pathways (STP – signal transduction pathways). The STP may simultaneously induce the plant to develop a hypersensitive response and systemically send signals to other parts of the plant to invest in defensive mechanisms as well, a process known as SAR (systemic acquired resistance). The result of STP response is the near-simultaneous shift in gene expression for hundreds of genes (Ersoz et al. 2010; Robert et al. 2010).

Case Studies

We have chosen several examples to illustrate genetic variation in resistance to insect and disease organisms expressed in conifers.

Resistance to *Pissodes strobi* (White Pine Weevil) Found in *Picea sitchensis* (Sitka spruce)

The white pine weevil is a destructive pest of sapling and pole-size pine and spruce species in North America. Repeated attacks kill the main leaders of young trees, often resulting in multi-stemmed and bushy-crowned trees of little commercial value. The prevalence of this insect along the northern Pacific Coast has led to sharp declines in reforestation of Sitka spruce in much of its northern range. Concerted efforts by scientists in federal and provincial research organizations of Canada, beginning in the early 1990s, identified geographical patterns of resistance to the insect in Sitka spruce provenance trials (Fig. 14.5). The regions identified as

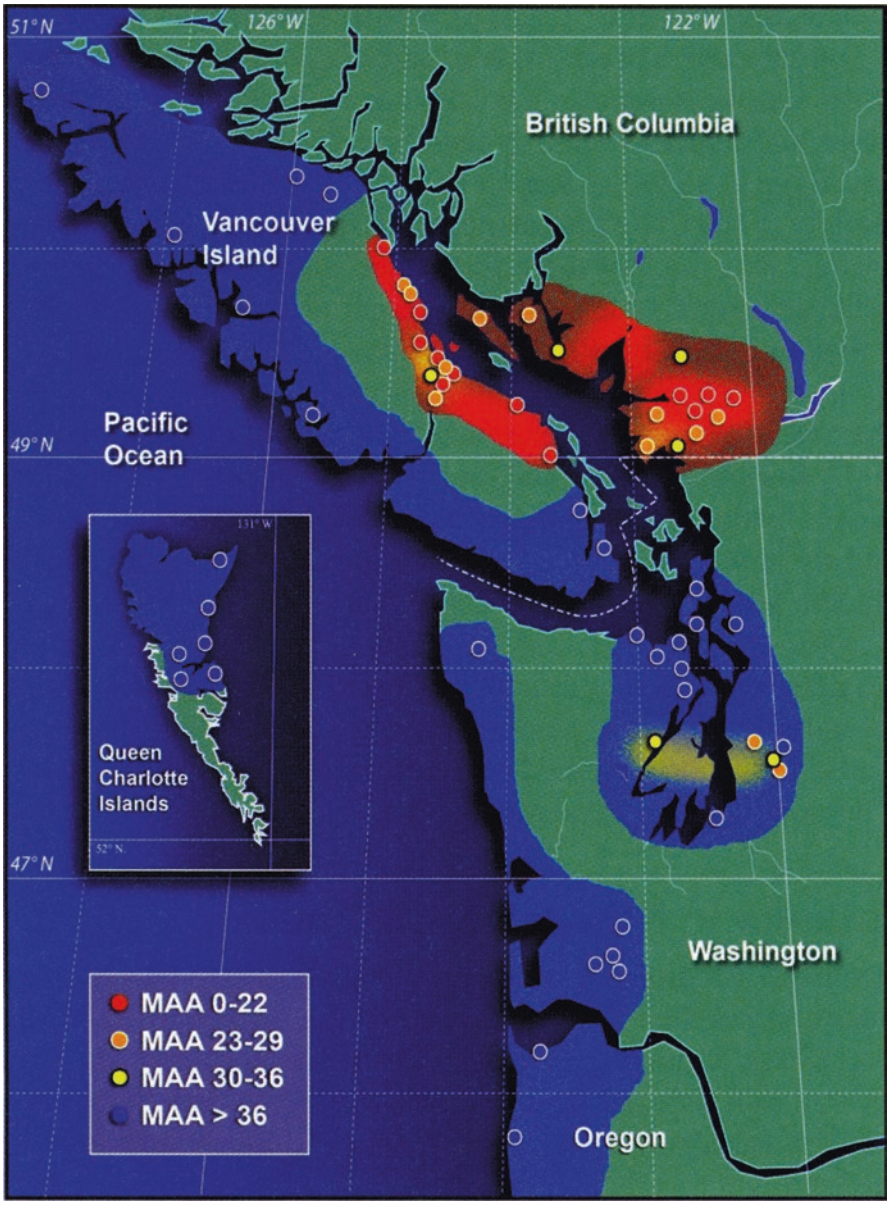


Fig. 14.5 Resistant provenances identified by mean annual attack (MAA) level and zones of high and low natural resistance; MAA is expressed as a percentage. The resistant region is shown in red, the susceptible region in blue shades (intermediate in yellow). (From King et al. 2012)

possessing high levels of resistance to weevil attack have allowed tree breeding programs to capture resistance in progeny of selected parents, and foresters are once again growing trees for reforestation. In the process, they have identified several mechanisms of resistance that appear to be operating, often in combination, most of which appear to be hindrances to insect attack, including: induced and constitutive resin cell production, production of sclereid or stone cells, and terpene defenses (King et al. 2012). The authors point out that these mechanisms are complex and likely rely on many genes (quantitative genetic variation) to deliver resistance, an observation consistent with the historical perception that, in conifers, resistance to insects is almost exclusively quantitative in nature.

However, the authors of this and previous studies (reviewed in King et al. 2012) have identified one or more trees that appear to be totally resistant to successful attack by the white pine weevil, strongly suggesting a major gene or R-gene resistance may exist. Though R-gene resistance systems for insects have been identified in some crop species, the spruce example would be rather novel for conifers. At the time of this study, the gene or specific mechanism responsible for the total resistance to the weevil has yet to be identified. Overall, the results of this program of study and reforestation are quite remarkable in terms of successful identification of resistance and the speed with which resistant trees could be redeployed on the landscape.

Resistance to an Introduced Pathogen (*Phytophthora lateralis*) in *Chamaecyparis lawsoniana* (Port-Orford-cedar)

Port-Orford-cedar (POC) is native to a limited range in the Klamath Mountains of southwestern Oregon and northwestern California where it grows to considerable size and age (Zobel et al. 1985). The species is genetically quite diverse with respect to crown traits and consequently has been widely planted as an ornamental with over a dozen recognized varieties officially named. In the 1950s, a nonnative root-infecting pathogen, *Phytophthora lateralis*, appeared and began killing POC trees in portions of the species' range. The pathogen was likely introduced on nursery horticultural stock as early as the 1920s. Areas particularly susceptible were those in riparian habitats, where mortality was very high. Mortality in upland, drier areas proceeded more slowly. In 1996, scientists in the United States Forest Service began a program to develop disease resistant POC (Sniezko et al. 2012a, b). This was done via an array of inoculation approaches of seedlings, grown from trees collected in the forest, and later, from select trees growing in clonal seed orchards, and planted in test beds (Hansen et al. 2012). Still other trials were conducted with wind- and control-pollinated families planted in field test sites where heavy mortality was observed in native stands.

In one such study, 16 families were selected for outplanting based on their previous performance in nursery test beds, with families known to range from 0% to 100% mortality. After 9 years in the field plantings, large and statistically significant

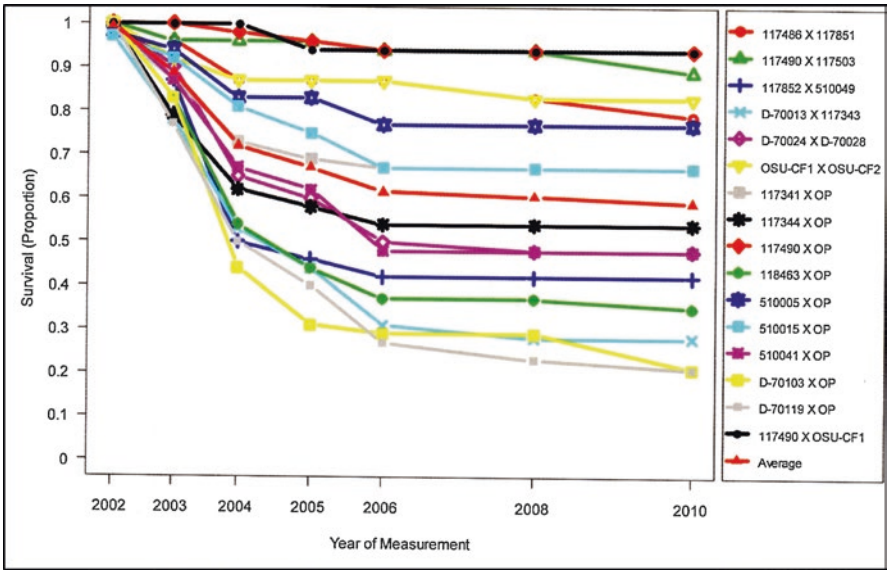


Fig. 14.6 Nine-year survival of 16 *Chamaecyparis lawsoniana* families inoculated with *Phytophthora lateralis* in a resistance field planting trial. (From Sniezko et al. 2012a)

differences in survival were noted among families at the most vulnerable test site, ranging from 20.8% to 93.8% (Fig. 14.6). Crosses from parents that had proven resistant to the pathogen in nursery root-dip trials consistently performed best in field tests, both confirming the value of short-term tests and illustrating the durable nature of the resistance to this pathogen.

The distribution of variation in a family's performance noted in Fig. 14.6 likely reflects the genetic basis of resistance observed in this species. The near-complete resistance noted in some families (e.g., 117490, CF1) suggests major gene resistance (MGR) exists, while intermediate levels of resistance noted in other families implies quantitative resistance (Sniezko et al. 2012b; R. Sniezko, personal communication, April 18, 2017). Research on the mechanisms of resistance in POC have identified a number of features consistent with hypersensitive response (HR) in the host including cell-wall thickening, cell collapse, and deposition of electron-dense materials around fungal hyphae (Oh et al. 2006; Oh and Hansen 2007). HR response is typical of MGR.

The rather rapid identification and use of resistant genotypes in species recovery from an introduced pathogen, as shown in the POC program, is exceptional and implies an evolutionary store of genetic variability in this species dating from its split with related species, most of which are Asian in distribution. It may also reflect a relatively narrow genetic base in the nonnative pathogen. As noted previously, conifers appear to deviate from angiosperm forest trees in this regard, the latter often lacking significant sources of resistance to nonnative pathogens.

Resistance to Stem Rusts in North American White Pines and Southern Yellow Pines

Stem rust pathogens are of major economic and ecologic importance to pine forests of North America. In the southeastern United States, *Cronartium quercuum* sp. *fusiformae,* the native fusiform rust, infects commercially valuable species, most notably *Pinus taeda* (loblolly pine) and *P. elliottii*, (slash pine), with estimated losses in productivity exceeding $140 million annually (Cubbage et al. 2000). In the United States and Canada, particularly in the west, the introduced *Cronartium ribicola* (white pine blister rust) has caused widespread mortality in commercially valuable species, such as *P. monticola* (western white pine), *P. lambertiana* (sugar pine), and *P. strobus* (eastern white pine), and has spread to most of the remaining native white pine species in the United States and Canada (five of six), many of which occupy ecologically critical high-elevation ranges. Among the latter are *P. albicaulis* (whitebark pine; see Fig. 14.2), *P. flexilis* (limber pine), *P. balfouriana* (foxtail pine), *P. aristata* (bristlecone pine), *P. strobiformis* (Southwestern white pine), and *P. longaeva* (Great Basin bristlecone pine). Efforts to identify sources of resistance to these pathogens through breeding and testing represent some of the longest-running such programs in the world, exceeding 50 years in both cases (reviewed in Sniezko et al. 2014). Genetic resistance to the rusts has been found in all the species evaluated, but damage has been profound and, for some species, will require decades or centuries from which to recover. Here we briefly discuss the differences between the two *Cronartium* pathosystems and the types of variation in resistance observed in their host species.

Resistance to fusiform rust has been known for decades, but the genetic basis of the host-pathogen interaction eluded scientists for over 35 years. In 1996, a study using RAPD (random amplified polymorphic DNA) markers, single-aeciospore isolates of the pathogen, and pedigreed loblolly pine families known to exhibit resistance identified major gene resistance (*Fr* locus) to the disease (Wilcox et al. 1996). This study closely mimicked an early study that identified single gene resistance to *C. ribicola* in the white pines (Devey et al. 1995). In the intervening years, investigators have identified an additional eight *Fr* genes in other families of loblolly pine (Isik et al. 2012; Amerson et al. 2015). This host-pathogen interaction has proven to be a classic instance of gene-for-gene resistance to a disease. That is, for each *Fr* gene in the host plant, the fungal species has a gene that can overcome that resistance (an allele for virulence). The expression of resistance by the host organism is thus dependent on the tree (1) possessing a *Fr* resistance allele at one or more of the nine known *Fr* loci and (2) the frequency of the avirulence/virulence alleles at those loci in the fungal population common to the area in which the tree is planted.

A comprehensive experiment illustrated the extent to which resistance, provided by the nine *Fr* loci, held up when confronted with ten different mixed pathogen inocula (Isik et al. 2012; see Table 14.4). In this trial, seven loblolly pine families known to possess resistance alleles (*R*) at one or more *Fr* loci were inoculated under controlled environments with the ten inocula, each presumably containing different fungal populations. Mean resistance figures, based on the formation of galls on inoculated seedlings, varied significantly by *Fr* gene, *Fr* genotype, and inocula mix

Table 14.4 Disease incidence means obtained from the *R* and *r* genotypes of seven *Pinus taeda* families inoculated with ten different *Cqf* inocula. (From Isik et al. 2012)

Fr	Inocula										Fr Genotype
Genotype	301	302	303	304	305	306	307	308	309	310	Mean
R1	0	0	0	0	0.19	0.1	0.27	0.09	0.03	0.46	*0.11*
r1	0.6	0.59	0.61	0.66	0.63	0.59	0.79	0.67	0.5	0.64	*0.63*
R2	0.58	0.68	0.69	0.48	0.82	0.57	0.69	0.59	0.43	0.57	*0.61*
r2	0.59	0.48	0.58	0.46	0.82	0.33	0.77	0.71	0.5	0.76	*0.6*
R3	0.03	0	0.03	0.04	0.11	0.04	0.03	0.14	0.07	0.03	*0.05*
r3	0.6	0.61	0.56	0.5	0.55	0.57	0.72	0.71	0.6	0.73	*0.62*
R4	0.09	0.28	0.38	0.17	0.03	0	0.09	0.03	0.03	0.5	*0.16*
r4	0.58	0.46	0.77	0.67	0.87	0.78	0.7	0.69	0.77	0.93	*0.72*
R5R9	0.1	0.15	0.25	0.57	0.5	0.14	0.14	0.59	0.23	0.5	*0.32*
R5r9	0.46	0.81	0.76	0.64	0.71	0.55	0.71	0.7	0.33	1	*0.67*
r5R9	0	0.21	0.5	0.31	0.53	0.1	0.46	0.6	0.29	0.44	*0.34*
r5r9	0.77	0.5	0.5	1	0.83	0.64	0.93	0.69	0.64	1	*0.75*
R6r7	0.27	0.56	0.08	0.39	0.32	0	0.17	0.44	0.26	0.44	*0.29*
R7r6	0.33	0.3	0.59	0.55	0.03	0.06	0.16	0.14	0.09	0.46	*0.27*
R8	0	0.1	0.21	0.17	0.04	0.24	0.59	0.64	0.38	0.7	*0.31*
r8	0.6	0.59	0.47	0.71	0.77	0.71	0.71	0.65	0.58	0.8	*0.66*
Mean_R	**0.19**	**0.31**	**0.35**	**0.33**	**0.33**	**0.18**	**0.33**	**0.4**	**0.21**	**0.51**	***0.31***
Mean_r	**0.62**	**0.54**	**0.58**	**0.67**	**0.75**	**0.6**	**0.77**	**0.69**	**0.6**	**0.81**	***0.66***

Incidence for each genotype and inoculum combination (example R1 × 301) is shown in the body of the table in plain type, with the mean for each genotype across all inocula (Genotype Mean) shown on the far right in italics. Colors also show the disease incidence level for inocula by genotype cells. Darker colors represent greater incidence. At the bottom of the table, Mean R values (in bold type) represent the mean for each inoculum across all genotypes having at least one known R allele, and Mean r values represent the mean for each inoculum across all genotypes having no known R alleles. At the bottom right of the table, the ***0.31*** value is the grand mean incidence for all genotypes having at least one known R allele, while the ***0.66*** value is the grand mean incidence for all genotypes having no known R alleles (r only) by all inocula combinations

(Table 14.4). The results of the study demonstrated the complex nature of the host-pathogen interaction in these coevolved organisms. For instance, the virulent allele of the fungus that overcomes the *R* allele in *Fr 1* (*R1*) occurs in very low frequencies in most of the inocula, and overall, only 11% of seedling with the *R* genotype were ultimately infected. Similarly, the virulent alleles against *R3* and *R4* were also conspicuously absent but quite common for *R2*.

Efforts to identify the *R* genes have so far been elusive, but they have been located in the loblolly pine genome (Fig. 14.7), genetically mapped using a large array of RAPD markers (Amerson et al. 2015). At least one of the *R* genes may have been identified in the *P. taeda* reference genome sequencing project (Neale et al. 2014). The *R* genes likely represent a multigene family, many of which co-locate on a single chromosome (*LG2*).

In loblolly pine, deploying resistant plant materials in commercial plantations is based entirely upon using these major *R* genes (qualitative genetic variation).

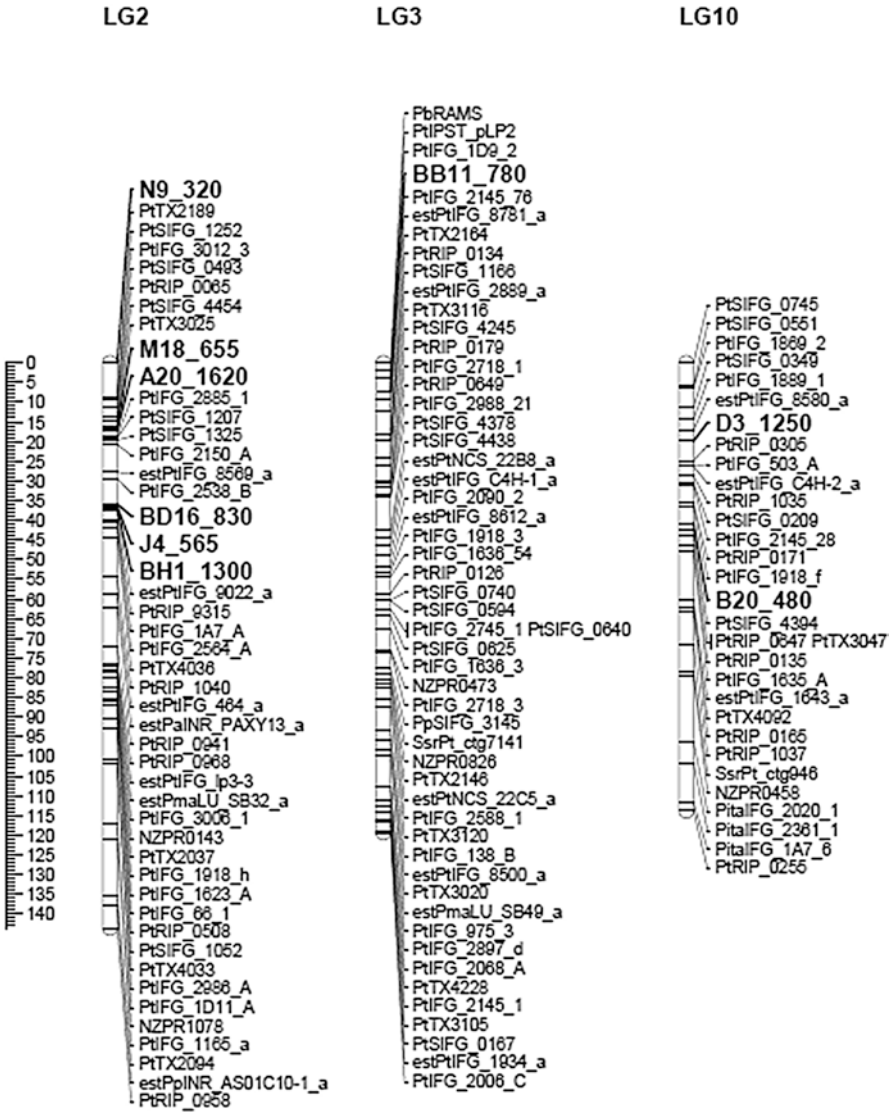


Fig. 14.7 Loblolly pine reference genetic map linkage groups 2, 3, and 10 showing the positions of the *Fr* gene-linked RAPD markers. RAPD markers are in larger font and bold type. (From Amerson et al. 2015)

In contrast, approaches to developing greater resistance to the white pine blister rust (*C. ribicola*) have relied heavily on quantitative or partial resistance mechanisms though major gene resistance, evidenced by hypersensitive-like responses (Fig. 14.8), which have been found in four of the nine native white pine species in North America (*P. lambertiana*, *Cr1*; *P. monticola*, *Cr2*; *P. strobiformis*, *Cr3*; and *P. flexilis*, *Cr4*; Devey et al. 1995; Sniezko et al. 2014, 2016; Liu et al. 2016b, 2017;

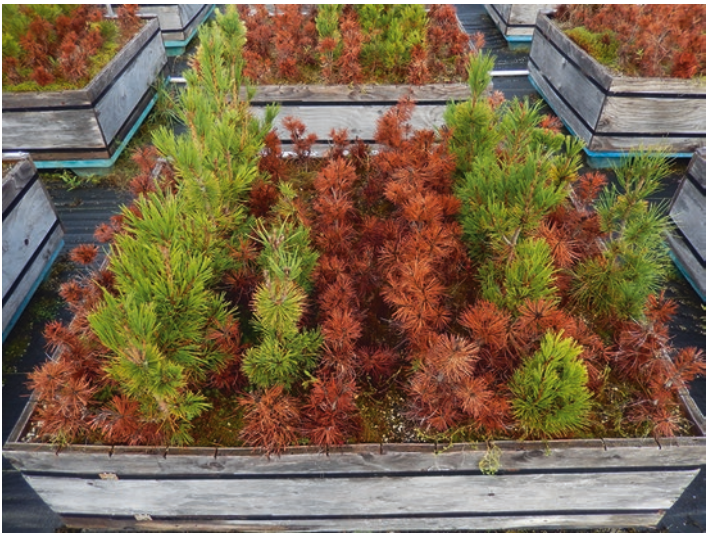


Fig. 14.8 Genetic variation in resistance to white pine blister rust among *Pinus albicaulis* (whitebark pine) families in a seedling inoculation trial at the USDA Forest Service's Dorena Genetic Resource Center, Cottage Grove, Oregon. Shown are 12 families (half-sib progeny of parent trees from natural stands) in 10-tree row plots, where survival ranges from 0% to >80%. The underlying mechanisms and inheritance of resistance in *P. albicaulis* are still to be resolved, but unlike four other western North American white pine species, no hypersensitive-like response in the needles has been documented. (Sniezko and Hansen 2001)

see also Table 14.3). In at least two species, virulent forms of the rust have been found. However, use of partial resistance mechanisms is necessitated by the very low frequencies of major *R* genes in natural populations (0–0.08 in *P. lambertiana*; 0–0.01 in *P. monticola*, and 0–0.129 in *P. flexilis*) and the presence of virulence in the rust populations. The exception to this is found in *P. lambertiana* where reliance on major *R* genes remains critical to deployment of the species.

Several types of partial resistance responses have been documented in white pines; however, relatively little is known about the underlying mechanisms and their genetic control (Quesada et al. 2010, 2014; Vázquez-Lobo et al. 2017). It is assumed that the mode of inheritance for most of these responses is more complex than for major gene resistance. Families with partial resistance exhibit responses described as (1) reduced number of stem infections, (2) latent stem infections following inoculation, (3) bark reactions, (4) higher survival of seedlings with cankers, and (5) delayed mortality. Though the descriptions of these responses do not embolden one to envision thriving young trees beating down disease, expression of these traits can lead to trees reaching maturity and reproducing. Breeders continue to cross trees with different responses (and presumably different mechanisms) in hopes of pyramiding traits to enhance resistance. Partial resistance is often viewed as having greater durability over time though in conifers the evidence in favor of this hypothesis seems incomplete.

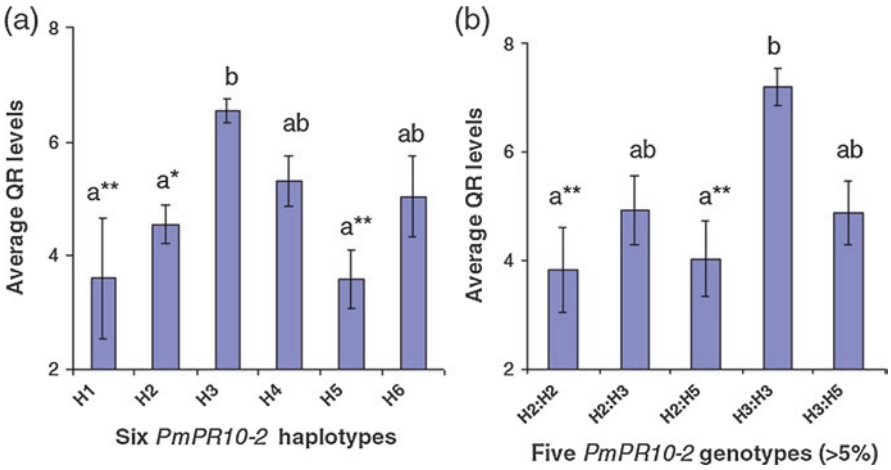


Fig. 14.9 Association of *PmPR10–2* alleles and genotype with white bark pine stem QDR levels. (**a**) Mean QDR levels were shown for tree groups with six *PMR10–2* alleles (H1 to H6) or (**b**) five common *PMR10–2* genotypes (H2:H2, H2:H3, H2:H5, H3:H3, and H3:H5). Standard error (SE) was calculated based on the entire population of each tree group. Statistical difference is significant (one-way ANOVA test and Tukey's highly significant difference test, $*P < 0.05$, $**P < 0.01$) between tree groups labeled with different letters. (From Liu et al. 2013b, c)

While it remains true today that the underlying mechanisms of partial resistance to the white pine blister rust organism are largely unknown, molecular genetic studies are beginning to reveal some of the genetic basis for these responses. Liu et al. (2013b, c) identified a pathogenesis-related protein (PR-10) in several families of western white pine that expressed different levels of quantitative disease resistance (QDR). The QDR score was based on the combined phenotypic scores of three different partial resistance responses. The PR-10 protein represents one of 17 families of proteins known to have pathogenesis-related effects. In this study, the authors identified a single locus with six novel alleles and five common genotypes (of 17 possible combinations) in the population studied. Significantly different levels of QDR were found among trees, and the variation was associated with the PR-10 alleles and genotypes (Fig. 14.9). Trees possessing the H3 allele, either in the heterozygous or homozygous condition, exhibited significantly greater QDR scores than trees with other alleles and genotypes. The PR-10 protein is thought to interact with the fungal hyphae as it moves from the site of infection in needles into the stems (Liu et al. 2013b, c). Overall, the rust-caused mortality, at 5 years, in families with positive QDR ratings ranged from 12% to 94% in this study, compared to 97–100% in families with QDR scores at or near zero. Information gained from this study will enable scientists to screen trees for resistance at this locus using molecular marker technology.

Summary

Most common and widespread conifer species display significant genetic variation in resistance to coevolved insect and disease pests and tolerance to abiotic stresses such as cold and drought. However, conifers, like their angiosperm tree counterparts, appear increasingly vulnerable to introduced pests and rapidly changing climatic conditions. Loss of forest productivity and declining ecosystem health have substantially accelerated in recent decades due to these and other factors. Efforts to characterize sources of genetic variation in pest resistance and abiotic tolerances are enumerated here. Significant progress in identifying the genetic basis of abiotic tolerance has been made (see also Chaps. 6, 7, 8, 10, 12) and breeders routinely deploy adapted stock for reforestation and restoration. Progress in identifying the genetic basis of disease and insect resistance is slower but will likely improve rapidly as molecular genetic resources expand.

Mechanisms of pest resistance are many and varied, and most are under genetic control. Specific mechanisms vary widely by category of pest (i.e., needle, shoot, or root pathogens, bark beetles, insect herbivory, etc.), tree species, and conifer family and are difficult to classify or categorize unambiguously. Defenses have been variously described as preformed or inducible, mechanical or chemical, local or systemic, and specific or general, but all bifurcated systems are simplifications of complex interactions. Major gene resistance to insect pests is relatively rare but occurs more frequently for disease organisms. A commonly held view is that resistance based on multiple mechanisms (quantitative genetic-based resistance) is more durable than major gene resistance.

Examples of identified sources of genetic resistance to insect and disease pests are reviewed, including studies on (a) *Pissodes strobi*, the native white pine weevil, in western North America, (b) the introduced root pathogen *Phytophtora lateralis* on *Chamaecyparis lawsoniana* (Port-Orford-cedar), and both native and introduced species of stem rusts (species of *Cronartium*) on North American pines.

Part III

Evolution

Hybridization and Introgression

15

Introduction

Spontaneous hybridization between taxa provides a source of genetic variation upon which selection may act. Though individual genetic variants are not new, the combinations of variants from diverse genomes found in hybrids are new (novel epistatic gene interactions), and hybrid populations may, and often do, contain more genetic diversity than their parental species (Stebbins 1959; Lewontin and Birch 1966; Grant 1971). While investigators have long debated the evolutionary significance of hybridization and the subsequent incorporation of genetic variation from one species into another, several studies reviewed in this chapter would suggest these phenomena play an important role in conifer evolution. This chapter seeks to summarize the scientific literature on hybridization and introgression in conifers with particular attention given to studies from the last quarter century. As in previous decades, much of this literature is descriptive. Increasingly, however, investigations have expanded to investigate such important principles as which speciation model is most appropriate, what are the mechanisms by which species boundaries are maintained, and how methods used to detect and measure introgression affect our interpretation of the results. Such important queries are reviewed by looking at specific case studies.

The study of hybridization in conifers has a rather long history, much of which was dominated by investigations of artificial crosses and the occurrence of spontaneous hybrids in gardens and arboreta. Species, hybrid, and provenance trials dominated the study of forest genetics for most of the first half of the last century (Wheeler et al. 2015). In the United States, and elsewhere around the world, concerted efforts were made to produce and test hybrids in hopes of finding trees exhibiting hybrid vigor and for clarifying phylogenetic relationships among taxa through crossability trials. In the United States, this work was concentrated in the US Forest Service (Duffield and Righter 1953; Wright 1955, 1959; Critchfield and Krugman 1967; Critchfield 1977, 1984b, 1986, 1988; Garrett 1979; Millar and Critchfield 1988). Results of some of these crossability studies are elaborated on at the end of this chapter, and a worldwide listing of known artificial hybrids, surely noncomprehensive, is provided.

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_15

Definitions and Background

Definitions

Several terms are reliably and repeatedly invoked in any discussion of hybridization and introgression, including the very terms themselves. Defining these terms is necessary though not entirely straightforward. *Hybridization* has been defined as the "crossing between individuals belonging to separate populations which have different adaptive norms" (Stebbins 1959). This definition can be interpreted as broadly as to include successful mating between virtually any two outcrossing individuals that are not genetically identical, but here we subscribe to the generally accepted definition of successful mating between individuals in different species. What is a species? Many definitions of what constitutes a species have been articulated but the classical definition, called the *biological species concept* (BSC), is a collection of individuals that successfully interbreed among themselves but not with members of other species (Niklas 1997). Obviously, the rigidity of this species concept would preclude hybridization which, as we shall see, occurs rather frequently among conifer taxa. Taxa that have diverged enough to be awarded species status but for which breeding barriers remain incomplete offer the opportunities to study the mechanisms of speciation and maintenance of species boundaries, topics to be discussed further in this chapter.

Hybrids between species range from being infertile to fully fertile. In the latter case, hybrid individuals may interbreed freely among themselves or with members of either parental species. Populations consisting of hybrids, their derivatives, and the parental species are called *hybrid swarms*. *Introgression* or *introgressive hybridization* has been defined as the repeated backcrossing of a natural hybrid to one or both parental populations, resulting in the transfer of genes from one species to another (Anderson 1949). Hybrid swarms typically occur in locations where species ranges overlap, or are *sympatric*. The ranges of *allopatric* species are geographically or ecologically separated and gene flow between them is nil or nearly so though long-distance transport of pollen or even seed is occasionally reported. Allopatry is the primary barrier to hybridization among taxa. Populations of the same species that are geographically or ecologically separated, or both, can begin to diverge, a gradual process called *allopatric speciation*. Other barriers to successful hybridization are genetic, and include such things as non-overlapping floral phenology of two taxa, inability to effect fertilization due to pollen tube growth inhibition, inviable embryos, and infertile hybrids. Alternative hypotheses concerning the mechanisms of genetic barriers, whether genic or chromosome based, continue to attract research attention.

The appearance of genetic differences in populations that hinders or prevents cross-fertilization and leads to divergence is labeled *sympatric speciation* (White et al. 2007). To our knowledge, there are few, if any, verified examples of sympatric speciation in conifers. *Hybrid speciation*, the origin of a new species directly from a natural hybrid, has been much discussed but infrequently described (Grant 1971, Chap. 13). Successful hybrid speciation requires stabilization of the hybrid's

breeding behavior. Grant (1971) lists seven methods by which this may occur, but only one, segregation of a new type isolated by external barriers (geographic isolation) from its parents, is relevant to conifers. Such conditions might occur following large-scale disturbances such as glaciation events; we will discuss a few likely conifer examples in later pages. Hybrid speciation results in *reticulate evolution*, a confounding of the otherwise bifurcating lineage concept of phylogenetics. Reticulate evolution may be a significant force in the genus *Pinus* (Willyard et al. 2009; Xing et al. 2014) and very possibly other conifer genera (Bouilli et al. 2011).

In nature, conifer hybrids arise almost exclusively when geographical or ecological barriers break down creating *hybrid environments* in which hybrids can thrive and compete with their parental species (Anderson 1948, 1949; Grant 1971). The concept was extended to suggest the hybrid environment must possess a varied array of ecological niches for introgression among subsequent generations of hybrids and backcrosses to be successful (Anderson 1949). While early studies of natural hybridization in plants focused on human-caused environmental disturbances such as agriculture, logging, and road building, by far the greatest disturbances experienced by forest trees have been caused by climate change, often coupled with the terraforming power of montane and continental glaciation. Indeed, most of the examples of large hybrid swarms and introgressive hybridization enumerated in this chapter (Table 15.1) are from northern temperate and sub-boreal forests that now exist in or near areas that were covered in ice during the last glacial advances.

The introduction of additional concepts and definitions will occur in later sections of this chapter to insure they are discussed in context with the topic being considered.

Background

Natural hybridization between recognized species is considered common in conifers (Wright 1976; Zobel and Talbert 1984; White et al. 2007), but the presence of the phenomenon is highly skewed across families and genera. A review of over 175 citations from the primary literature as well as reference materials such as Eckenwalder (2009) reveals that confirmed or highly likely natural hybrids have been identified for few (12–14) of the ~71 conifer genera (see Table 15.1). Moreover, hybrids are largely confined to the Cupressaceae (4 of 32 genera) and Pinaceae (5 of 11 genera) and to their genera that have experienced relatively recent phylogenetic radiation, like the cypresses, junipers, pines, spruces, firs, and larches. Recent divergence of incipient species is likely characterized by incomplete breeding barriers that lead to hybridization in secondary contact zones of these species. Hybridization in conifers almost always occurs among closely related species and virtually never among species in different genera. The dearth of verified hybrids in genera of the Podocarpaceae, Araucariaceae, and Taxaceae no doubt reflects one or more of the following conditions: (1) strong reproductive barriers that have developed over long periods of time, (2) modest numbers of species with discrete or allopatric native ranges, and (3) a general lack of close inspection by botanists (Webby et al. 1987).

Table 15.1 Naturally occurring conifer hybrids grouped taxonomically by family and genus. Parental species are identified for each citation and are separated by a "/"

Genus	Subgenus/ section/ subsection	Species	Purpose	Approach	Results	Citations
Araucariaceae						
Araucaria	*sect. Eutacta*	*A. montana/ A. laubenfelsii/ A. biramulata/ A. scopulorum*	Phylogenetics, population genetics	cpDNA, nDNA	Evaluates taxonomic relationships and interprets results in terms of evolutionary rates of divergence, cryptic hybrid/introgression, and incomplete lineage sorting	Gaudeul et al. (2014)
Cupressaceae						
Athrotaxus		*A. cupressoides/ A. selaginoides (A. x laxifolia)*				Eckenwalder (2009)
Chamaecyparis		*C. obtusa/C. pisifera*	Organelle inheritance	cpDNA, mtDNA	Paternal inheritance of both organelle markers	Kondo et al. (1998)
Cupressus		*C. macnabiana/ C. sargentii*	Hybridization	nDNA, cpDNA	Documents hybrids in the genus, notes reports of hybrids in *Athrotaxus* and *Callitris*	Little (2004)
Juniperus		Many	Hybridization	Morphology	A species review of variation and hybridization	Hall (1952)
Juniperus	*sect. Sabina*	*J. maritima/J. scopulorum*	Hybridization, introgression	nDNA, cpDNA	Reports allopatric hybrids and introgression as a function of long-distance gene flow via pollen	Adams (2015)
Juniperus	*sect. Sabina*	*J. osteoperma/J. occidentalis*	Taxonomy, introgression	cpDNA, morphology	Evidence of intermediacy in portions of the western range of *J. osteosperma* and the eastern range of *J. occidentalis*	Terry et al. (2000), Terry (2010)

Juniperus	*sect. Sabina*	*J. tibetica/ J. convallium/ J. saltuaria/ J. przewalskii*	Population genetics	nDNA	Genetic evidence of complex evolutionary histories for junipers on the high-altitude Qinghai-Tibetan plateau (QTP)	Li et al. (2012b)
Juniperus	*sect. Sabina*	*J. scopulorum/ J. horizontalis (J. x fassettii)*				Fassett (1944), Eckenwalder (2009)
Juniperus	*sect. Sabina*	*J. chinensis/ J. sabina (J. x pfitzeriana)*				Eckenwalder (2009)
Pinaceae						
Abies	sect. *Balsamea*	*A. lasiocarpa/ A. balsamea*	Hybridization, introgression	cpDNA, mtDNA	Integrated paleoecology and phylogeography to demonstrate that the hybrid zone between these species is stable	Cinget et al. (2015)
Abies	*sect. Nobilis*	*A. magnifica/ A. procera* (occasionally referred to as *A. magnifica* var. *shastensis*)	Hybridization, introgression	cpDNA	Suggests Shasta red fir is a complex group of hybrids of *A. magnifica* and *A. procera*	Oline (2008)
Abies	*sect. Balsamea*	*A. nephrolepis/ A. sachalinensis*	Hybridization, introgression	spSSR, mtDNA, AFLP	Populations on Sakhalin Is., listed as *A. sachalinensis*, cluster primarily with *A. nephrolepis* (based on STRUCTURE)	Semerikova et al. (2011)
Abies	*sect. Grandis*	*A. grandis/ A. concolor*				Steinhoff (1980)
Abies	*sect. Abies*	*A. alba/A. cephalonica (Abies x borisii-regis)*				Eckenwalder (2009)

(continued)

Table 15.1 (continued)

Genus	Subgenus/ section/ subsection	Species	Purpose	Approach	Results	Citations
Abies	*sect. Abies*	*A. cephalonica/ A. nordmanniana (Abies x equi-trojani)*				Eckenwalder (2009)
Abies	*sect. Momi*	*A. chinensis/ A. fargesii*				Wang et al. (2014)
Larix		*L. sibirica/ L. gmelinii (L. x czekanowskii)*				Eckenwalder (2009)
Larix		*L. lyallii/ L. occidentalis*				Bobrov (1973) in Wright (1976)
Larix		*L. mastersiana/ L. potaninii*				Bobrov (1973) in Wright (1976)
Picea		*35 species*	Hybridization, introgression	cpDNA, mtDNA	An investigation of virtually all spruce species shows incongruence of different markers regarding proof of hybridization/introgression. Reflects on reticulate evolution as a major factor in spruce diversification and speciation	Bouille et al. (2011)
Picea		*P. abies/P. obovata*	Hybridization, introgression	mtDNA, cpDNA, morphology	Karelian spruces between Urals and Scandinavia represent a large hybrid swarm. Equivocal on support of hybrid species designation *P. x fennica*	Krutovskii and Bergmann (1995), Volkova et al. (2014)

Picea		*P. asperata/ P. crassifolia/ P. obovata/ P. meyeri/ P. koraiensis*	Gene flow	mtDNA, cpDNA	CpDNA markers better differentiate between species and introgress less strongly than mt markers due to much higher gene flow. Population differentiation is greater with mtDNA, and diversity is higher with cpDNA	Du et al. (2009)
Picea		*P. engelmannii/ P. glauca*	Taxonomy, diversity	Allozymes	No unique species-specific allozyme markers. Concludes that *P. engelmannii* is a subspecies of *P. glauca*	Rajora and Dancik (2000)
Picea		*P. engelmannii/ P. glauca*	Hybridization, introgression	RAPD	Khasa distinguishes between species with RAPD markers. Hybrid swarms are so common that much of the overlapping range is referred to as interior spruce. Ogilvie suggests most hybrids occur in intermediate elevational bands above 5000 feet, with *P. engelmannii* being subalpine and *P. glauca* sub-boreal. Strong and Hill suggest paleo-hybridization explains morphological indicators of hybridization well to the north and east of current zones of sympatry	Ogilvie and von Rudloff (1968), Khasa and Dancik (1996); Rajora and Dancik (2000), Strong and Hills (2006), Haselhorst and Buerkle (2013), De la Torre et al. (2014a, c)
Picea		*P. engelmannii/ P. glauca*	Hybridization, introgression	SNPs	Contrasts the genic vs genomic model of species boundary maintenance. Suggests it is genic, related to genes under selection for cold adaptation	De la Torre et al. (2014a, c)

(continued)

Table 15.1 (continued)

Genus	Subgenus/ section/ subsection	Species	Purpose	Approach	Results	Citations
Picea		*P. engelmannii/ P. glauca*	Population genetics	nSSR	Suggests abundant evidence for introgression of *P. glauca* genes and even pure *P. glauca* in stands throughout Wyoming, evidence of past hybridization between the two species. Example of trailing edge concept, as species continue to move north post glaciation	Haselhorst and Buerkle (2013)
Picea		*P. glauca/ P. sitchensis*	Hybridization, introgression	cpDNA	Demonstrated hybridity in populations	Szmidt et al. (1988b)
Picea		*P. glauca/ P. sitchensis (P. x lutzii)*	Adaptation	Morphology, isozymes, SNPs	Little first describes the natural hybrid in Alaska (P. x lutzii) or Lutz spruce. Roche describes the hybrid zone in BC. Morphological and isozymes provide clear evidence of hybrids. Grossnickle et al. showed physiological traits were consistent with a hybrid index derived from nDNA measures of hybridity. Sutton et al. demonstrate complex hybrids occur, indicative of true introgression and trend of variation in mix consistent with topographical trends. Hamilton et al. test hypotheses on how hybrid zones are maintained, nature of selection, and how hybrids might ameliorate effects of climate change	Little (1953), Roche (1969), Yeh and Arnott (1986), Szmidt et al. (1988b), Sutton et al. (1994), Grossnickle et al. (1996, 1997), Bennuah et al. (2004), Eckenwalder (2009), Hamilton et al. (2013a, b)

Picea		*P. sitchensis/ P. glauca*	Hybridization, introgression	nSTS, allozymes	Measured introgression across large hybrid zone using a hybrid index for STS markers. Determined longitude better predicted genetics than climate or ecological variables	Bennuah et al. (2004)
Picea		*P. mariana/P. rubens*	Evolution	nSNPs	SNPs used to reveal permeable and impermeable loci to introgression in support of the genic view of species and the effect of selection speciation	Jaramillo-Correa and Bousquet (2003), De Lafontaine et al. (2015)
Picea		*P. mariana/P. rubens*	mt genome recombination	mtDNA	Demonstrates genome recombination	Manley (1972), Bobola et al. (1996), Perron and Bousquet (1997), Jaramillo-Correa and Bousquet (2003, 2005), De Lafontaine et al. (2015)
Picea		*P. meyeri/ P. wilsonii/ P. neoveitchii*	Hybridization, introgression, gene flow	nrDNA, cpDNA, mtDNA	Cp markers best delineate species; mt markers show introgression mostly from resident populations into spreading species	Wang et al. (2016)
Picea		*P. glehnii/ P. jezoensis (P. x notha)*				Eckenwalder (2009)

(continued)

Table 15.1 (continued)

Genus	Subgenus/ section/ subsection	Species	Purpose	Approach	Results	Citations
Picea		*P. purpurea/ P. likiangensis*	Gene flow, introgression	mtDNA, cpDNA	Tested the theories of Currat and Du regarding the direction and extent of introgression depending on which marker types are used. Introgression of mt markers from resident populations into invading populations is typical and strong, but cp markers from invaders should have difficulty sticking unless resident populations are very small	Du et al. (2011)
Picea		*P. koyamai/ P. alcoquiana (P. x shirasawae)*				Eckenwalder (2009)
Picea		*P. wilsonii/ P. likiangensis*	Hybridization, introgression	cpDNA/ mtDNA	Several analytical methods used to conclude that *P. purpurea* is an ancient hybrid of these species, occurring likely 1.3 Ma in the Qinghai-Tibet Plateau and surviving by occupying higher elevations and latitudes than the parental species during glacial advances and retreats. Much like the *Pinus densata* story in the same region	Sun et al. (2014)
Picea		*P. wilsonii/ P. morrisonicola/ P. neoveitchii*	Phylogenetics	nDNA, cpDNA, mtDNA	Results reveal a complex evolutionary history for these three species, introgression being only a slight factor	Zou et al. (2013)

Pinus	*Pinus/Pinus/ Pinus*	*P. yunnanensis/ P. tabulaeformis (P. densata)*	Hybridization, introgression	cpDNA	Propose currently recognized species *P. densata* is of hybrid origin, established during the tertiary period. *P. densata* exhibits cp variation, while the two prospective parental species do not	Wang and Szmidt (1994), Wang et al. (2011), Xing et al. (2014)
Pinus	*Pinus/Pinus/ Pinus*	*P. yunnanensis/ P. tabulaeformis (P. densata)*	Population structure	Allozymes	*P. densata* is relatively diverse compared to other conifers, Fst 12%, P = 71.4 at pop level, H = 0.19. Attributed to hybrid origin. Introgression from *P. yunnanensis* may still be contributing variation	Yu et al. (2000)
Pinus	*Pinus/Pinus/ Pinus*	*P. henryi/ P. tabuliformis*	Hybridization, introgression	nSSR	*P. henryi* is a rare, subtropical pine endemic to China and not widely accepted as a species, often lumped with *P. tabuliformis*. Authors suggest some recent hybridization between the two, possibly as a result of climate warming	Liu et al. (2012)
Pinus	*Pinus/Pinus/ Pinus*	*P. hwangshanensis/ P. massoniana*	Hybridization, introgression	mtDNA, cpDNA, morphology	Authors suggest high rates of intraspecific gene flow should reduce introgressed genes and favor lineage sorting. Using cp and mtDNA markers, they believe they provide evidence of that hypothesis	Zhou et al. (2010a), Luo et al. (2000)

(continued)

Table 15.1 (continued)

Genus	Subgenus/ section/ subsection	Species	Purpose	Approach	Results	Citations
Pinus	*Pinus/Pinus/ Pinus*	*P. mugo/P. sylvestris (P. x rhaetica)*	Population structure	Allozymes	Found populations in hybrid swarms to be more diverse than either parental species. Other traits well studied: morphological, chemical, etc. Wachowiak et al. (2006) concluded hybrids existed, but evidence of a swarm was unfounded at a specific reserve in the northern foot hills of the Tatra mts	Bobowicz and Danielewicz (2000), Christensen (1987), Staszkiewicz (1993), Wachowiak et al. (2006), Heuertz et al. (2010)
Pinus	*Pinus/Pinus/ Pinus*	*P. nigra/P. sylvestris (P. x neilreichiana)*				Eckenwalder (2009)
Pinus	*Pinus/Pinus/ Pinus*	*P. sylvestris (4 subspecies)/ P. densiflora*	Hybridization, introgression	cpDNA	Markers differentiated 3 subspecies of *P. sylvestris* and *P. densiflora*, but not *P. sylvestris* var. s*ylvestriformis*, the latter being more closely related to *P. densiflora*. Allozymes and cpDNA were admixed in this var. suggesting introgressive origin	Szmidt and Wang (1993); Ren et al. (2012)
Pinus	*Pinus/Pinus/ Pinus*	*P. sylvestris* var. *nevadensis*	Gene flow	cpSSR	Rates of pollen introgression in seed (40%) and seedling recruits (15%) suggest gene flow from surrounding plantations is affecting the conservation value of these rare populations in southern Spain	Robledo-Arnuncio et al. (2009)
Pinus	*Pinus/Pinus/ Pinus*	*P. sylvestris/ P. uncinata*				Jasinska et al. (2010)

Pinus	*Pinus/Pinus/ Pinus*	*P. densiflora/ P. thunbergii (P. x densithunbergii)*				Eckenwalder (2009)
Pinus	*Pinus/ Trifoliae/ Australes*	*P. echinata/P. taeda*	Hybridization, introgression	Allozymes, nSSR	A review of many papers but specifically related to SSR work that shows rates of hybridization between the species has greatly increased over the years, due mostly to human causes, and this can have a profound adaptive effect on species	Zobel (1953), Hare and Switzer (1969), Edwards-Burke et al. (1997), Xu et al. (2008), Stewart et al. (2010, 2012), Tauer et al. (2012)
Pinus	*Pinus/ Trifoliae/ Australes*	*P. palustris/P. taeda (P. x sondereggeri)*				Eckenwalder (2009)
Pinus	*Pinus/ Trifoliae/ Ponderosae*	*17 species*	Phylogenetics	nDNA, cpDNA	A look at taxonomic issues in pine in fine detail. Contrast issues related to incomplete lineage sorting and reticulate evolution as a function of hybridization and introgression. Many species cases discussed, including alternative explanations for how *P. coulteri* and *P. washoensis* evolved. A good review of observed hybrid reports in pine	Willyard et al. (2009)
Pinus	*Pinus/ Trifoliae/ Ponderosae*	*P. coulteri/P. jeffreyii*	Biogeography	Allozymes	Presence of private alleles in coulter populations reflects hybridization with jeffrey in many populations. Rare alleles may also represent hybrizymes	Ledig (2000)

(continued)

Table 15.1 (continued)

Genus	Subgenus/ section/ subsection	Species	Purpose	Approach	Results	Citations
Pinus	*Pinus/ Trifoliae/ Ponderosae*	*P. ponderosa/ P. arizonica*	Phylogenetics	cpSSR	A study of two species in the *Ponderosae* that have historically been thought of as having hybrid swarms. Based on 4 cpSSRs, authors concluded that hybrids were actually an unnamed taxon more closely related to Mexican species, and not the product of introgression	Epperson et al. (2009)
Pinus	*Pinus/ Trifoliae/ Ponderosae*	*Many*	Phylogenetics	cpDNA	A complete treatment of phylogenetics of the subsection *Ponderosae* implies considerable historical hybridization and introgression	Gernandt et al. (2009)
Pinus	*Pinus/ Trifoliae/ Ponderosae*	*Intraspecific: P. ponderosa*	Hybridization/ introgression	Allozymes, cpDNA, and mtDNA	Upheld a theory that variable Fst across loci represents selection or historical patterns of separation and reconnection of populations. Very modest levels of introgression across the zone of contact in Montana regardless of marker type, but pollen shows greater exchange than seed	Latta and Mitton (1999)
Pinus	*Pinus/ Trifoliae/ Contortae*	*P. contorta/ P. banksiana (*P. x murraybanksiana*)*	Hybridization, introgression	Morphology	Review of natural and artificial hybrids in *P. contorta* and subsection *Contortae* in general	Critchfield (1980)
Pinus	*Pinus/ Trifoliae/ Contortae*	*P. contorta/ P. banksiana*	Hybridization, introgression, population genetics	Allozymes, morphology	Quantifies introgression based on multiple measures and indices, with speculation on nature of introgressive events and biogeographic interpretations	Wheeler and Guries (1987)

Pinus	*Pinus/ Trifoliae/ Contortae*	*P. contorta/ P. banksiana*	Hybridization, introgression	Gall rust susceptibility	*P. contorta* and hybrids were significantly more susceptible to rust infection than *P. banksiana*, regardless of source of rust. Wu et al. show data on proximity of *P. contorta* stands to *P. banksiana* range related to resistance to 4 pests	Yang et al. (1999), Wu et al. (1996)
Pinus	*Pinus/ Trifoliae/ Contortae*	*P. contorta/ P. banksiana*	Hybridization, introgression, population genetics	RAPD	Ye argues that the pattern of hybrids is indicative of secondary contact, and less so, primary intergradation. Supports introgression of genes into both species	Ye et al. (2002)
Pinus	*Pinus/ Trifoliae/ Contortae*	*P. contorta/ P. banksiana*	Hybridization, introgression	Morphology	Evaluates PCA and DFA for quantifying individual tree and population level of introgression or gene mixing. Reinforces relationship between taxa and environment: Hybrid sites = hybrid trees. DFA works a little better than PCA	Gleiker and Carroll (2011)
Pinus	*Pinus/ Trifoliae/ Contortae*	*P. contorta/ P. banksiana*	Hybridization, introgression	nSSR	They calculate genetic ancestry of seedlings based on molecular marker data using Bayesian algorithms. Show ancestry can be closed predicted by environmental factors like soil moisture and elevation. Hybrid zone much larger and more complex than previously recognized. Range limits changed	Cullingham et al. (2012)

(continued)

Table 15.1 (continued)

Genus	Subgenus/ section/ subsection	Species	Purpose	Approach	Results	Citations
Pinus	*Pinus/ Trifoliae/ Contortae*	*P. contorta/ P. banksiana*	Hybridization, introgression	mtDNA, cpDNA	mtDNA shows extensive introgression (mt capture) into *P. banksiana*, well into Saskatchewan, likely paleo events prior to the hipsothermal. Cp introgression far less evident and reflective of recent events	Godbout et al. (2012)
Pinus	*Strobus subgenus*	*18 species*	Phylogenetics	cpDNA, mtDNA, nrDNA	Purpose was to tease out incongruence of gene trees in the subgenus *Strobus*. Discusses the premise that cryptic introgression is the cause of incongruence	Tsutsui et al. (2009)
Pinus	*Strobus/ Quinquefoliae/ Strobus*	*P. lambertiana/ P. albicaulis*	Phylogenetics	cpDNA	A classic study of cryptic introgression showing that sugar pine in nNorthern California has a very clear cut geographic boundary between populations that harbor a cp haplotype identical to whitebark pine and one to the south that is different. The line of demarcation also coincides with levels of Cr1 resistance to white pine blister rust	Liston et al. (2007)

Pinus	*Strobus/ Quinquefoliae/ Strobus*	*P. pumila/ P. parviflora* var. *pentaphylla (P. x hakkodensis)*	Hybridization, introgression	cpDNA-SSCP, Senjo cp, and mtDNA	Results suggest pollen-mediated unidirectional introgression from *P. parviflora* var. *pentaphylla* to *P. pumila* across an elevational band at higher elevations. Ito et al. relate how flowering phenology differences restrict F1 hybrids, but hybrids are intermediate in phenology to both parents, and they interbreed quite frequently with both parent species. The 2004 paper shows mt and cp markers moving in opposite directions (introgression). Senjo story is different: hybrid zones exhibit cp haplotypes of *P. parviflora v. peta* and mt mitotypes of *P. Pumila* and mt capture in pure stands of *P. parviflora* – cross always in one direction. Suggests seed flow introgression	Watano et al. (1995), Watano et al. (2004), Ito et al. (2008), Senjo et al. (1999)
Pinus	*Strobus/ Parrya/ Cembroides*	*P. monophylla/ P. edulis*	Hybridization, introgression	Morphology	Single-leaf pinyon may have 3 subspecies, all of which appear to exhibit evidence of hybridization/ introgression with *P. edulis* and *P. juarezensis*, suggesting the latter yielded *P. quadrifolia*	Lanner and Phillips (1992)
Tsuga		*T. chinensis/ T. dumosa (T. x forrestii)*				Eckenwalder (2009)
Tsuga		*T. heterophylla/ T. mertensiana (T. x jeffreyi)*				Eckenwalder (2009)

(continued)

Table 15.1 (continued)

Genus	Subgenus/ section/ subsection	Species	Purpose	Approach	Results	Citations
Podocarpaceae						
Phyllocladus		*P. toatoa/P. alpinus/P. trichomanoides* (unverified)				Eckenwalder (2009)
Podocarpus	*sect. Australis*	*P. acutifolius/ P. cunninghamii/ P. nivalis/P. totara*				Webby et al. (1987), Eckenwalder (2009)

In some cases, more than two parents are included in a study of hybridization among all taxa. For most citations, the purpose of the research, the approach used to identify hybrids, and the major results of the study are listed. Hybrid listings missing these data have largely been extracted from Eckenwalder (2009). Hybrids achieving named status are identified parenthetically

Approaches to Identifying Hybrids and Quantifying Levels of Introgression

Before the development of protein and molecular genetic markers, putative hybrids were identified using morphological, physiological, and chemical (terpene) profiles. Parental species were characterized for multiple traits, typically associated with cones, seeds, needles, or chemical content of cortex oleoresins, and putative hybrid individuals were compared with the parental standards. The introduction of hybrid indices (Anderson 1948, 1949) enhanced the power of discrimination by combining the results of multiple traits into a single observation. The hybrid index value, though a quantitative measure of intermediacy, is likely only a crude surrogate or estimator of actual genetic constitution, but is still used today as a relatively cost-effective way of identifying hybrids (e.g., Gleiker and Carroll 2011).

Protein and molecular markers significantly changed the investigative landscape of introgressive hybridization studies and expanded dramatically the types of questions that could be addressed. They also led to the development of several approaches for estimating or quantifying the extent of gene exchange among species, including the use of software such as STRUCTURE 2.3.3 (Pritchard et al. 2000), INTROGRESS 1.1 (Gompert and Buerkle 2010), and TESS 2.3 (Chen et al. 2007).

Early studies employed allozymes to identify species-specific alleles or allele frequency differences between species, and characterize putative hybrid populations based on those allelic differences (Wheeler and Guries 1987; Ledig et al. 1999; Edwards-Burke et al. 1997; Yu et al. 2000). Allozymes were replaced for these types of studies by RAPD, AFLP, and SSR markers in the nuclear genome, but the biggest changes in the ability to identify and analyze hybrids were brought by the introduction of markers originating in the chloroplast and mitochondrial genomes. Today, most studies of conifer introgressive hybridization feature the use of both cytoplasmic marker types, occasionally in combination with nuclear markers (Tsutsui et al. 2009; Semerikova et al. 2011; Zou et al. 2013; Wang et al. 2016). Regardless of genome source, neutral markers have been used almost exclusively in studies of introgressive hybridization in conifers though a few exceptions have appeared in recent years (Hamilton et al. 2013a, b; De La Torre et al. 2014a, c).

As noted previously (Chaps. 2 and 9), organellar and nuclear genomes differ greatly in size, pattern of inheritance, and rates of mutation and gene flow. Organelle genomes in the conifers are almost always uniparentally inherited (Neale and Sederoff 1988), paternally for the chloroplast and maternally for the mitochondria. Gene flow is restricted in the mutationally conservative mitochondrial genome, while the chloroplast genome experiences high rates of gene flow and increased levels of genetic diversity. Results obtained from the three genomes may give incongruous interpretations of (1) the extent and direction of gene flow between species in introgressive events (Watano et al. 2004; Du et al. 2009; Godbout et al. 2012; Wang et al. 2011) and (2) inferences made regarding the evolutionary relationships among species (Cronn and Wendel 2004; Willyard et al. 2009; Currat et al. 2008; Tsutsui et al. 2009).

Evolving Insights

In a seminal paper, Currat et al. (2008) addressed the underlying mechanisms guiding or controlling introgression events. The study was based on spatially explicit simulations that included demographic processes (e.g., population densities) and competition (e.g., selection) and provided several insights and predictions, most of which are supported by studies of introgressive hybridization in conifers. Key points of the paper, contingent on the premise that interbreeding between taxa in essentially unimpeded, include:

- Introgression between two species is asymmetric, moving primarily from the resident species into the invading species in a newly formed zone of hybridization. This is reinforced by the fact that the invading species will occur in lower densities and most frequently mate with members of the local species.
- Introgression should be most common for genes experiencing little gene flow with conspecific neighboring populations. This somewhat counterintuitive point may explain why mitochondrial genes seem to introgress more rapidly in conifers than either nuclear or chloroplast genes. Low densities of the invading species at the onset of hybridization events will ensure fixation of local genes in the population.
- Phylogenetic inferences based on markers showing high rates of introgression, such as mitochondrial genes, will be incongruent with the history of the species studied. Such genes may become entrenched or "fossilized" in the colonizing species and remain long after the local species has disappeared. Patterns of introgression can therefore reveal which species occupied an area first and which was the invader. They will also reveal traces of reticulation.

A corollary to the points noted above is that genes experiencing greater gene flow (e.g., cpDNA) should introgress less, and therefore should be more species specific, a finding that has been reinforced frequently in conifer studies (reviewed in Du et al. 2009). Other distinctions between the organelle genomes are noteworthy. CpDNA markers tend to exhibit considerable genetic diversity, but relatively modest population differentiation – in keeping with most neutral nuclear genes. mtDNA markers exhibit relatively modest or low levels of diversity, but high levels of population differentiation, typically reflected in patterns of geographical variation.

Case Studies of Introgressive Hybridization in Conifers

Among the dozens of examples of species hybridization and putative introgression in conifers, a few cases stand out as being of significant and long-lasting interest. A sampling of these has been chosen to illustrate some of the points made in the preceding sections of this chapter.

Pinus contorta (Lodgepole Pine) x *P. banksiana* (Jack Pine)

Pinus contorta and *P. banksiana* are closely related species in subsection *Contortae* and two of the most widely distributed conifers in the world (Critchfield and Little 1966). They have occupied their extensive sub-boreal ranges in Canada only within the last 4000 to 11,000 years following glacial retreat. Based on morphological data, *P. contorta* appears to have moved far to the east of its current range but began to retreat with the warming and drying climate during the hypsithermal period, concurrent with the ingress of *P. banksiana* from the east (Argus 1966; Schoenike 1976; Rudolph and Yeatman 1982; Critchfield 1985). The current ranges of the two species are sympatric over relatively large areas of central and west-central Alberta and the southwest corner of the Northwest Territories of Canada (Fig. 15.1) within which they commonly hybridize and create hybrid swarms. The populations inhabiting the zone of species overlap likely migrated from one or more refugia east of the Rocky Mountains and south of the continental ice sheets (Critchfield 1980; Wheeler and Guries 1982b; MacDonald and Cwynar 1985) though *P. contorta* may have also survived in refugia north of the ice sheets, in the Yukon (Wheeler and Guries 1982b; Godbout et al. 2008; Strong 2010). Within the zone of sympatry, species occupy different ecological niches defined primarily by edaphic and moisture conditions.

The *P. contorta*/*P. banksiana* complex has been extensively studied for over 50 years. Early identification of hybrids in the zone of overlap relied on cone morphology and orientation traits (Moss 1949) and chemical and physical properties of turpentine (Mirov 1956) to verify hybridity and frequency of occurrence. Increasingly in-depth studies employed morphological markers alone (Gleiker and Carroll 2011) or in combination with allozymes (Wheeler and Guries 1987), measures of gall rust susceptibility (Wu et al. 1996; Yang et al. 1999), RAPD markers (Ye et al. 2002), nSSR markers (Cullingham et al. 2012), and organelle markers (cpDNA or mtDNA) (Wagner et al. 1987, 1991b; Godbout et al. 2012).

The first in-depth study was conducted by Wheeler and Guries (1987). They evaluated 7 diagnostic cone and seed morphology traits and allelic frequencies at 42 allozyme loci for 27 lodgepole pines, 6 jack pines, and 4 putative hybrid populations (Fig. 15.1) to quantify the extent of introgression of neutral genes in the nuclear genome. Additionally, measures of genetic diversity were calculated for all populations based on the allozyme data (Table 15.2). Four of the lodgepole pine populations were outliers, occurring on soils unfavorable to jack pine, east of the contiguous lodgepole pine species boundary. Two outlier populations occurred within the jack pine range and two in areas not traditionally assigned to either species range. Morphological traits were used to build hybrid indices (Fig. 15.2) and, independently, were analyzed using principle component analysis (PCA). Allele frequencies for 11 loci that most distinguished between parental species were subjected to PCA and least-squares analysis. Each of the four data sets was used to calculate quantitative estimates of introgression, interpreted as the proportion of genes in the population derived from *P. banksiana* (Table 15.3).

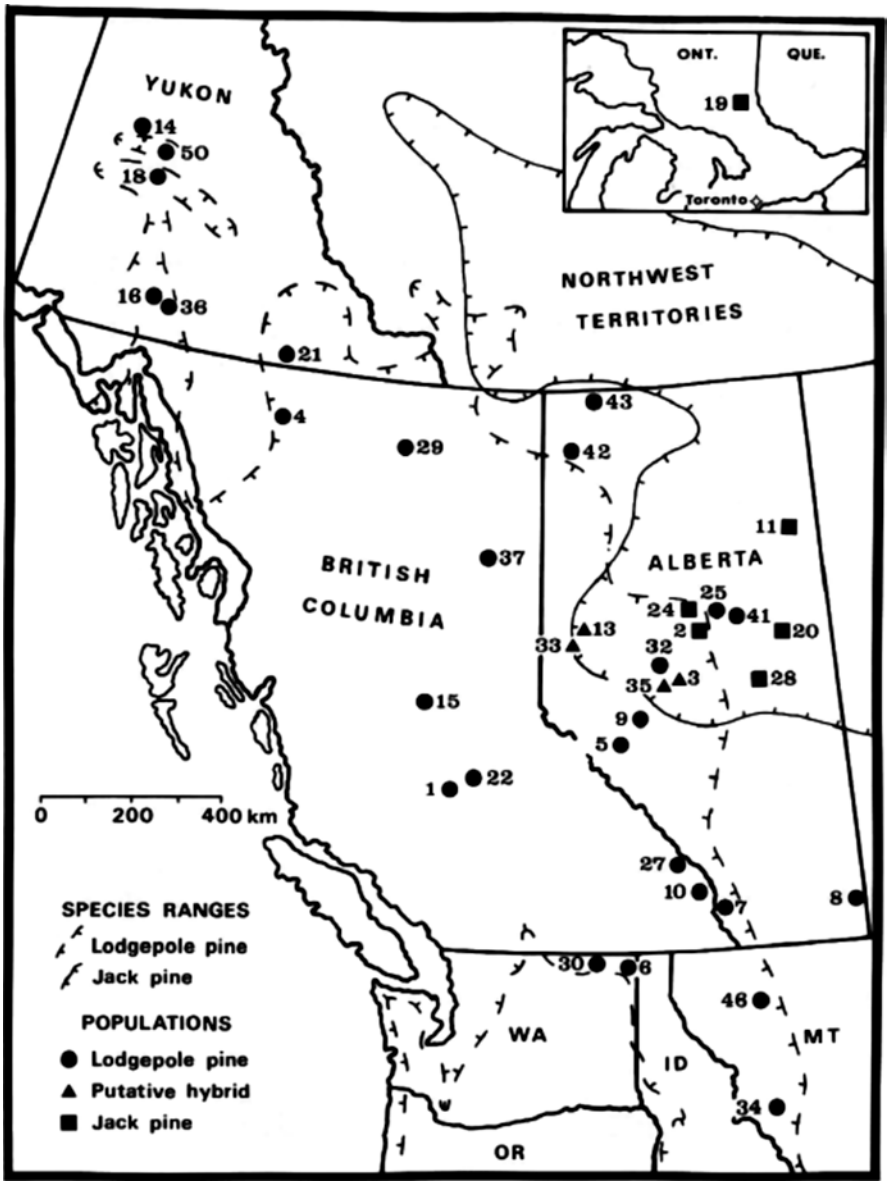


Fig. 15.1 Natural ranges for *P. contorta* var. *latifolia* (lodgepole pine) and *P. banksiana* (jack pine) including sympatric zones and the distribution of 37 sampled populations. (Figure was obtained from Wheeler and Guries (1987), Fig. 1. Page 1877)

Table 15.2 Mean and range of expected heterozygosities (H_e), proportion of polymorphic loci (P), and average number of alleles per locus (A) for *P. banksiana* and *P. contorta*, putative hybrids, and *P. contorta* outlier populations

	H_e	*P*	*A*
P. banksiana			
Mean	0.104 (0.010)	0 44 (0.03)	1.70 (0.07)
Range	0.087–0.117	0.40–0.48	1.62–1.81
Hybrids			
Mean	0.146 (0.013)	0.59 (0.04)	1.94 (0.09)
Range	0.127–0.155	0.55–0.62	1.81–2.02
Outliers			
Mean	0.132 (0.015)	0.59 (0.06)	2.02 (0.09)
Range	0.116–0.137	0.52–0.64	1.93–2.14
Pinus contorta			
Mean	0.118 (0.011)	0.69 (0.06)	1.86 (0.13)
Range	0.102–0.144	0.61–0.82	1.69–2.24

Values in parentheses are standard deviations for among-population within taxa estimates. Table modified from Wheeler and Guries (1987), Table 5. P 1882 (see previous)

Results of the study revealed that:

- *Pinus contorta* and *P. banksiana* are discrete taxonomic entities that can be readily distinguished both morphologically and electrophoretically.
- Estimates of genetic diversity (H_e, P, A; see Chap. 9) were, with a single exception, higher in hybrid and outlier populations than in parental species, supporting the view that introgressive hybridization results in increased genetic variability upon which selection may act.
- The two data sets and four analytical approaches produced remarkably concordant results with respect to parental species characterization and the detection and quantification of introgressive hybridization events.
- A detailed study of the hybrid indices showed individual trees segregating across the *P. contorta*/*P. banksiana* spectrum, implying a complex mix of hybrids and backcrosses. The evolutionary significance of this is that genes are moving freely among taxa and potentially contributing to the adaptability of progeny in a continuously changing environment.
- Multiple localized introgressive events have occurred in current sympatric zones and there is evidence of cryptic introgression in the Yukon, eastern Alberta, and Montana areas.

More recent studies have significantly added to our understanding of the ecological influences on hybridization on this species pair. Cullingham et al. (2012) calculated genetic ancestry of seedlings using Bayesian algorithms based on nSSR marker data. They showed ancestry could be closely predicted by environmental factors like soil moisture and elevation, and concluded that the zone of hybridization is much larger and more complex than previously recognized. The relationship

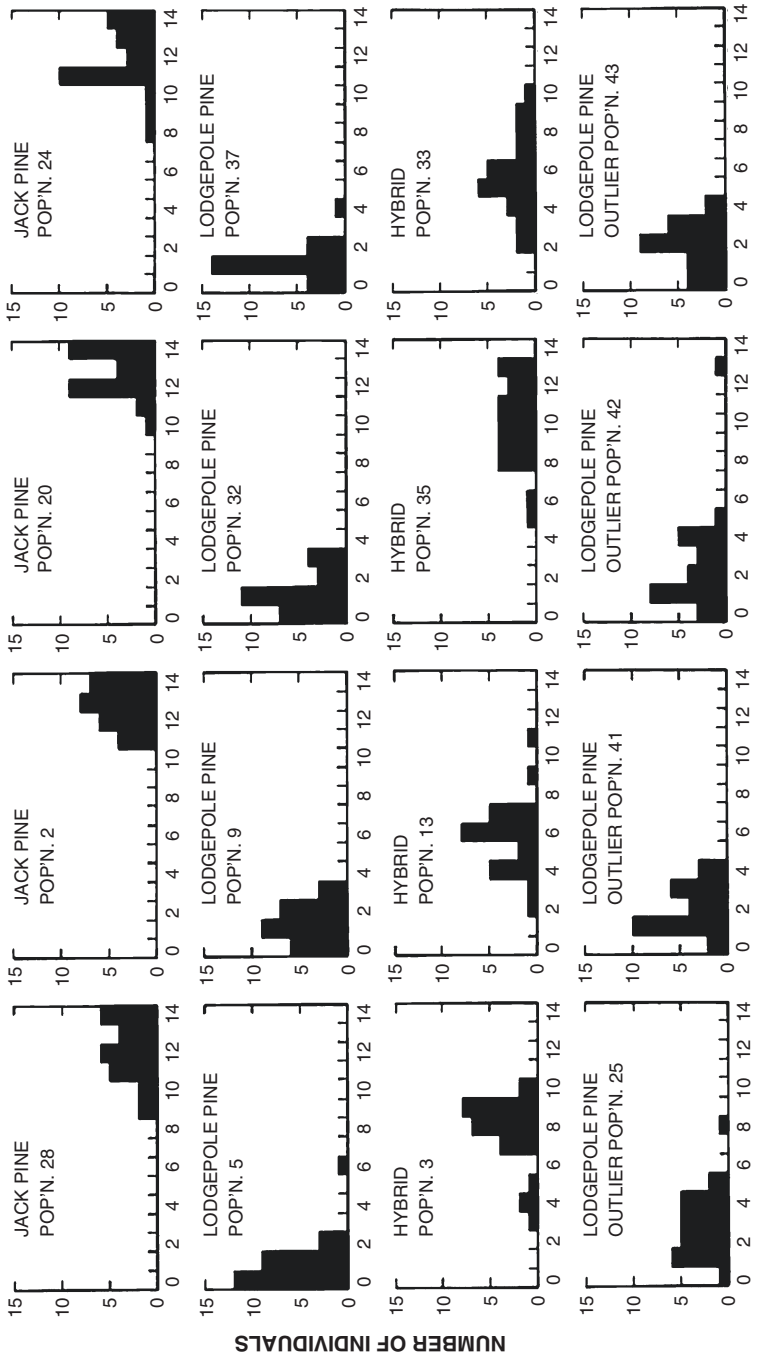


Fig. 15.2 Aggregate hybrid indices for *P. banksiana, P. contorta,* and putative hybrid pine populations, using seven diagnostic cone and seed traits. (Figure obtained from Wheeler and Guries (1987), fig. 2, p 1878, as noted for Fig. 15.1)

Table 15.3 Quantitative estimates of introgression obtained from hybrid indices ($L1$), PCA scores of seven morphological traits ($L2$), PCA scores of population gene frequencies ($L3$), and least-squares analysis (m) of population gene frequencies

Population	L_1	L_2	L_3	$m \pm SE_m$
Pinus banksiana				
2	100	100	100	100.0 ± 0.8
11	–	–	98	96.6 ± 0.7*
19	–	–	100	100.0 ± 1.2
20	100	100	100	98.9 ± 1.3
24	95	93	91	93.8 ± 1.8**
28	98	98	98	98.9 ± 1.3
Hybrids				
3	60	75	82	74.3 ± 3.2***
13	42	43	55	52.5 ± 5.0***
33	34	39	55	54.9 ± 4.9***
35	80	76	79	75.6 ± 3.6***
Outliers				
25	15	12	14	9.9 ± 2.1***
41	7	3	9	6.1 ± 1.7***
42	8	10	28	25.0 ± 3.8***
43	7	6	12	8.2 ± 1.5***
Pinus contorta				
1	–	–	–	0.7 ± 0.8
4	–	–	–	2.9 ± 4.4
5	0	0	14	7.4 ± 3.0*
6	–	–	–	3.2 ± 2.2
7	–	–	4	0.0 ± 2.1
8	–	–	–	2.3 ± 2.9
9	2	9	4	0.0 ± 2.0
10	–	–	–	1.6 ± 1.4
14	–	–	14	10.2 ± 3.0**
15	–	–	0	0.0 ± 1.9
16	–	–	9	5.9 ± 2.8*
18	–	–	5	5.7 ± 2.9
21	–	–	0	0.0 ± 2.9
22	–	–	0	0.0 ± 1.6
23	–	–	0	0.0 ± 1.6
27	–	–	–	2.9 ± 1.7
29	–	–	–	2.4 ± 2.5
30	–	–	–	0.8 ± 1.3
32	1	2	0	0.0 ± 1.5
34	–	–	17	8.5 ± 3.5*
36	–	–	–	2.4 ± 3.3
37	0	2	7	2.4 ± 1.8
50	–	–	–	2.6 ± 2.1

All values are interpreted as the proportion of genes derived from *P. banksiana*. Table modified from Wheeler and Guries (1987), Table 3. P 1880

NOTE: Statistical significance denoted at the 0.05 (*), 0.01 (**), and 0.001 (***) vels

between taxa and environment (hybrid sites = hybrid trees) was reinforced by Gleiker and Carroll (2011) who quantified individual tree and population-level introgression using PCA with morphological data.

More recently, a fascinating study based on combined analyses of cpDNA and mtDNA markers looked at patterns of introgression resulting from past and recent gene flow events (Godbout et al. 2012). The authors sampled extensively (50 populations) across the species' ranges in Canada and subjected results to cluster analyses to identify genetic discontinuities among groups of populations and canonical analysis to detect putative associations among cytoplasmic DNA variation, tree morphology, and site ecological features. Genetic data were illustrated by showing the geographic distribution of mitotype and chlorotype frequencies for the populations sampled (Fig. 15.3).

Study results showed that mtDNA introgression was extensive and asymmetric, with *P. contorta* mitotypes extending from the current hybrid zone in Alberta (AB) into central Saskatchewan (SK), consistent with results suggested from earlier observation of morphological and allozyme data (Wheeler and Guries 1987; Gleiker and Carroll 2011). There was virtually no evidence of *P. banksiana* mitotypes introgressing into *P. contorta*. Very weak cpDNA introgression was observed and only in *P. banksiana* populations. The patterns observed support the view that Central Canada was first colonized by migrants from a *P. contorta* glacial refugium located west of the Rocky Mountains, before being replaced by *P. banksiana* migrating westward during the Holocene. Extensive gene flow from conspecific *P. banksiana* populations eventually erased the traces of *P. contorta* cpDNA. These findings directly validate two of the mechanisms controlling introgressive events postulated by Currat et al. (2008), namely that introgression is directionally asymmetric and introgression should be most common for genes experiencing little gene flow with conspecific neighboring populations.

Picea sitchensis (Sitka Spruce) x *P. glauca* (White Spruce) and *P. engelmannii* (Engelmann Spruce) x *P. glauca*

The Canadian province of British Columbia in western North America is home to four spruce species, three of which hybridize relatively freely in areas of sympatry. *Picea glauca* is a boreal species with a range that extends across the entirety of Canada, including most of the province north of the 51st parallel and east of the coastal mountain range. It is sympatric with *P. engelmannii* throughout most of the southeastern corner of the province and with *P. sitchensis* in the Nass and Skeena river valleys in northwest British Columbia. The latter pair of species also hybridizes on the Kenai Peninsula of Alaska, where the hybrid swarms are referred to as Lutz spruce (*Picea x lutzii*; Little 1953). *P. sitchensis* thrives along the Pacific Coast, from California to Alaska, but extends up river valleys where the maritime influence persists till it clashes with continental climates of the interior. The extent of mixing between *P. engelmannii* and *P. glauca* is so considerable that provincial foresters refer to spruce forests as "interior spruce," without attempting to distinguish between

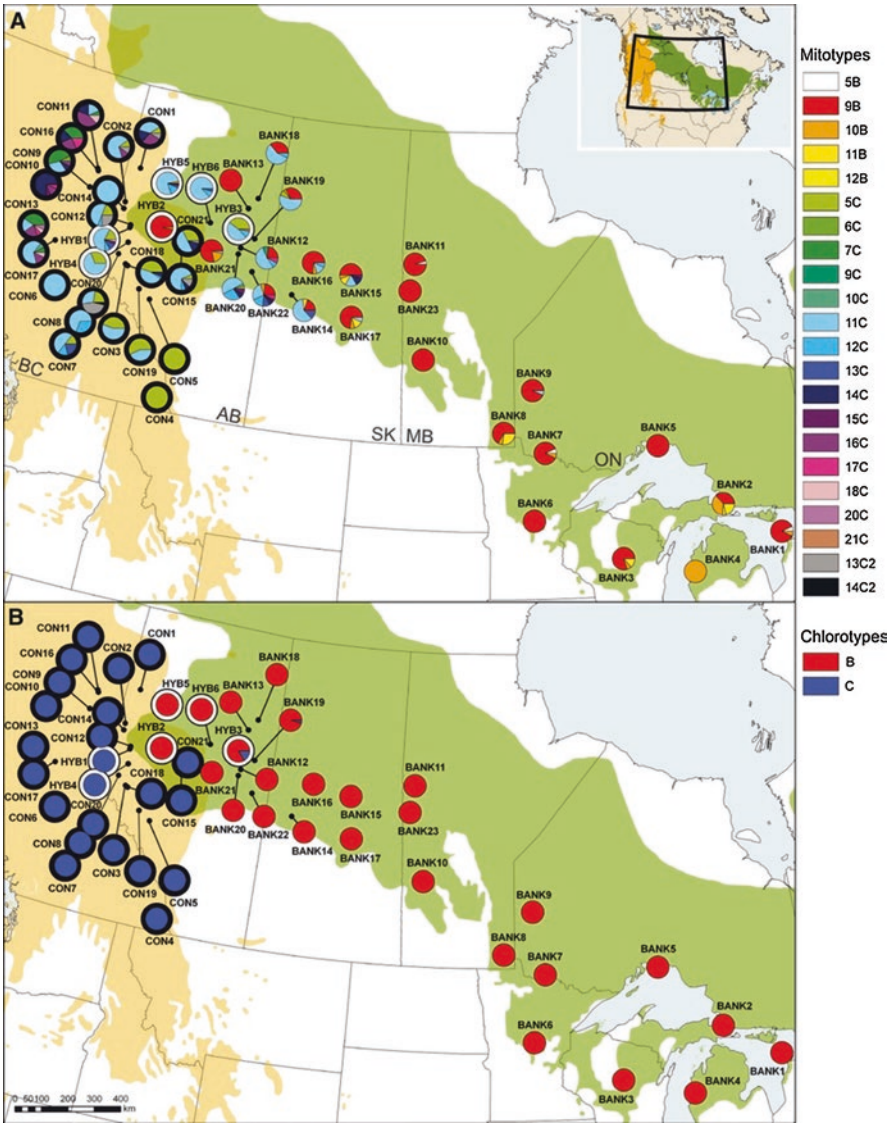


Fig. 15.3 Geographic distribution and frequency of mitotypes (**a**) and chlorotypes (**b**) observed in 50 *Pinus contorta* and *Pinus banksiana* populations. Pie charts highlighted by black circles correspond to *P. contorta* populations, those by white circles to hybrid populations, and the other ones to *P. banksiana* populations. Orange background, the natural range of *P. contorta*; green background, that of *P. banksiana*. Abbreviations of Canadian provinces: AB Alberta, BC British Columbia, MB Manitoba, ON Ontario, SK Saskatchewan. (**a**) Mitotypes are coded with a first number corresponding to the number of repeats observed at locus *nad7* intron 1, followed by a letter and a possible other number indicative of a predominantly *banksiana* type (B) or *contorta* type (C or C2) according to the polymorphism observed at locus *nad1* intron b/c. (**b**) Chlorotype B was defined as *banksiana* type and chlorotype C as *contorta* type. (Figure obtained from Godbout et al. (2012). see http://onlinelibrary.wiley.com/store/10.1002/ece3.294/asset/image_t/ece3294-fig-0001-t.gif?v=1&t=izadayhr&s=8039e5e5964d824baac89892057218e07484dcc8)

the species. Parental species are ecologically separated across much of the zone of sympatry by elevational gradients, *P. engelmannii* occurring at higher, colder elevations than *P. glauca*, but mixing at intermediate elevations. Natural hybrids of both pairs have been extensively studied (Table 15.1), in part driven by the needs of operational forestry practices that seek to reforest cutover lands with the most adapted materials and to market the harvested materials for optimal purposes. Several recent studies have used plant collections from across their respective zones of hybridization, grown in common gardens, to address important questions about the nature of speciation and the maintenance of species boundaries despite long histories of interspecific gene flow. These studies are novel in that they use single nucleotide polymorphisms (SNPs) from candidate genes to assess the extent of admixture and introgression. Such markers may be under selection.

Based on ecological niche modelling, it is likely that *P. glauca* and *P. engelmannii* may have enjoyed moving zones of sympatry for most of the last 21,000 years (De La Torre et al. 2014a) during which time species integrities have remained intact. What are the genetic mechanisms that permit interspecific gene flow but maintain species integrity? In studies conducted by De La Torre et al. (2014a, c) on the *P. glauca*/*P. engelmannii* complex in British Columbia, the authors addressed this question by posing a more specific question: are genomes or genes the units of speciation? They note that under the most widely recognized biological species concept (BSC), "the genomes of species are coadapted units that are separated from other units by reproductive barriers." As noted in the definitions section of this chapter, the BSC precludes calling taxa that hybridize true species. The "genic" view of speciation proposes that the gene is the unit of species differentiation (Wu 2001) and that reproductive isolation is a consequence of natural selection acting on individual genes. Under this model, it is assumed that species genomes may be semipermeable, where some regions may share introgressed genes and other regions, that house genes conferring adaptation and fitness in specific ecological niches, do not share genes or individuals that possess such introgressed alleles at these loci are rapidly selected against (Strasburg et al. 2012).

This study estimated genome-wide admixture in 745 trees from 9 populations using a panel of 311 candidate gene SNPs to identify loci putatively involved in adaptive differences or reproductive barriers between species (Fig. 15.4). Admixture/introgression was estimated using the programs STRUCTURE 2.3.3, INTROGRESS 1.1, and TESS, and tests for loci potentially under selection were conducted using BayeScan v2.0 and Bayenv 2.0 (see Chap. 12). High levels of admixture and introgression, primarily from Engelmann spruce into white spruce, in the zone of sympatry were detected, with most, but not all, alleles shared among the parental species and their hybrids. The genetic architecture of the hybrid populations suggested that most introgression was historical, and evidence for recent admixture was very modest. While these two species appear to have highly porous genomes, a small number of widely distributed genes, identified as being under selection, were not shared and likely contribute to the maintenance of species differences. These putative adaptive loci represent gene functions associated with adaptation to winter climatic regimes including length of growing season and snow depth. The authors conclude that the

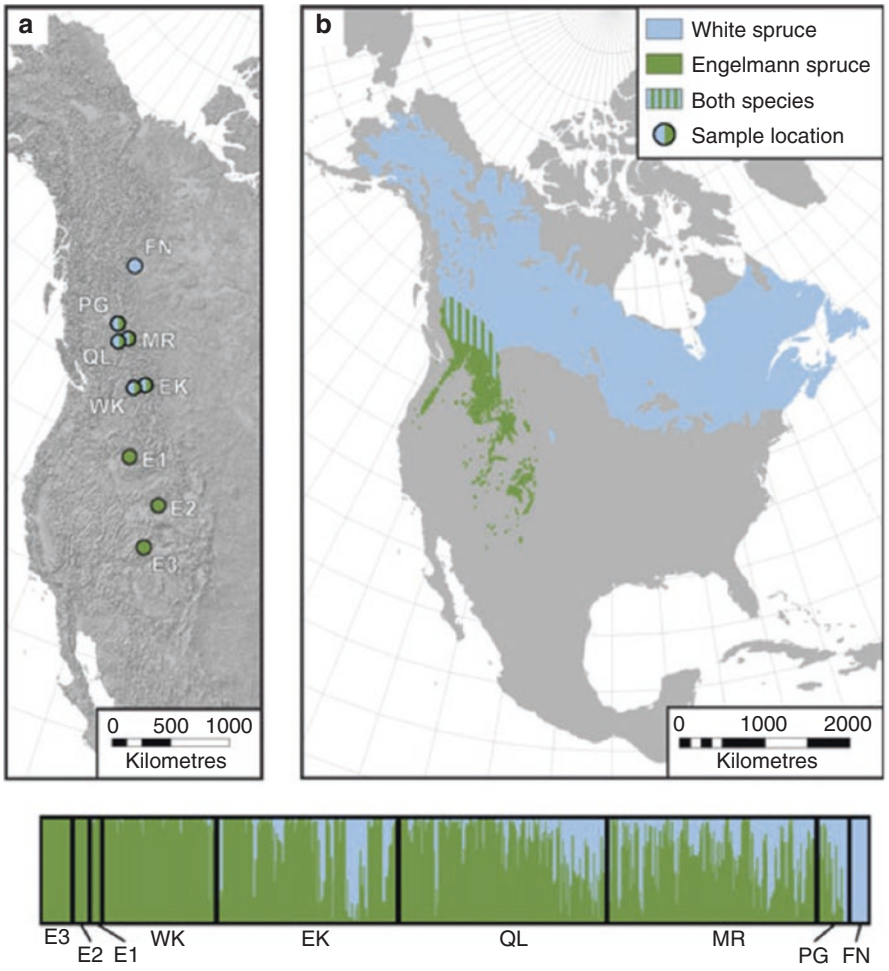


Fig. 15.4 (**a**) Geographical locations of populations of pure *Picea glauca* (FN), pure *P. engelmannii* (E1, E2, and E3), and their hybrids (remaining populations). (**b**) Map showing the location of the current hybrid zone in North America. (**c**) Posterior estimates of the cluster membership for the *P. glauca* x *P. engelmannii* hybrid zone with TESS for $K = 2$. Populations are ordered by increasing latitude from left to right. (Figure obtained from De La Torre et al. (2014a) page 2049)

genic model of speciation holds in this instance, and given the current rate of interspecific gene flow, hybrid populations in this complex may be in the early stages of ecological speciation (incipient homoploid species).

In a separate study, conducted in the same lab as the previous example, Hamilton et al. (2013a, b) evaluated admixture and introgression in the *P. sitchensis*/*P. glauca* complex of northwestern British Columbia, again relying on a set of SNPs putatively associated with adaptive traits in *Picea*. Admixture was estimated using two methods: a maximum likelihood approach using the program *hindex* and a Bayesian

clustering approach implemented in STRUCTURE. The results were used to create a molecular hybrid index. Interspecific heterozygosity (IH) was also estimated, and these values were plotted against the hybrid index as a means of categorizing individuals into genotypic classes (Fig. 15.5). The expectation here is that parental species will register IH and hybrid index values of 0 (for pure *P. glauca*), IH values near 1, and hybrid index values near 0.5 for F_1 individuals and IH near 0 and a hybrid index near 1 for pure *P. sitchensis*. These values would assume that allele frequencies are diagnostic or species specific for loci being studied, which was not the case in this study, though a modest subset of 31 SNPs were nearly diagnostic. Figure 15.5b demonstrates that (1) most of the individuals sampled in the zone of species overlap were hybrids, (2) most of the hybrids were of advanced generations of intermixing and backcrossing with parental species, and (3) the direction of introgression was asymmetric, favoring the movement of *P. sitchensis* into *P. glauca*. In both this case and the former case (*P. engelmannii* x *P. glauca*), the asymmetric gene flow favored the local or resident species, as predicted by Currat et al. (2008).

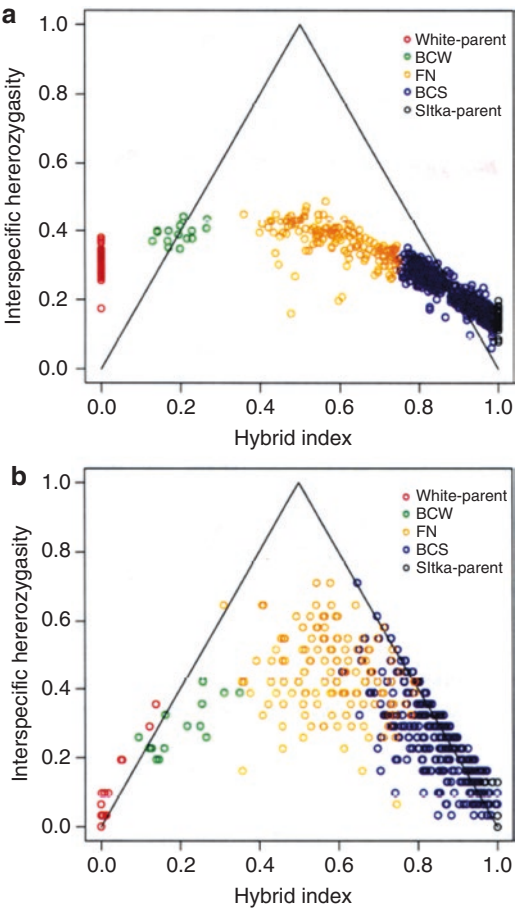


Fig. 15.5 Interspecific heterozygosity (vertical axis) vs. hybrid index (horizontal axis) for 268 candidate gene SNPs (**a**) and a subset of 31 SNPs that exhibited an allele frequency differential (δ) > 0.90 (**b**). Plot based on all loci was used to assign individuals to genotypic classes, where hybrid index 0 = *Picea glauca* (white spruce) and 1 = *Picea sitchensis* (Sitka spruce). Individuals are classified as pure parental species, backcrosses toward either *Picea sitchensis* (BCS) or P. glauca (BCW), or advanced generation hybrids (FN). (Figure obtained from Hamilton et al. 2013b. Page 832, fig. 2)

Hybrid index values were subsequently regressed against an array of geographic and climatic variables (Hamilton et al. 2013b; Fig. 15.6). The results clearly illustrate that marker-based hybrid index values are strongly correlated with climate and geography, particularly annual mean precipitation levels and distance from the ocean. *Picea sitchensis*- and *P. sitchensis*-like hybrid derivatives are clearly favored in wetter environments, while *P. glauca*- and *P. glauca*-like hybrid derivatives are favored in cooler, drier continental climates, as expected. These results are indicative of an environmentally determined "bounded hybrid superiority model of hybrid zone maintenance," much like elevation played a role in determining population structure in the *P. glauca*/*Picea sitchensis* complex noted previously. Finally, interspecific patterns of differentiation, based on F_{ST} values, identified three candidate genes that were clear targets of long-term divergent selection between the parental species and may be acting as part of the genome that is keeping species barriers intact (Hamilton et al. 2013a).

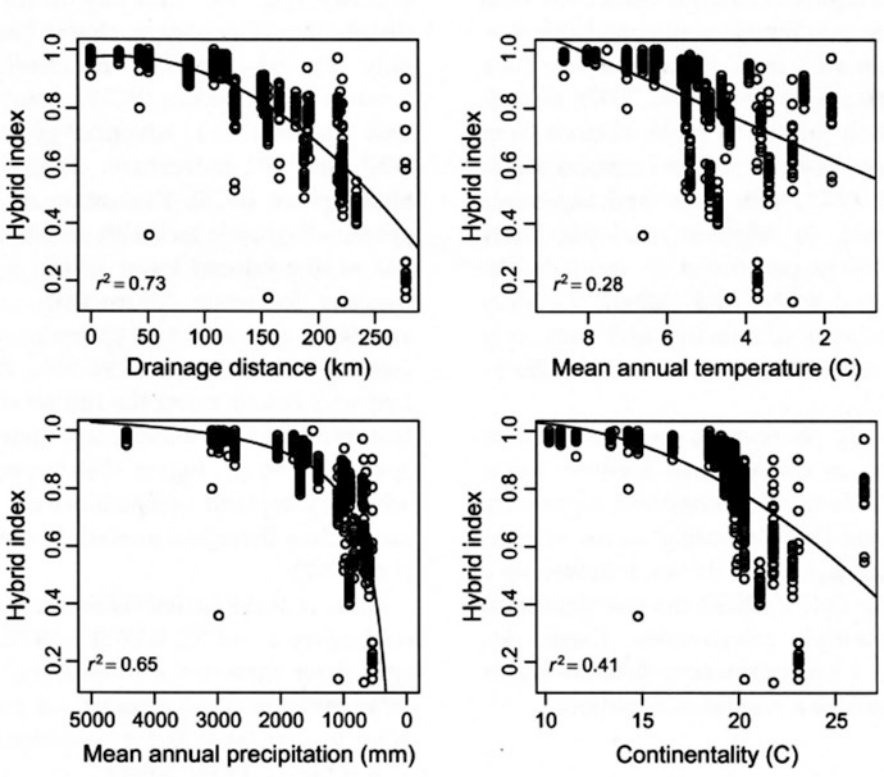


Fig. 15.6 Relationship between 268 SNP-based hybrid index (0 = *P. glauca*, 1 = *P. sitchensis*) and geographic and climatic variables across 721 individuals spanning the introgression zone, including drainage distance (km), mean annual temperature (°C), mean annual precipitation (mm), and continentality (°C). (Figure obtained Hamilton et al. 2013b. Page 834. Fig. 3)

Pinus taeda (Loblolly Pine) x *P. echinata* (Shortleaf Pine)

Our final case study of introgressive hybridization in conifers is a tale of how intensive forest management (plantation forestry) can significantly influence the extent of hybridization in surrounding natural stands. *Pinus taeda* and *P. echinata* are common species of the southeastern United States where they are sympatric across much of their ranges. *Pinus echinata* is favored on upland, drier, and lighter soils and tolerates fire through its ability to sprout new shoots. *Pinus taeda* is favored on heavier, wetter soils common to riparian bottomlands and is fire intolerant. Both species are opportunists and will readily occupy disturbed sites or abandoned fields. *Pinus taeda* was a relatively minor component of pre-colonial forests but has now become the preferred species for reforestation throughout much of the region, with annual reforestation efforts exceeding two billion trees. It is commonly planted beyond its native range on sites once considered too cold or dry.

Natural hybridization between the two species is hindered by differences in the timing of pollen shed and ovulate cone receptivity, but annual weather variations may condense the flowering period for both species to be coincidental. Hybrids between the species have been reported for decades though historically they have occurred in low frequencies in natural stands (Zobel 1953; Hare and Switzer 1969; Edwards-Burke et al. 1997; Stewart et al. 2010). A recent study, based on nSSR markers, of seed collections made in natural stands of both species in the 1950s revealed hybrid frequencies of 4.5% in natural *P. taeda* stands and 3.3% in natural *P. echinata* stands (Stewart et al. 2010). A comparison of stands west and east of the Mississippi River showed hybridization to be significantly higher in the west (8.1% vs 2.1%). A subsequent study made new collections from the same stands (natural regeneration) sampled over 60 years ago, including populations in both sympatric and allopatric portions of the species' ranges (Stewart et al. 2012; Tauer et al. 2012). The findings were quite remarkable and point to significant evolutionary and forest management issues:

- The frequency of hybrids in natural stands of both species increased dramatically: from 4.5% to 27.3% in *P. taeda* and from 3.3% to 45.7% in *P. echinata*. Hybrids were found in all populations sampled, with a single exception, including stands from allopatric regions (Fig. 15.7). These significant and troubling changes to the genetic makeup of pine populations across the southeastern United States are the result of several natural and human-caused factors. Reforestation following timber harvest has been done with *P. taeda* almost to the exclusion of other southern pines, even beyond their natural range. The presence of *P. echinata* has greatly diminished. Climate change has resulted in flowering times of the two species overlapping more than in previous decades. Fire suppression has favored the establishment and growth of *P. taeda* to the detriment of *P. echinata* and other pine species.
- Estimates of population differentiation also changed appreciably over the years, increasing from 0.061 to 0.115 in *P. taeda* stands and from 0.080 to 0.146 in *P. echinata* stands. A reduction in intraspecies gene flow relative to overall pollen loads increases the divergence of stands.

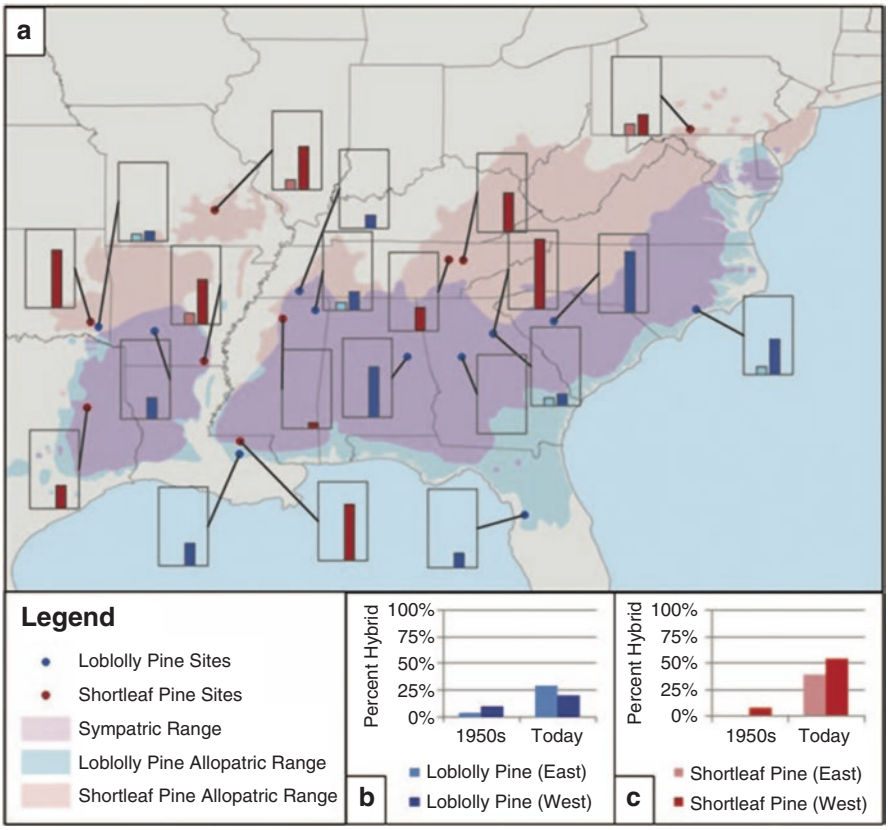


Fig. 15.7 Hybrid proportion of *Pinus echinata* (shortleaf pine) and *P. taeda* (loblolly pine) over time and range. The rate of hybridization of loblolly pine and *P. echinata* has increased since the 1950s. (**a**) Map of the ranges of *P. taeda* and *P. echinata*, including their sympatric range, along with sample sites. The graphs indicate pre-management hybrid percentages on the left and present-day hybrid percentages on the right. The height of each graph is 100% hybrid. (**b** and **c**) The percentage of individuals with hybrid character from pre-management trees and modern trees for (**b**) *P. taeda* and (**c**) *P. echinata*. (Figure obtained from Tauer et al. 2012. fig. 2, page 220)

- The hybrids are predominantly second- and third-generation backcross and intercross genotypes. Species integrities are being eroded across the range, the consequences of which are likely to remain uncertain for years to come.

Hybrid Speciation

The previous examples described introgressive hybridization events that have left clear evidence of interspecific gene flow and apparent long-term success of hybrid genomes in hybrid habitats. In these cases, the eventual outcome is uncertain. Will hybrid swarms eventually differentiate, under changing environmental conditions, and become separate species? Ledig (1998) argues that hybrid speciation should not

occur very frequently in conifers because geographical separation of hybrids and parents seldom occurs. There are, however, a few cases suggesting that new species have arisen or are in the process of doing so.

The ranges of *Picea abies* (Norway spruce) and *P. obovata* (Siberian spruce) are sympatric over large areas of Eastern Europe and the Ural Mountains, where they are known to hybridize rather freely. The hybrids are typically referred to as Karelian spruce and have been treated as a separate species (*Picea fennica*; Volkova et al. 2014). An earlier study was less certain of the species status of Karelian spruce (Krutovskii and Bergmann 1995). Similarly, taxonomic uncertainty over the proper treatment of hybrid populations of the true firs *Abies magnifica* (California red fir) and *A. procera* (noble fir) that occur over large areas of Southern Oregon and Northern California, in the western United States, remains even after a century of discussion (Oline 2008). Generally treated as a variety of red fir (*A. magnifica* var. *shastensis*), many of these populations remain isolated from either parental species following range retraction in recent millennia and may be slowly evolving toward species status.

The somewhat bizarre case of *Pinus maximartinezii* was discussed at some length in Chap. 9. This may represent a unique example of a hybrid pine species that developed from a single seed or two, possibly because of human selection for large seeds in the Mexican highlands (Ledig et al. 1999).

A reasonable case has been made for homoploid hybrid speciation in Northeast China, where both *Pinus funebris* and *P. takahasii* are thought to have been derived from natural crosses between *P. sylvestris* var. *mongolica* and *P. densiflora* (Ren et al. 2012). It should be noted that neither of the putative hybrid pine species identified in the Ren et al. (2012) paper are listed as accepted species in Table 16.3 of this volume.

Perhaps the most convincing evidence of hybrid speciation in conifers comes from another area of China, the Tibetan Plateau. The eastern region of the plateau is believed to have experienced significant increases in altitude (uplift) beginning around 8 million years ago, and the area has experienced many cycles of glaciation and climate change since that time. The region appears to be a locus of hybridization events for conifers and other plants, including reports for spruce (Du et al. 2011; Sun et al. 2014), juniper (Li et al. 2012b), and pine, the last of which is the subject of this discussion. *Pinus densata* has been identified as being a homoploid hybrid species, derived from natural hybridization events between two other regional pines, *P. yunnanensis* and *P. tabulaeformis*. The three species occupy distinctly different ecological ranges today (Fig. 15.8). *Pinus densata* occurs in western Sichuan province and the eastern part of the Tibetan Plateau, where it is endemic to high mountain elevations (2700 to 4200 m; Wang and Szmidt 1994). *Pinus tabulaeformis* has an extensive range in northern and central China (0–2600 m), and *P. yunnanensis* is primarily a lower elevation (600–3100 m) species of southern Sichuan and much of Yunnan provinces with modest extension to higher elevations. The three species have narrow zones of overlap today, but at one time (Late Tertiary), the parental species were believed to have been in contact across a rather broad zone of overlap, likely in the northeastern area of the current range of *P. densata*, leading to multiple introgressive hybridization events. Genetic evidence of these apparent ancient events remains today.

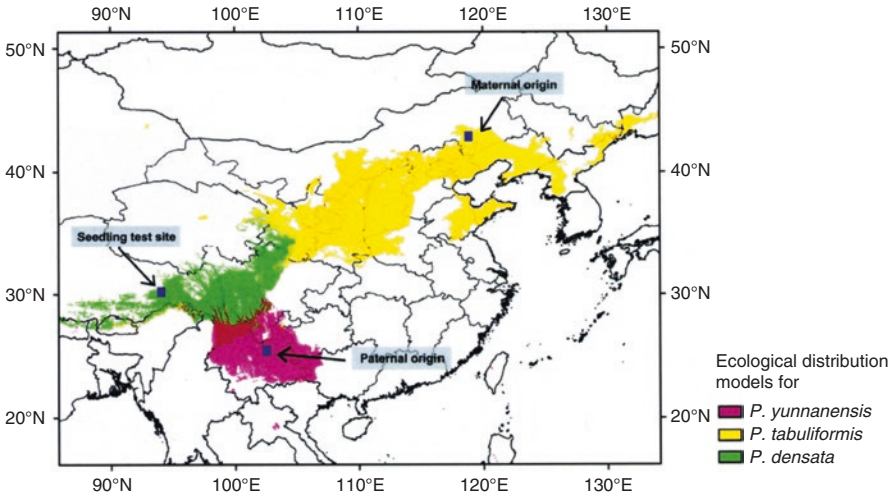


Fig. 15.8 Ranges of three Chinese pine species as defined by an ecological distribution model. *Pinus densata* is believed to have arisen by hybrid speciation from parental species *P. yunnanensis* and *P. tabuliformis*. (Figure obtained from Xing et al. 2014. Fig. 1, p 1892)

The species complex has been extensively studied using morphological traits (Wu et al. 1996; Xing et al. 2014), allozymes (Wang and Szmidt 1994; Yu et al. 2000), and cpDNA and mtDNA molecular markers (Wang and Szmidt 1994; Wang et al. 2011), the results of which suggest a complex evolutionary history in the region. *Pinus densata* exhibits rather high levels of genetic diversity, relative to parental and other pine species for most traits studied (Wang and Szmidt 1994; Yu et al. 2000; Wang et al. 2011), though the most recent study shows that measures of diversity and divergence among stands vary widely across the species range, declining significantly from east to west (Wang et al. 2011; Fig. 15.9). The pattern is most apparent in mtDNA markers but exists in cpDNA markers as well.

Of the 29 mitotypes detected in the study of these three species, 10 are associated with *Pinus tabuliformis* (Pt), 13 with *P. yunnanensis* (Py), and 6 with *P. densata* (Pd), though 4 parental mitotypes were found in adjacent populations of *P. densata* in the eastern areas of the species range. All six *P. densata* mitotypes appear to be putative recombinants of polymorphisms found in the parental species, likely following at least two recombination events. Though considered to be relatively rare, such recombinant mt genomes have been reported in other conifers (Jaramillo-Correa and Bousquet 2005).

As shown in Fig. 15.9, the pattern of mitotypes and chlorotypes is rather complex, with the greatest diversity occurring in the northeastern-most populations of *P. densata* and the virtual fixation of otherwise rare types in the far western populations. Study authors (Wang et al. 2011) suggest the northeastern part of the *P. densata* range represents the ancestral zone of introgression and recurrent migratory events occurred with westward expansion over time. The highly differentiated populations of the species probably occurred because of the species surviving repeated

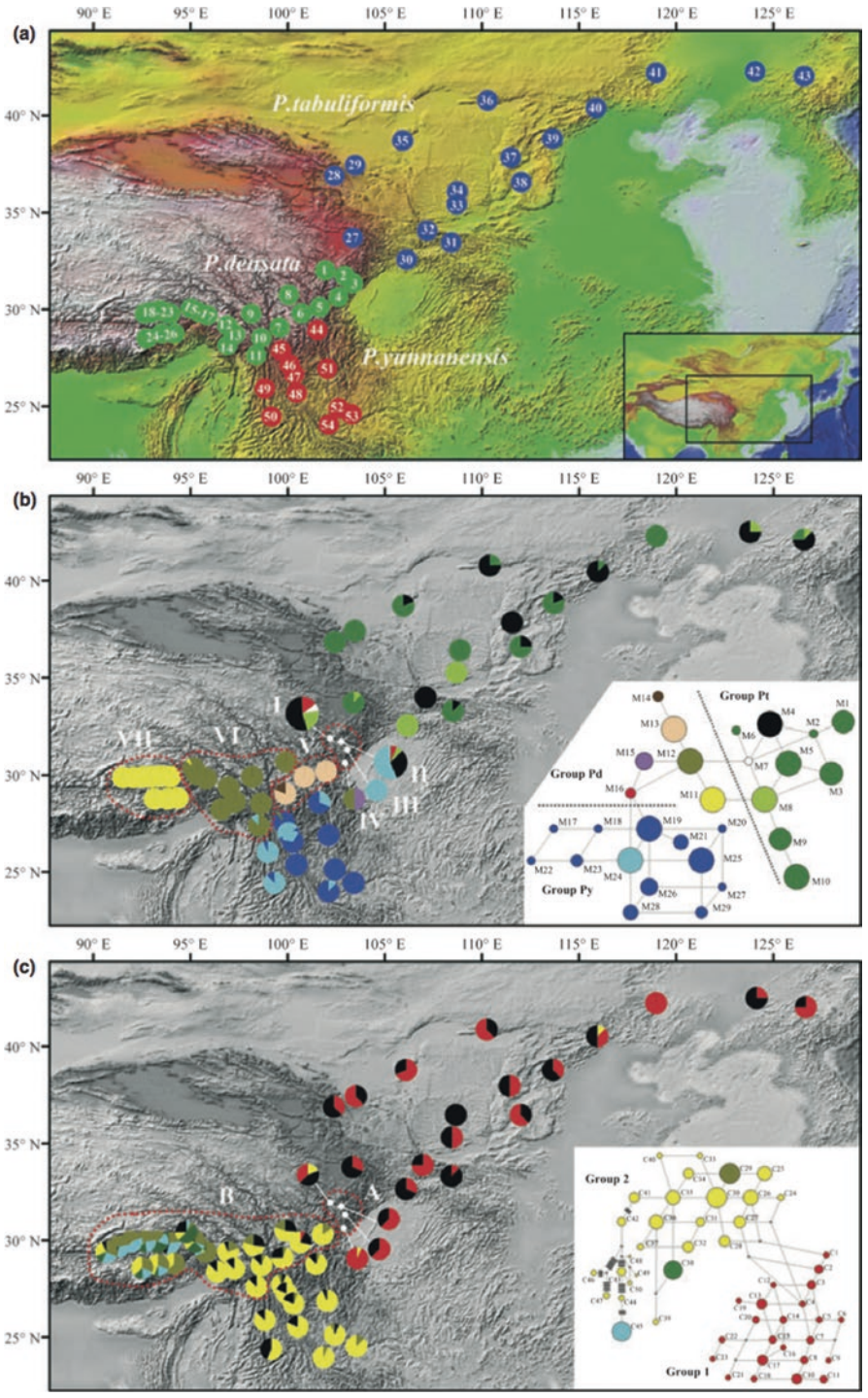


Fig. 15.9 (**a**) Geographic distribution of the 54 sampled populations of 3 pine species. (**b**) The distribution of the 29 mitotypes detected in 3 pine species. Pie charts show the proportions of

(continued)

glaciation events in deep Tibetan valleys – events accompanied by low gene flow and genetic drift. Overall, the species complex in this region represents a fascinating attempt to reconstruct the evolutionary history of speciation.

Artificial Hybrids

As noted in the introduction of this chapter, the creation and evaluation of artificial hybrids in conifers occupied many investigators during the twentieth century, in part because of the considerable success enjoyed by early plant breeders working with crop species. Early work in this area sought to identify hybrids that were superior to either parental species for traits such as growth or tolerance to insect and disease pests (Duffield and Righter 1953; Wright 1959; Garrett 1979). Successful crosses were often planted in permanent plantations for evaluation. Though these efforts revealed some modest successes, pursuit of hybrid vigor has largely been abandoned in conifers. A few exceptions exist, however. Dungey (2001) reported an annual planting rate of about 5200 hectares of hybrid pines across the globe, primarily in subtropical Queensland, Australia, South Africa, and the United States.

In Queensland, testing and subsequent establishment of commercial pine plantations began in the early 1960s (Nikles 1992, 2000). As much as 13% of their plantations were established with hybrids at one time (reviewed in Dungey 2001), the most common of which is that derived from crosses between *Pinus caribaea* var. *hondurensis* and *P. elliottii* var. *elliottii*. In Korea, F_1 hybrids between two North American pine species, *P. rigida* and *P. taeda*, have been planted for 40 or more years (Hyun 1977; Byun et al. 1989). Testing of pine hybrids appears to be ongoing in Zimbabwe and South Africa, as well as in Japan (Dungey 2001).

The larches are notoriously easy to hybridize, and several crosses have found utility in commercial plantations or as ornamentals. The most commonly used hybrid is *Larix x marschlinsii*, until recently referred to as *L. x eurolepis*. This hybrid is a cross between the *L. decidua* (European larch) and *L. kaempferi* (Japanese larch). The seedlings are produced from seed obtained from orchards planted with alternating rows of parental species clones. Seed orchards of other larch species pairs are also reported (*L. sibirica* x *L. decidua*, Lewandowski et al. 1994; *L. gmelinii* var. *japonica* x *L. kaempferi*, Moriguchi et al. 2008). The orchard designs noted here highlight the difficulties of producing commercial quantities of propagules for potentially highly valued hybrids. Until methods to cost-effectively clone and reproduce outstanding conifer hybrids are developed, as it currently can be with poplars, eucalypts, and acacias, the use of conifer hybrids will likely always be restricted to specialty applications.

Crossability studies based on controlled crosses between recognized taxa are also viewed as important contributions to determining phylogenetic relationships.

Fig. 15.9 (continued) mitotypes in each population. (**c**) The distribution and relationships of chlorotypes. Fifty common chlorotypes (which each occurred more than twice) were clustered into two major groups. (Figure and caption modified from Wang et al. 2011. From fig. 1, page 3798)

Table 15.4 Conifer hybrids created through controlled crosses or occurring spontaneously in arboreta and botanical gardens grouped taxonomically by family and genus

Genus	Subgenus/section/ subsection	Species	Results	Origin	Citation
Cupressaceae					
Cupressus		*C. macrocarpa/C. nootkatensis (C. x leylandii)*			Eckenwalder (2009), Kou et al. (2014)
Cupressus		*C. arizonica/C. nootkatensis (C. x notabilis)*			Eckenwalder (2009)
Cupressus		*C. lusitanica/C. nootkatensis (C. x ovensii)*			Eckenwalder (2009)
Taxodium		*T. distichum/T. macronatum*	Concluded that a specific hybrid clone grew faster and was as salt tolerant as *T. distichum* parent	Artificial	Zhou et al. (2010b)
Pinaceae					
Abies	*sect. Grandis; sect. Nobilis; sect. Bracteata*	*A. bracteata/A. concolor/A. grandis/A. magnifica*	Concluded *Abies* has genetic barriers that prohibit crossability among sections	Artificial	Critchfield (1988)
Abies	*sect. Balsamea; sect. Grandis*	*A. balsamea/A. concolor/ A. lasiocarpa/A. fraseri*			Hawley and DeHayes (1985)
Abies	*sect. Balsamea*	*A. koreana/A. veitchii (Abies x arnoldiana)*			Eckenwalder (2009)
Abies	*sect. Abies*	*A. nordmanniana/A. pinsapo (Abies x insignis)*			Eckenwalder (2009)
Abies	*sect. Abies; sect. Grandis; sect. Nobilis; sect. Momi*	*A. alba/A. nordmanniana/ A. procera/A. grandis/A. holophylla/A. numidica*			Kormutak et al. (2013)
Abies	*sect. Pseudopicea; sect. Abies*	*A. pindrow/A. pinsapo (Abies x vasconcellosiana)*			Eckenwalder (2009)

Abies	*sect. Abies*	*A. cephalonica/A. pinsapo (Abies x vilmorinii)*			Eckenwalder (2009)
Larix		*L. decidua/L. kaempferi (L. x marschlinsii) Dunkeld larch*	QTL detected but inconsistently across years and crosses	Artificial	Arcade et al. (2002), Eckenwalder (2009)
Larix		*L. decidua/L. laricina (L. x pendula)*			Eckenwalder (2009)
Larix		*many*			Larsen (1956)
Larix		*L. laricina/L. decidua/ L. kaempferi/L. sibirica*			Meirmans et al. (2014)
Larix		*L. gmelinii* var. *japonica/ L. kaempferi*			Moriguchi et al. (2008)
Larix		*L. sibirica/L. decidua*			Lewandowski et al. (1994)
Picea		*P. engelmannii/P. pungens*	No natural hybrids observed; a very few observed following controlled crosses	Artificial	Ernst et al. (1990)
Picea		*P. jezonensis/P. glauca (P. x saaghyi)*			Eckenwalder (2009)
Picea		*P. mariana/P. omorika (P. x mariorika)*			Eckenwalder (2009)
Picea		*P. jezonensis/P. mariana (P. x moser)*			Eckenwalder (2009)
Picea		*14 species*	Sixteen of 70 controlled crosses produced successful hybrids. Results used to postulate phylogenetic origin of the genus	Artificial	Wright (1955)
Pinus	*Pinus/Trifoliae/ Attenuatae*	*P. muricata/P. radiata/P. attenuata*	Lack of crossability among races of *P. muricata* at distal ends of species range; southern populations of *P. muricata* crosses with many races of *P. radiata*	Artificial	Critchfield and Krugman (1967), Millar and Critchfield (1988)

(continued)

Table 15.4 (continued)

Genus	Subgenus/section/ subsection	Species	Results	Origin	Citation
Pinus	*Pinus/Trifoliae/ Australes*	*P. elliottii/P. caribaea*	Several small effect QTLs for adventitious rooting were found in both species (congruent) but only 1 very large effect locus (40%) in the F2 suggesting that alternative alleles were fixed in the two species. Oddly, the favorable allele came from the unexpected (*P. caribaea*) parent in the synthetic cross	Artificial	Shepherd et al. (2006)
Pinus	*Pinus/Trifoliae/ Australes*	*P. elliottii* var. *elliottii/ P. caribaea* var. *hondurensis*	Study on branch architecture traits revealed QTL in F1 populations	Artificial	Shepherd et al. (2002)
Pinus	*Pinus/Trifoliae/ Ponderosae*	*7 species*	A thorough treatment of crossability of the western yellow pines. Natural hybrids occur for *P. ponderosa* with *P. jeffreyii*, *P. washoensis*, *P. arizonica*, and *P. engelmanni*. Crossability varies by race. Most yellow pines in the west cross to some extent, including the Mexican pines. *P. ponderosa* probably expanded to its current range only in the last 8000 years	Artificial	Conkle and Critchfield (1988)
Pinus	*Pinus/Trifoliae/ Ponderosae*	*P. washoensis/P. ponderosa/ P. jeffreyii*	Crossability and relationships of *P. washoensis* pine	Artificial	Critchfield (1984b)
Pinus	*Strobus/ Quinquefoliae/ Strobus*	*15 species*	Crossing results of *Pinus* section *Strobus*. The white pines generally hybridize rather freely regardless of geography, distance, or continent; the exception is *P. lambertiana* which only hybridizes with *P. armandii* and *P. koraiensis*	Artificial	Critchfield (1986)
Pinus	*Strobus/ Quinquefoliae/ Strobus*	*P. ayacahuite/P. wallichiana (P. x holfordiana)*			Eckenwalder (2009)

Pinus	*Strobus/ Quinquefoliae/ Strobus*	*P. parviflora/P. strobus (P. x hunnewellii)*			Eckenwalder (2009)
Pinus	*Strobus/Parrya/ Balfourianae*	*P. aristata/P. longaeva/ P. balfouriana*	Northern and southern populations of *P. balfouriana* cross readily; *P. longaeva* crosses poorly with northern populations of *P. balfouriana*, but well with southern populations. *P. longaeva* crosses relatively poorly with *P. aristata*	Artificial	Critchfield (1977)
Pinus	*Many*	*Review*	Western yellow pines do not cross with eastern yellow pines in the United States, or elsewhere in the world, but both cross freely within their own groups. Generally poor crossability among the big cone trees (*P. coulteri*, *P. sabiniana*, and *P. torreyana*)	Artificial	Critchfield and Krugman (1967)
Pinus	*Many*	*Many*	A listing of crosses made and tested	Artificial	Garrett (1979), Wright and Gabriel (1958), Wright (1959)
Pinus	*Many*	*Many*			Duffield and Righter (1953)
Tsuga		*T. canadensis/T. caroliniana/ T. chinensis*			Pooler et al. (2002)
Taxaceae					
Taxus		*T. canadensis/T. cuspidata (T. xhunnewelliana)*			Eckenwalder (2009)
Taxus		*T. baccata/T. cuspidata (T. xmedia)*			Eckenwalder (2009)

Parental species are identified for each citation and are separated by a "/". In some cases, more than two parents are included in a study of hybridization among all taxa. Hybrids achieving named status are identified parenthetically

Many of the early studies served both to identify crossability and heterotic potential (Duffield and Righter 1953; Larsen 1956; Wright 1959; Garrett 1979). Subsequent efforts focused on elucidating phylogenetic relationships targeting specific taxonomic groupings (Table 15.4). Notable studies include those for *Abies* (Hawley and DeHayes 1985; Critchfield 1988), *Picea* (Wright 1955), *Pinus* subsection *Attenuata* (Critchfield 1967; Millar and Critchfield 1988), *Pinus* subsection *Balfourianae* (Critchfield 1977), *Pinus* subsection *Ponderosae* (Critchfield 1984b; Conkle and Critchfield 1988), and *Pinus* subsection *Strobus* (Critchfield 1986).

Overall, crossability among conifer species is poor. When it occurs, it is typically between species within closely related groups such as sections of *Abies* (Critchfield 1988) or subsections of *Pinus* (Critchfield 1986). Crosses among subsections of *Pinus* are rare and, between the subgenera, *Pinus* and *Strobus*, nonexistent. Crossing behavior within the two subgenera does differ, however (Critchfield 1986). In the subgenus *Pinus* (hard pines), the ability to hybridize is closely tied to geography. Verified crosses between species native to Eastern and Western Hemisphere hard pines do not exist and are rare even between western and southeastern US hard pines. Quite the opposite is true of the soft pines (subgenus *Strobus*) which appear to have few barriers to crossability as defined by geography. About half of all verified soft pine crosses are between pairs of species native to different hemispheres (Critchfield 1986). More recently Canadian scientists report ~90 successful hybrids among white pine species, though many of these are reciprocal crosses (Daoust and Beaulieu 2004). In spruce, successful crosses involved species that were morphologically similar or had neighboring ranges, or both (Wright 1955).

Summary

The study of hybridization in conifers began with observations of spontaneous hybrids in arboreta and greatly expanded in the mid-twentieth century with artificial crossing experiments to identify superior performing hybrids and clarify taxonomic relationships. Subsequent studies investigated spontaneous hybrids in natural populations and characterized the direction and extent of introgression. In recent decades, the evolutionary importance of hybridization and introgression has been illuminated by studies that address such important principles as which speciation model is most appropriate, what are the mechanisms by which species boundaries are maintained, and how reticulate evolution may confound our phylogenetic interpretations.

Naturally occurring hybrids in conifers seem restricted to a relatively few taxa (~ 14 of 71 genera) and mostly to genera that have experienced relatively recent phylogenetic radiations within the Pinaceae and Cupressaceae. Hybrids commonly occur among closely related species with sympatric ranges that meet in hybrid environments. Historically, hybrids were identified using morphological traits and hybrid indices. Increasingly, sophisticated approaches to the study of hybridization and introgression have benefitted from the use of molecular markers, especially those found in uniparentally inherited organelles (chloroplasts and mitochondria), and improved analytical software. Both theory and observation suggest that

introgression from one species into another in zones of hybridization follows consistent patterns: (1) introgression between two species is asymmetric, moving primarily from the resident species into the invading species in a newly formed zone of hybridization, (2) introgression should be most common for genes experiencing little gene flow with conspecific neighboring populations, and (3) phylogenetic inferences based on markers showing high rates of introgression, such as mitochondrial genes, will be incongruent with the history of the species studied.

Paleobotany, Taxonomic Classification, and Phylogenetics

16

Introduction

The preceding chapters reviewed our knowledge of genetic diversity in conifers from the genome to populations and species, from SNPs to phenotypic variation in morphological and adaptive traits controlled by dozens or even hundreds of genes. In this chapter we look at the diversity of conifers at the species, genus, and family levels, and the evolutionary relationships among them. As the title of this chapter implies, the scope of discussion is large which will almost assuredly result in the superficial treatment of some areas that have garnered considerable scientific enquiry. Our objective, nevertheless, is to capture current views on the number of extant conifer species, how they are classified, from whence they came, and how they are related to one another. In doing so, we hope to avoid confusion associated with the numerous and often nuanced definitions of terms such as taxonomy, systematics, classification, and phylogenetics (Mayden 1992; Stevens 1994; Singh 2004; Wiley and Lieberman 2011). In the treatment presented here we define *taxonomy* simply as the discipline of discovery, description, naming, and classification of groups or taxa, typically species, genera, and families. Taxonomy today is based upon, and richly informed by, phylogenetics, the study of the evolutionary relationships between taxa. Current hypotheses of phylogenetic relationships among and within conifer families, and thoughts on the origins of conifers are summarized here, along with estimates of times of divergence for major taxa based on fossil records and molecular clock studies.

Paleobotany

The conifers and their gymnosperm allies share a conspicuous fossil presence in sediments beginning in the late Paleozoic, well over 300 million years ago (Rothwell et al. 1997; Taylor et al. 2009). They are represented by a diverse and heterogeneous group of fossils that continue to defy construction of a consensus phylogeny.

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_16

Though common, conifer fossils do not represent a continuous evolutionary history of the large and diverse radiation of taxa. The fragmentary and discontinuous nature of fossil representation has resulted in an evolving view of the origin of conifers. It is generally accepted that progymnosperms were the ancestral group from which gymnosperms evolved (Beck and Wight 1988). Found in the fossil record extending from the Middle Devonian (ca. 395 mya) to the Lower Mississippian (ca. 345 mya), the progymnosperms were notable for possessing secondary xylem and a bifacial cambium, with a phloem-like layer peripheral to the xylem – essentially wood that looks very much like conifers today. But unlike seed plants, the progymnosperms reproduced by spores.

The most recognized of the progymnosperms is *Archaeopteris*, considered to be the first modern tree (Beck 1960; Meyer-Berthaud et al. 1999; Fig. 16.1). During much of the Late Devonian (ca. 400 to 370 mya), *Archaeopteris* spread worldwide, thoroughly dominating many environments. It was a large, deeply rooted, and long-lived tree that stabilized stream banks and likely contributed to soil development and aquatic trophic richness (Scheckler 2003). With wood anatomically like modern conifers, *Archaeopteris* grew to impressive size, with boles over 1 m in diameter and 30 meters in height. The trees also exhibited heterospory, with both female megaspores and male microspores, a reproductive system that was a likely precursor of modern seed reproduction.

Fig. 16.1 Reconstruction of the progymnosperm *Archaeopteris*, widely considered to be the first modern tree. (Photo credit: Dennis C. Murphy (Copyright 2005, 2006))

The ancestral origin of modern conifers remains uncertain, however. Early gymnosperms, including the now extinct orders Cordaitales (Rothwell 1988) and Bennettitales, have been identified as potential ancestors, though the latter may be most closely related to the extant Gnetales. The earliest recognized conifers, dating from the Upper Carboniferous Period strata of Europe and North America are presently placed in many families in the order Voltziales. Collectively they are commonly referred to as "walchian" conifers. Extant conifer families begin to appear in the Triassic Period (ca. 250 to ca. 200 mya) and are generally considered to have descended from the walchian conifers. The Mesozoic Era (ca. 250 to ca. 65 mya) is often considered the golden age of conifers given their expansive radiation and floristic dominance following the end of the Permian extinction event(s). Fossil evidence of vesicular-arbuscular mycorrhizal associations in Voltzialean conifers in the Triassic (over 200 mya) has been credited, in part, for the survival and spread of the gymnosperms in extreme paleo-environments (Harper et al. 2015). A visual representation (Fig. 16.2) of a hypothetical timeline for the origin and abundance of extinct and extant conifer families, noting the vague and uncertain relationships among taxa, was provided by Farjon (2008). The figure presumes taxa existed some time before the earliest known fossil records for said taxa.

The fossil record for extant families and genera continues to expand and remains an active area of research for a modest number of investigators worldwide. Revisions of fossil assignment to taxa contribute to changing perspectives of conifer evolution. The earliest known dates for modern families and genera are variously given, often over quite large ranges, due to difficulties in unambiguously assigning fragmentary evidence to specific taxa and dating of geological strata. Generally recognized reliable fossil finds support the appearance of the Podocarpaceae in the Triassic (Townrow 1967), the Araucariaceae (Stockey 1982), Cupressaceae (Escapa et al. 2008), and Taxaceae (Nathorst 1908) in the early Jurassic, the Pinaceae in late Jurassic (Rothwell et al. 2012), and the Sciadopityaceae in the Cretaceous (Ohsawa et al. 1991). Eckenwalder (2009), in his thorough treatment of extant conifers of the world, recognizes reliable dates for the first appearance in the fossil record of many important genera, many of which occur in the early to mid-Tertiary (65 to 40 mya: *Dacrycarpus*, *Dacrydium*, *Agathis*, *Thuja*, *Cephalotaxus*, *Abies*, *Picea,* and *Tsuga*). Notable genera appearing in the Cretaceous (145 to 65 mya) include *Podocarpus*, *Wollemia*, *Metasequoia*, *Pinus*, and *Sciadopitys*.

The earliest conifer fossil records for extant genera are those for *Araucaria* (ca. 190 mya) and the Taxaceae genera *Taxus* and *Torreya* (both ca. 160 mya). Eckenwalder (2009) notes that genera such as *Phyllocladus* and *Cephalotaxus*, while assigned to the Tertiary unambiguously, may be represented by pollen fossils dating as early as 150 mya.

Fossils often occur in assemblages in accurately dated geologic formations, an example of which occurs in the Okanogan Highlands stretching across portions of British Columbia, Canada, and northern Washington State, in the United States. These formations date very nearly to 50 mya. Greenwood et al. (2005) report the assemblages of over 65 taxa, 21 of which are conifers, in various combinations across seven sites in the Highlands, representing a remarkable spectrum of biological, and presumably ecological, diversity. Many of the conifer fossils noted here are

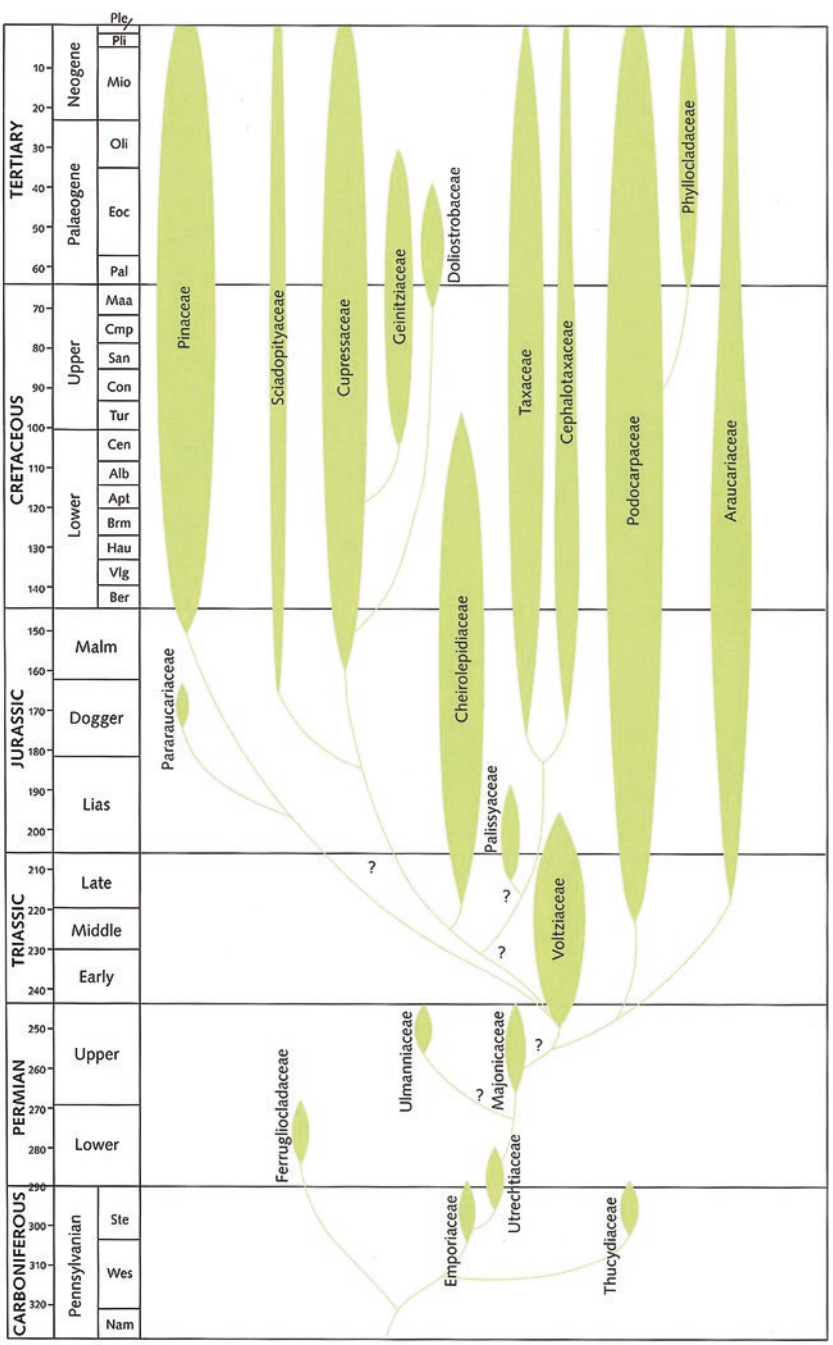


Fig. 16.2 Opposite: A hypothetical evolutionary tree showing origins, relationships, duration, and diversity of the 20 extinct and extant conifer families. The geological time scale is given in intervals of 10 million years in the column on the left, with the era, periods, and abbreviated names of epochs or stages

representatives of some of the earliest occurrences for those genera. The collections included 9 of the 11 extant genera of Pinaceae, missing only *Cathaya* and *Nothotsuga*; the Taxaceae genera *Taxus*, *Torrey*, and *Amentotaxus*; and nine genera of the Cupressaceae (*Calocedrus*, *Chamaecyparis*, *Juniperus*, *Sequoia*, *Taxodium*, *Thuja*, *Cryptomeria*, *Glyptostrobus*, and *Metasequoia*). Interestingly, this section of North America lies near the boundary of two sub-continental terranes that accreted to the North American continental plate between 100 and 50 mya. Whether the floras observed were from those migrating terranes, the original land mass, or some combination of the two remains unknown.

Fossil records for extant species are infrequent and or difficult to confirm, for it is impossible to assess the genetic similarity of fossils and living materials. Critchfield (1984a) summarized the literature for several North American conifers and provided estimates for *Picea glauca* (5.7–23 mya; Wolfe 1969), *Abies balsamea* (0.038 mya), *Pseudotsuga menziesii* (13 mya; Axelrod 1980), and *Pinus monticola* (ca. <10 mya; Wolfe 1969). Ages of two to five million years for recently radiated species of genera such as *Juniperus* and *Pinus* are commonly cited. Developing an understanding of the temporal and spatial appearance of extinct and extant conifer taxa is hampered by the incomplete and fragmentary fossil evidence available, much like building a house with only some of the lumber and hardware in hand, still concerted efforts to do so have continued for over 90 years.

The first widely recognized systematic treatment of fossils recognized as conifers is credited to Florin (1951). His published works, spanning five decades beginning in the 1920s, resulted in a broadly accepted depiction of the systematic diversity of conifers which stood largely unchallenged into the 1980s (Miller 1977; Beck 1988; Rothwell et al. 2005). More recently, studies summarizing findings from the last 30 years have re-evaluated Florin's interpretations and developed new approaches for identifying walchian specimens and resolving systematic relationships among fossil and living conifers (Hernandez-Castillo et al. 2001; Rothwell et al. 2005, 2012). While uncertainty will continue to pervade phylogenetic relationships among extinct conifer taxa, clarity in the evolutionary relationships between known fossil and living taxa will surely improve as more fossils are found, and existing fossils are more carefully characterized. A relatively recent example is provided by a re-examination of a seed cone (*Pararaucaria patagonia*) from Patagonia (Escapa et al. 2012). Using refined methods of study of cone anatomy, the authors assigned the fossil to the extinct family Cheirolepidiaceae and documented anatomical features that provided evidence for the antiquity of pinoid (pine-like) conifers, a possible ghost lineage leading to the origin of the Pinaceae.

Fig. 16.2 (continued) (Alb Albian, Apt Aptian, Ber Berriasian, Brm Barremian, Cen Cenomanian, Cmp Campanian, Con Coniacean, Eoc Eocene, Hau Hauterivian, Maa Maastrichtian, Mio Miocene, Nam Namurian, Oli Oligocene, Pal Palaeocene, Ple Pleistocene, Pli Pliocene, San Santonian, Ste Stephanian, Tur Turonian, Vlg Valanginian, Wes Westphalian). The relative diversity of families is represented by the width of each branch on the tree. The length represents the duration, and the termination when each family became extinct (if it did). Diversity is based on evidence from the fossil record and is in all cases incompletely known, as is the duration of most of the extinct families. (From Farjon (2008))

Taxonomic Classification

The first documented effort to classify plants, including conifers, was produced by the Greek Theophrastus in his *Historia Plantarum*, written sometime between 350 and 287 BC. His classification was based on how plants reproduced. Linnaeus, considered by many as the father of modern taxonomy, published *Species Plantarum* in 1753, and so established the Latin binomial naming convention used today. Linnaeus recognized only 26 conifer species in his early works (from Farjon 2008), and his system of classification, based on sexual traits, was known by him to be artificial. More natural systems of classification that grouped species according to affinities of anatomical and morphological traits such as reproductive structures, leaves, phyllotaxy, and wood structure and, more recently, on biochemical and DNA traits matured over the ensuing 250 years (Pilger 1926; Lawrence 1951; Florin 1951; Hart 1987; Eckenwalder 2009; Farjon 2010).

As previously noted in Chap. 1, conifers have been variously recognized at the level of division or phylum (Pinophyta, Coniferophyta), class (Pinopsida, Coniferae), sub-class (Pinidae), and order (Coniferales). The latter is currently considered as widely accepted (Gernandt et al. 2011), though Christenhusz et al. (2011) recognized three different taxa at the level of order: the Pinales consisting solely of the family Pinaceae, the Araucariales which includes the Araucariaceae and Podocarpaceae, and the Cupressales with families Sciadopityaceae, Cupressaceae, and Taxaceae. We follow the treatment of Gernandt et al. (2011). Taxonomic support for these six families, given recent morphological and molecular studies, appears strong (Eckenwalder 2009; Gernandt et al. 2011; The Angiosperm Phylogeny Website). These six families are recognized here (see Fig. 1.4 this volume), though recent treatments (Farjon 2001, 2010; Farjon and Filer 2013) recognize as many as eight families, including Cephalotaxaceae and Phyllocladaceae, with the previous six. The number of recognized genera and species of conifers also varies modestly, by authority (Eckenwalder 2009; Farjon 2010; Mao et al. 2012; Table 1.1 this volume; Table 16.1).

Taxonomic classification of conifers remains an active and dynamic, though relatively mature field of study. Discoveries of new taxa are now rather rare, but many of the existing taxa remain poorly studied, and opportunities for taxonomic revision remain. Since 1947, 16 new genera have been described, 12 of which were assigned to the Podocarpaceae to reflect reassignment of species previously included in the genus *Podocarpus*. The remaining four represented newly identified species, all of which are considered to represent relicts, with ranges restricted to one or a few populations (Farjon 2008, 2010). All four of these genera are arguably monotypic, and three of them, *Cathaya*, *Wollemia*, and *Metasequoia*, are well known from the fossil record. Two, *Cathaya* and *Metasequoia*, are from China; one, *Xanthocyparis*, from Vietnam; and the other (*Wollemia*) from Australia. Disagreements on the naming and recognition of taxa continue to exist and will likely expand as newly available DNA markers and sequence find increased use in taxonomic and phylogenetic studies (see discussion below).

Table 16.1 Enumeration of extant conifer species, presented by genus and family, as circumscribed in recent encyclopedic treatments by Eckenwalder (2009) and Farjon (2010), with an added perspective on the Cupressaceae provided by the treatment of Mao et al. (2012). Ranges for the number of genera and species recognized reflect the authors' variable taxonomic treatments, as discussed in the text

		Number of species[a]			Total number	
Family (subfamily) [clade]	Genus	A	B	C	Genera	Species
Araucariaceae					3	35
	Agathis	15	15			
	Araucaria	19	19			
	Wollemia	1	1			
Cephalotaxaceae					1	8
	Cephalotaxus	8	0			
Cupressaceae					28 to 32	136 to 162
(Cunninghamioideae)	*Cunninghamia*	2	2	2		
(Taiwanioideae)	*Taiwania*	1	1	1		
(Athrotaxoideae)	*Athrotaxus*	3	2	3		
(Sequoioideae)	*Metasequoia*	1	1	1		
	Sequoia	1	1	1		
	Sequoiadendron	1	1	1		
(Taxodioideae)	*Cryptomeria*	1	1	1		
	Glyptostrobus	1	1	1		
	Taxodium	2	2	2		
(Callitroideae)	*Actinostrobus*	3	3	3		
	Austrocedrus	1	1	1		
	Callitris	15	17	15		
	Diselma	1	1	1		
	Fitzroya	1	1	1		
	Libocedrus	5	6	5		
	Neocallitropsis	1	1	1		
	Papuacedrus	3	1	3		
	Pilgerodendron	1	0	1		
	Widdringtonia	4	4	4		
(Cupressoideae)	*Callitropsis*	0	0	1		
	Calocedrus	4	3	4		
	Chamaecyparis	5	5	5		
	Cupressus	15	17	12		
	Fokienia	1	1	1		
	Hesperocyparis	0	0	16		
	Juniperus	53	54	67		
	Microbiota	1	1	1		
	Platycladus	1	1	1		
	Tetraclinis	1	1	1		
	Thuja	5	5	5		
	Thujopsis	1	1	1		
	Xanthocyparis	2	0	1		

(continued)

Table 16.1 (continued)

		Number of species[a]			Total number	
Family (subfamily) [clade]	Genus	A	B	C	Genera	Species
Phyllocladaceae					1	4
	Phyllocladus	4	0			
Pinaceae					11	195 to 231
(Pinoideae)	*Pinus*	113	97			
	Picea	38	29			
	Cathaya	1	1			
	Larix	11	10			
	Pseudotsuga	4	4			
(Abietoideae)	*Cedrus*	3	2			
	Abies	47	40			
	Keteleeria	3	2			
	Pseudolarix	1	1			
	Nothotsuga	1	1			
	Tsuga	9	8			
Podocarpaceae					17 to 19	152 to 179
[Prumnopityoid]	*Lepidothamnus*	3	3			
	Phyllocladus	0	5			
	Halocarpus	3	3			
	Parasitaxus	2	2			
	Lagarostrobos	1	1			
	Manoao	1	1			
	Prumnopitys	9	8			
	Sundacarpus	1	1			
[Dacrydioid]	*Saxegothaea*	1	1			
	Microcachrys	1	1			
	Pherosphaera	2	2			
	Acmopyle	2	2			
	Dacrycarpus	9	9			
	Falcatifolium	6	5			
	Dacrydium	22	21			
[Podocarpoid]	*Retrophyllum*	5	4			
	Afrocarpus	5	2			
	Nageia	5	5			
	Podocarpus	97	82			
Sciadopityaceae					1	1
	Sciadopitys	1	1			
Taxaceae					5 or 6	23 or 24
	Taxus	10	8			
	Pseudotaxus	1	1			
	Austrotaxus	1	1			
	Cephalotaxus	0	5			
	Torreya	6	6			
	Amentotaxus	6	2			
		613	547			

[a]A = Farjon (2010); B = Eckenwalder (2009); C = Mao et al. (2012)

The current view of conifer taxonomic classification has been influenced by major revisions in the Podocarpaceae and Cupressaceae. The number of currently recognized genera in the Podocarpaceae increased significantly, from 7 to 18 or 19, based on a series of studies by de Laubenfels (1969, 1972, 1987) and Quinn (1982), which reduced and reassigned a number of species previously assigned to the genera *Podocarpus* and *Dacrydium*. Eckenwalder's (1976) recommendation that the Taxodiaceae be folded into the Cupressaceae has been roundly supported, given the distinction between the families seems to have derived almost exclusively from leaf form and phyllotaxy (opposite versus alternate needle/scale arrangements). The result of this change was to reduce the number of generally recognized families from seven to six but preserve the existing number of genera. Recent differences of opinion on placement and naming of taxa throughout the Coniferales may best be discussed following the enumeration of species by genera and family (Table 16.1) based upon the encyclopedic treatments of the conifers by Eckenwalder (2009) and Farjon (2010).

The presentation of extant conifers (Table 16.1) generally follows the arrangement of taxa as derived from Eckenwalder (2009) but has been modified to reflect modest changes and a recent treatment of the Cupressaceae by Mao et al. (2012). In addition, genera are listed, within families, based upon current phylogenetic hypotheses, from the most primitive to the most derived. The most obvious contrasts in classification treatments are reflected in the number of conifer families recognized by the authors as noted above. Eckenwalder places *Phyllocladus and Cephalotaxus* within the Podocarpaceae and Taxaceae, respectively, as has Gernandt et al. (2011) and Christenhusz et al. (2011) while Farjon and Filer (2013) retained them in their own families. Additionally, Farjon (2010) recognizes two genera, *Pilgerodendron* and *Xanthocyparis*, not included in Eckenwalder's (2009) treatment. Christenhusz et al. (2011) concur with Farjon on this matter. The treatments also diverge in the number of species assigned to specific genera, notably *Picea*, *Pinus*, and *Abies* in the Pinaceae, *Podocarpus* in the Podocarpaceae, and *Cupressus*, *Juniperus*, and *Hesperocyparis* in the Cupressaceae. Mao et al. (2012) recognize the genera *Hesperocyparis* and *Callitropsis*, while the others do not, but agree with Farjon (2010) on recognition of *Xanthrocyparis* and recognize many more species of *Juniperus* than either of the other authors.

While these treatments of conifer systematics are presented here as generally accepted and built on consensus of active investigators, challenges remain and will continue to be raised as more information is obtained. Taxonomic treatments today are strongly influenced by detailed phylogenetic studies that increasingly are expanding the number of taxa surveyed and the number and type of traits evaluated, most notably DNA sequences with different cellular origin (nuclear, chloroplast, mitochondrial). Treatments often change as more expansive studies are concluded. For instance, in the past two decades considerable attention has focused on two genera, *Cupressus* (Adams et al. 2014) and *Pinus* (Gernandt et al. 2005), each of which has posed challenges to classification for years. We discuss each briefly here, and later, in the section on phylogenetics.

Cupressus

In recent years many different proposals to erect new genera within the family Cupressaceae or to subdivide *Cupressus* into multiple genera have been made (Farjon et al. 2002; Little 2006; Adams et al. 2009; Mao et al. 2010, 2012; Terry et al. 2012). The progression of these studies has followed expanded use of DNA sequence information, though some, like Little (2006), combine DNA sequence with a wide range of morphological, anatomical, biochemical, and reproductive traits to derive their hypotheses. A common theme to all studies is the naming of one of the most recently described conifer species, *Xanthocyparis vietnamensis.* Farjon et al. (2002) proposed the genus *Xanthocyparis* to include both the newly described species and the New World species known as *Chamaecyparis nootkatensis* (the Alaska yellow cedar). Little (2006) concluded, based on the broad array of traits noted above, that *Cupressus* was paraphyletic, and therefore erected the genus *Callitropsis* and assigned to it all New World cypress species plus the recently named *Xanthocyparis nootkatensis*. Adams et al. (2009), based on a different set of nuclear and chloroplast DNA sequence regions, assigned the New World cypresses to a new genus, *Hesperocyparis*, but left Alaska yellow cedar in the genus *Callitropsis*. A comprehensive study (Mao et al. 2010) concluded the New and Old World cypresses along with *Xanthocyparis nootkatensis* were monophyletic,and sister to *Juniperus.* However, more recent phylogenetic studies (Mao et al. 2012; Terry and Adams 2015), to be reviewed later in this chapter, continue to recognize *Hesperocyparis* to include all New World cypresses, with the exception of *Callitropsis nootkatensis*. The apparent ease with which *C. nootkatensis* hybridizes with other species of New World cypresses (Chap. 15, Table 15.4) might suggest it should be left in *Cupressus* or *Hesperocyparis*, whichever prevails. Until further resolution of this matter occurs, the prudent position may be to reduce *Cupressus*, *Hesperocyparis, Callitropsis,* and *Xanthocyparis* (in North America) to synonymy, and follow Eckenwalder's (2009) assignment of all New World cypresses to *Cupressus* while leaving Farjon's (2010) *Xanthocyparis vietnamensis* standing.

Pinus

Pinus is the largest and best studied of conifer genera, containing as many as 113, or more, recognized species (Farjon 2010; Farjon and Filer 2013). The taxonomic treatment of pines has been the subject of dozens of studies. Krupkin et al. (1996) noted that there had been over 40 classification schemes proposed for the genus, and Syring et al. (2005) suggested that despite ca. 30 published studies in the previous two decades, a well-resolved phylogeny of *Pinus* remained "a work in progress." Several in-depth studies have been published since, and virtually all recognized questions remain regarding the relationships between species and their inclusion in higher taxa such as subsections, sections, and subgenera.

For many years the classification of Little and Critchfield (1969), based largely on comparative morphology and crossing (interspecific hybridization) studies of

pines, was accepted as a standard. They recognized 3 subgenera (*Pinus*, *Strobus*, and *Ducampopinus*, the latter erected to accommodate the unusual species, *P. krempfii*), 4 sections, and 15 subsections for 95 species. Subsequent studies, relying increasingly on DNA sequence or fragment analyses, have offered alternative treatments for portions or the entirety of the genus (Krupkin et al. 1996; Price et al. 1998; Liston et al. 1999; Wang et al. 1999; Geada Lopez et al. 2002; Syring et al. 2005; Gernandt et al. 2003, 2005; Hernandez-Leon et al. 2013). Eckenwalder (2009) adopted the treatment of Gernandt et al. (2005) in his comprehensive treatise on conifers of the world. A comparison of treatments based on morphological traits alone (Little and Critchfield 1969), on DNA sequence of restriction fragments from the chloroplast genome, morphological traits, and biochemical traits combined (Price et al. 1998), and from cpDNA sequence alone (Gernandt et al. 2005) shows a trend to fewer taxonomic subdivisions based on phylogenetic support (Table 16.2). As with *Cupressus*, classification of *Pinus* is sure to evolve as more traits and taxa

Table 16.2 Classification of the genus *Pinus* according to Little and Critchfield (1969), Price et al. (1998), and Gernandt et al. (2005). Parenthetical values represent the number of species assigned to that subsection

Little and Critchfield (1969)	Price et al. (1998)	Gernandt et al. (2005)
Subgenus *Pinus* Section *Pinea* Subsect. *Leiophyllae* (2) Subsect. *Canarienses* (2) Subsect. *Pineae* (1) Section Pinus Subsect. *Sylvestris* (19) Subsect. *Australes* (11) Subsect. *Ponderosae* (13) Subsect. *Sabiniana* (3) Subsect. *Contortae* (4) Subsect. *Oocarpae* (7) Subgenus *Strobus* Section *Strobus* Subsect. *Cembrae* (5) Subsect. *Strobi* (14) Section *Parrya* Subsect. *Cembroides* (8) Subsect. *Gerardianae* (2) Subsect. *Balfourianae* (2) Subgenus *Ducampopinus* Section *Ducampopinus* Subsect. *Krempfiani*	Subgenus *Pinus* Section *Pinus* Subsect. *Pinus* (19) Subsect. *Canarienses* (2) Subsect. *Halepenses* (2) Subsect. *Pineae* (1) Subsect. *Contortae* (4) Subsect. *Australes* (11) Subsect. *Ponderosae* (17) Subsect. *Attenuatae* (3) Subsect. *Oocarpae* (10) Subsect. *Leiophyllae* (2) Subgenus *Strobus* Section *Strobus* Subsect. *Cembrae* (5) Subsect. *Strobi* (16) Section *Parrya* Subsect. *Cembroides* (11) Subsect. *Gerardianae* (2) Subsect. *Balfourianae* (3) Subsect. *Krempfianae* (1) Subsect. *Rzedowskianae* (1)	Subgenus *Pinus* Section *Trifoliae* Subsect. *Australes* (22) Subsect. *Ponderosae* (14) Subsect. *Contortae* (4) Section *Pinus* Subsect. *Pinus* (17) Subsect. *Pinaster* (7) Subgenus *Strobus* Section *Quinquefoliae* Subsect. *Strobus* (18) Subsect. *Gerardianae* (3) Subsect. *Krempfianae* (1) Section *Parrya* Subsect. *Cembroides* (11) Subsect. *Balfourianae* (3) Subsect. *Nelsonia* (1)

are evaluated. What is less certain is whether complete resolution or consensus will ever be reached. A more detailed look at the nature of these studies, the inherent hurdles imposed by use of DNA traits, and the phylogenetic relationships hypothesized among taxa is provided later in this chapter.

Phylogenetics

Phylogenetics is the study of the evolutionary relatedness, or genealogical relationships, among groups of organisms (e.g., species, genera, families). Phylogenetic inferences, or hypotheses, are discovered and informed by comparing suites of characters among groups of taxa. Shared characters, called *synapomorphies*, are used to infer common ancestry among members of a group, or clade, and such groupings are recognized as natural and monophyletic. Phylogenetic systematics is firmly entrenched as the dominant paradigm of systematic biology and has defined how we study evolution (Wiley and Lieberman 2011).

The German entomologist, Willi Hennig (1950, 1966), is often considered the father of modern phylogenetics. Central to his approach was the concept of strict monophyly, and his early works outlined five basic ideas or propositions that defined the development of the discipline (see p. 2–3 in Wiley and Lieberman 2011). Inferred phylogenies are derived through comparative studies and analysis of carefully selected suites of characters, and are graphically represented by phylogenetic trees, or cladograms, many types of which have been developed.

The first comprehensive phylogenetic treatment of conifers was conducted by Hart (1987) using a data matrix of 123 binary and multistate characters representing leaf, stem, floral, embryo, and seed morphology traits as well as chromosome number. The original study generated cladistic relationships of conifer families and genera virtually all of which have been superseded by modern treatments that have included DNA fragment and sequence traits, many of which have been done in combination with the Hart (1987) dataset.

Character Selection

Characters vary in their phylogenetic content. Some similarities that arise through descent may have value for discovering relationships at one level of enquiry (e.g., defining genera) while others are applicable at other levels (e.g., defining families). Careful selection of character traits for the level of study of interest is important.

Modern phylogenetic systematics (i.e., the last 25–30 years) has been dominated by use of arrays or matrices of molecular data based on DNA and amino acid sequences, and molecular fragment size markers, but traditional morphological characters, in use by systematists for decades or longer, remain relevant and even necessary, as in the case of comparative studies between extant and extinct taxa (Gernandt et al. 2016). Miller (1988) noted the many types of characters used in studies on the origin of modern conifer families including karyotypes, antigens,

construction of ovulate cones, pollen grains, foliage types and phyllotaxy, vascular structure and types of secondary xylem, and resin canals. Embryological traits have been particularly well studied. For studies comparing extant taxa, the many and varied morphological and anatomical traits may serve phylogenetic purposes well (Hart 1987). A major challenge facing investigators using such traits is determining whether similarities in character states are truly synapomorphies or the result of convergent evolution. For those who study extinct taxa, the number and quality of traits available for comparison may be severely limited due to incomplete fossil evidence.

The fundamental allure of molecular data arises from the fact that the genomes of all living organisms contain a relatively complete record of their evolutionary past. The more closely related two taxa are, the more similar their DNA sequences. Plants contain three independent genomes: the large nuclear genome containing several thousands of genes located on multiple chromosomes and often millions or even billions of bases (bp) that do not code for functional proteins, the small chloroplast genome (ca. 120,000 to 130,000 bp on a single, circular chromosome), and the larger (ca. 9× > cp genome) but less well-characterized mitochondrial genome (see Chap. 2 for details on genome sizes).

As Sanger sequencing, and later, next-generation sequencing technologies developed, DNA-based characters proliferated at an ever-increasing pace. Choice of which character to use is highly dependent upon the objectives of the studies (Table 16.3). The first DNA markers to be widely used in comparative molecular studies of conifers (and other plants) came from the chloroplast genome, and included restriction fragment length variants and, later, sequence variants of genes and spacer regions, the most notable being the *rbc*L marker sequence (large chain of the ribulose bisphosphate carboxylase gene; Chase et al. 1993; Alverez and Wendel 2003). Pioneering studies by Strauss and co-workers in the late 1980s and early 1990s used cp restriction fragments for phylogenetic studies of *Pinus* and *Pseudotsuga* (Strauss and Doerksen 1990; Strauss et al. 1990) and revealed the absence of a large (ca. 25 kbp) inverted repeat in *Pinus* (Strauss et al. 1988). The absence of the repeat, later shown to be a shared character in other conifers (Raubeson and Jansen 1992), but not in other seed plants, contributed significantly to the view that the conifers are monophyletic.

By the mid-1990s, and for several years thereafter, the inclusion of a nuclear locus, the internal transcribed spacer (ITS) region of the 18S-5.8S-26S ribosomal cistron, became widespread in plant molecular phylogenetic studies, at the generic and infrageneric level, often to the exclusion of any other marker type (Alverez and Wendel 2003). Reliance on a single marker such as this likely led to many misguided phylogenetic hypotheses. Subsequently, other markers were characterized from the chloroplast and nuclear genomes. As technologies and genetic resources improved, phylogenetic studies increasingly included low-copy EST (expressed sequence tag)-based nuclear marker sequences. Today, genome skimming (shallow sequencing) of multiple genomes (Straub et al. 2012) or whole genome sequencing, particularly for the small cp genome, are proposed as a means of dramatically improving conifer phylogenetic and population studies (Cronn et al. 2008;

Table 16.3 Types, characteristics, and utility of molecular markers used in conifer phylogenetic studies.

Genome	Common markers	Characteristics	Utility
Chloroplast (cp)	Short segment sequences: rbcL-Ribulose bisphosphate carboxylase large chain; matK (intron maturase); trnK; rpL16	In conifers, typically uniparentally inherited (paternal), small genome, slow to rapid substitution rates, high-copy number per cell, stable structure, rarity of recombination, primer availability and ease of amplification.	Phylogeography and introgression studies; phylogenetic studies at multiple taxonomic levels depending on nature of the marker
	Restriction length polymorphisms	As above but increased number of potential markers.	
	Whole plastome sequences	Greatly expanded number of variants; reduced cost, increased simplicity of obtaining data	
Mitochondria (mt)	Restriction length polymorphisms (minisatellites) within introns or other noncoding regions	Generally uniparental inheritance (maternal); very slow rate of sequence evolution	Phylogeography and phylogenetic studies at higher taxonomic levels
	nad5	Slow rate of evolution	As above
Nuclear	Nuclear Ribosomal DNA Internally Transcribed Spacer region (nrDNA-ITS 18S-5.8S-25S cistron)	Easily amplified due to high copy number of rRNA genes, highly variable, low evolutionary pressure, nonfunctional sequence, biparental inheritance	Molecular phylogeny at generic and intrageneric (species and subspecies) levels
	Single-copy or low-copy nuclear genes	Biparental inheritance, not subject to concerted evolution, contain codons to limit alignment ambiguity, lower homoplasy than ITS data, nearly limitless numbers, evolutionarily functional	As above

Whittall et al. 2010, Chen et al. 2015, Ruhsam et al. 2015). Clearly, the development and use of molecular markers has revolutionized conifer phylogenetic studies over the last 30 years, leading to new or revised hypothesized evolutionary relationships for most conifer taxa. Their use, however, is not without issue.

All markers possess characteristics that require attention when analyzing and interpreting the data. Failed assumptions resulting from known or unknown trait peculiarities can lead to very different phylogenetic interpretations. These problems may be magnified in studies that seek to impose a molecular clock

interpretation on a phylogenetic tree, providing time estimates for divergence of taxa at one or more nodes of the tree. Such issues include but are certainly not restricted to the following:

- Variation in mutation rates (rate of substitution): the rate of substitution may vary widely among genomes (i.e., organelle versus nuclear: Syring et al. 2005; Willyard et al. 2007), among loci, within a locus across taxa, or between codon positions within a locus (Magallón and Sanderson 2002). Such variation can be useful, when studying phylogeny at different taxonomic levels, for instance, but disruptive if applied to molecular clock studies without knowledge of rate variation.
- Phylogenetic misinterpretations may occur due to natural interspecific hybridization events (reticulation) that are undetected as such (Liston et al. 2007). This is particularly relevant when studies are solely dependent on uniparentally inherited organelle markers.
- Sequencing errors: particularly relevant with first generation sequencing platforms, today's NGS approaches are typically redundant enough to remove casual sequencing errors.
- Multigene families, pseudogenes, and orthology: single-copy loci are infrequent in most conifers. Identifying orthologous genes in taxa being compared can be difficult, particularly when working with EST-based sequences (de la Torre et al. 2006).
- Gene trees of different loci within the same genome or among genomes may result in incongruous phylogenetic interpretations or may not reflect the species tree (Castresana 2007).
- The slow time to allele lineage coalescence in nuclear loci (time to monophyly) in conifer species may exceed species divergence estimates (reviewed in Syring et al. 2007).
- Molecular phylogenetic studies have not always been congruous with those based on morphological or anatomical traits.

While some of these potential pitfalls continue to provide challenges, investigators have addressed many others through careful selection of molecular markers. The trend has been to increase the number of markers used, often representing all three types of cellular DNA, and including loci with known and variable rates of substitution or other biases. Often, morphological or anatomical characters are included with molecular markers to increase the breadth of reference. Construction of hypothetical phylogenies using these large and often complex data sets has been facilitated by significant improvements in analytical approaches and computational capacity. There are three general analytical approaches to inferring phylogenetic relationships: parsimony, maximum likelihood (ML), and Bayesian (Felsenstein 2004; Wiley and Lieberman 2011). Adherents to one or another of these approaches may hold strong opinions as to their specific merits though all approaches appear to have utility, and many recent studies employ all three methods. A further discussion on this topic would exceed the scope of this chapter but interested readers may wish to view any of several books published on the subject in the last 10 years.

We next turn to providing current views or hypotheses of the phylogenetic relationships among extant conifer families, genera, subgeneric taxa, and, in some cases, species. Above all, it should be remembered the topologies shown here, drawn principally from the primary literature, are hypothetical constructs of the real evolutionary relationships that exist. In most cases cited hereafter, investigators have selected a single tree, or perhaps a few trees, from the dozens, hundreds, or even thousands produced, to represent phylogenetic relationships, even though others may have had equally strong statistical support. In short, acceptance of the phylogenetic relationships shown here must be tempered with the knowledge that the addition of further information may, and likely will, alter the results.

The Conifers and Related Gymnosperms

Before reviewing the currently held views of conifer family phylogenetics, it seems prudent to briefly discuss how the conifers relate to the other gymnosperm clades. At the heart of the debate is whether the gymnosperms, which include the cycads, *Ginkgo*, gnetophytes, and conifers, are monophyletic. Considerable disagreement and ambiguity regarding the phylogenetic relationships among the major clades remain, the most notable being the relationship of the gnetophytes to the other clades (e.g., Donoghue and Doyle 2000; Magallón and Sanderson 2002; Hajibabaei et al. 2006; Burleigh and Mathews 2007; Judd et al. 2008; Braukmann et al. 2009; Gernandt et al. 2011). Several hypotheses have been constructed to describe the relationships among the clades, each supported by independent suites of traits. For many decades, beginning in the early twentieth century, the gnetophytes were believed to be a sister clade to the angiosperms, based on shared morphological traits that included wood anatomy (vessels) and chemistry (syringal lignans), flower-like reproductive structures, leaf and pollen structure, etc. This was called the anthophyte hypothesis.

As shared morphological traits were reinterpreted to have likely arisen independently through parallel evolution, alternative hypotheses were put forward which held that the closest relatives of the gnetophytes were the conifers (*gnetifer hypothesis*), or more specifically the Pinaceae family (*gnepine hypothesis*), supported largely by a series of papers based on DNA sequences from plastid, mitochondrial, and nuclear genomes (such as Chaw et al. 1997; Chaw et al. 2000; Bowe et al. 2000; Gugerli et al. 2001). More recently, an extensive study based on two low-copy nuclear loci (*LFY* and *NLY*), taxa representing virtually all gymnosperm genera supported the gnepine hypothesis, placing the cycads, then Ginkgoaceae, sister to the Pinaceae, followed by the gnetophytes and finally the remaining conifer families (see Fig. 1, page 6 in Lu et al. 2014b). Others have found molecular inconsistencies in the gnepine hypothesis which led Gernandt et al. (2011) to exclude the Gnetales from their review of conifer phylogeny. Acceptance of the gnepine hypothesis would render the conifers paraphyletic, requiring a significant revision in thinking on the evolution of seed plants. Clearly, the true relationship between the gymnosperm clades remains unresolved. What does seem convincing is that each of the major gymnosperm clades appears to be internally monophyletic.

Araucariaceae

The Araucariaceae consists of three genera, *Araucaria* (19 species), *Agathis* (15 species), and the relatively recently discovered *Wollemia* (monotypic), all of which, including the family, are considered monophyletic (Setoguchi et al. 1998; Escapa and Catalano 2013). All extant species in the family, except for a few species of *Agathis* that have migrated just north of the equator in Malaysia, occur in the Southern Hemisphere and are of Gondwanan origin, though the fossil record shows the family to have been a major component of Mesozoic forests, with representatives once common in the Northern Hemisphere (e.g., Miller 1977; Stockey 1982). *Araucaria* is divided into four sections: *Araucaria*, with two species, *A. angustifolia* and *A. araucana*, in South America; *Bunya* with the single species *A. bidwillii* from coastal Queensland, Australia; *Intermedia* with the single species *A. hunsteinii* from New Guinea; and *Eutacta* with the remaining 15 species, most of which occur on the island of New Caledonia. DNA studies, including complete plastid genome sequencing comparisons (Ruhsam et al. 2015), reveal little differentiation among the New Caledonian species, implying relatively recent radiation events for this otherwise ancient family. Some, though not all, authorities recognize as many as three sections or groups of *Agathis* species (not treated here). Among the most recognized members of the family are *Araucaria heterophylla* (the Norfolk Island pine), *A. aracana* (the monkey puzzle tree) of Chile, *Agathis australis* (the kauri) from New Zealand, and *Wollemia nobilis* (Wollemia pine) believed to exist today as a single clone. The Araucariaceae is a sister family to the Podocarpaceae.

In recent years (1997–2015), the family has been the focus of several phylogenetic studies (Setoguchi et al. 1998; Knapp et al. 2007; Escapa and Catalano 2013; Ruhsam et al. 2015). Except for Escapa and Catalano (2013), all studies were based exclusively on molecular (DNA) information and, in general, relied on reduced taxon- or gene-sampling approaches or both combined. The most inclusive of these studies sampled 29 species but relied entirely on a single locus (the *rbc*L gene sequence, 1322 bp; Setoguchi et al. 1998). Escapa and Catalano (2013) conducted a more comprehensive study that included 31 of the family's 35 recognized species, eight fossil taxa, and outgroup species of all the remaining conifer families, using both molecular (19 plastid, 2 nuclear, and 2 mitochondrial genomic regions) and morphological (62) characters. There is agreement among the studies on the monophyly of the family and its three genera, but disagreement on how the genera are related to one another. The best supported view is that *Wollemia* and *Agathis* form a clade (agathoid) and are sister to *Araucaria* (Gilmore and Hill 1997; Kunzmann 2007; Liu et al. 2009; Escapa and Catalano 2013; see Fig. 16.3). Support is also shown for the monophyly of the four recognized sections of *Araucaria* (Escapa and Catalano 2013), but little support exists for resolving the relationship among species of the section *Eutacta*, as noted previously.

An interesting application of a phylogenetic study of the Araucariaceae is that which sought to inform the biogeographical query regarding the potential "drowning" of New Zealand during the Oligocene period (Knapp et al. 2007). There remains some controversy as to whether New Zealand was largely or completed submerged during the Oligocene, long after the Gondwanan breakup (ca. 80 mya) that

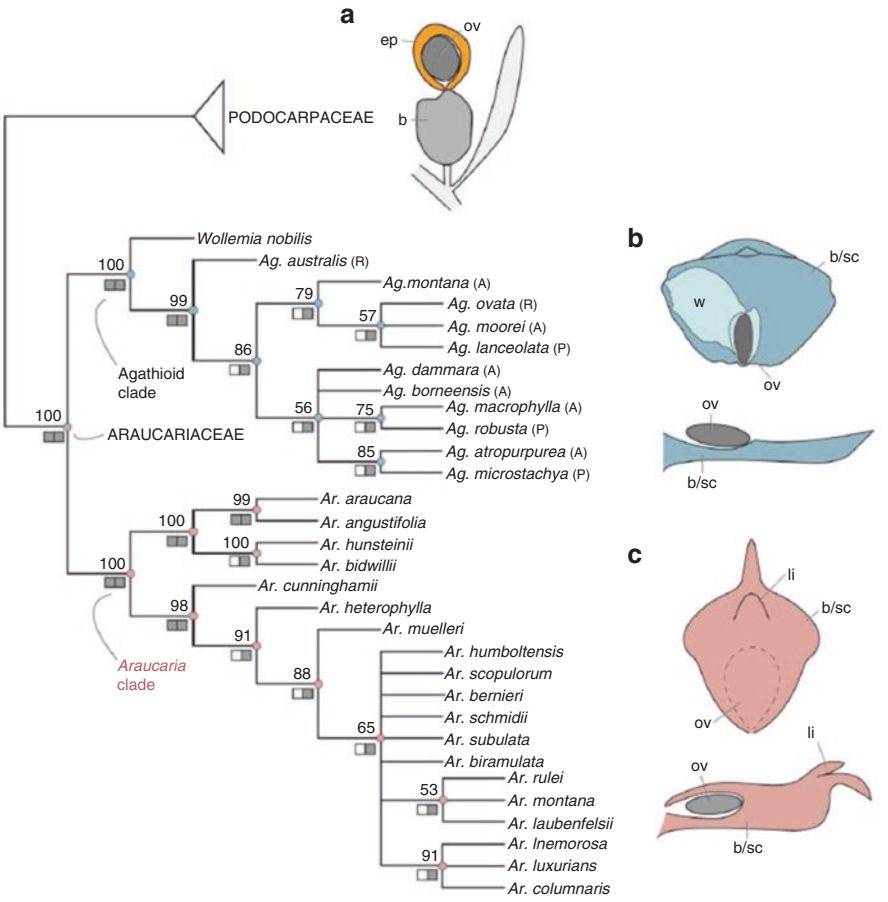


Fig. 16.3 Phylogenetic hypothesis among extant species of the Araucariaceae, including the genera *Wollemia, Agathis* (Ag), and *Araucaria* (Ar), obtained from a combined matrix of morphological and molecular data. Numbers on the nodes indicate jackknife support values; nodes with frequencies less than 50 were collapsed. Squares below each node indicate support of molecular data set (right) and morphological data set (left) if independently analyzed (supported = gray; unsupported = white). Bold letters on the right of *Agathis* species indicate previous assignment to sections: *Agathis* (A), *Prismatobracteata* (P), and *Rostrata* (R) which are no longer recognized. Line drawing indicate bract/scale morphologies in Podocarpaceae (*a*), agathioid clade (*b*), *Araucaria (c).* b bract, b/sc bract/scale complex, ep epimatium, li ligule, ov ovule, w wing. (Figure modified from Escapa and Catalano (2013))

separated the North and South Islands of New Zealand from the other land masses. Today, kauri (*A. australis*) occurs on the North Island. The question is whether it has been there since New Zealand separated during the Cretaceous, in which case it was not likely to have been wholly submerged, or whether kauri migrated to the New Zealand post-Oligocene. Using genomic sequences of the cpDNA loci *mat*K and *rbc*L, as well as the trnD-trnT intergene region, the authors constructed a phylogeny for a modest number of family taxa and imposed a molecular clock analysis

using known fossil dates. Averaging over all genes and analytical approaches, both constrained and unconstrained by fossil minimum dates, they concluded (1) the root split between *Araucaria* and the agathoid clade occurred about 207 mya, (2) *Wollemia* separated from *Agathis* about 150 mya, and kauri separated from other *Agathis* species about 65 mya (Knapp et al. 2007). The authors supported the view, based on dates and the degree of genetic divergence between kauri and native Australian *Agathis* species, that New Zealand was never totally submerged.

Cupressaceae

The Cupressaceae sensu *lato* is comprised of 28–32 genera and 136–162 species (Table 16.1). Eckenwalder's (1976) recommendation that the Cupressaceace sensu stricto be combined with the Taxodiaceae sensu Pilger (Pilger 1926) has been well supported by phylogenetic analyses of molecular (Brunsfeld et al. 1994; Stefanovic et al. 1998; Yang et al. 2012) and morphological or combined molecular and morphological data sets (Hart 1987; Gadek et al. 2000; Farjon and Garcia 2003). *Sciadopitys*, once considered to reside within the Taxodiaceae, was elevated to the monotypic family Sciadopityaceae by Hayata (1931 as cited in Farjon and Garcia 2003), a move subsequently supported by karyotypic (Schlarbaum and Tsuchiya 1984b), molecular (Brunsfeld et al. 1994), and anatomical (Farjon and Garcia 2003) data, among other studies.

The Cupressaceae is represented by a highly variable array of taxa and is the only family to be widely distributed in both Southern and Northern Hemispheres, growing on all habitable continents. The distribution of extant taxa reflects the breakup of the supercontinent Pangaea (Leslie et al. 2012; Mao et al. 2012) with distinct clades assigned to hemispheres. It is posited that the family originally evolved in moist to mesic environments and that Cenozoic climate change coupled with the dramatic vicariance events associated with continental shifting led to the dramatic adaptive responses that permit some taxa to occupy xeric environments today (Pittermann et al. 2012). Over half of the family's genera are monotypic, suggesting a retraction to relic status, particularly for the earliest-evolving clades, while about 65% of the family's species belong to only three genera (*Cupressus, Callitris,* and *Juniperus;* Table 16.1), all of which appear to have experienced relatively recent radiations in dry environments. Among the conifers, representatives of the Cupressaceae hold many distinctions, including the geographically most widely distributed species (*Juniperus communis*, circumboreal), one of the shortest-stature species (*Juniperus horizontalis*, <1.0 m), the tallest species (*Sequoia sempervirens*, >116 m), and the most massive species (*Sequoiadendron giganteum*; stem volume of the largest tree is approximately 1540 m^3). At least 12 species within the family have specimens that exceed 1000 years of age (see Table 1.3 this volume). Family representatives live in the Sahara Desert (*Cupressus dupreziana*) and survive in seasonally standing water (*Taxodium distichum*). Regardless of the great diversity of taxa mentioned here, phylogenetic studies have concluded the family is monophyletic, with all taxa having all descended from a common ancestor that lived around 250 mya (Mao et al. 2012).

With the collapse of the Taxodiaceae into the Cupressaceae, the remaining phylogenetic questions for the family have focused on how the recognized genera are related, and how they might be grouped to reflect their line of descent. Several early phylogenetic studies based on anatomical, morphological, and flavonoid data disagreed on assignment of clades to tribal and subfamilial groupings of the Cupressaceae (Li 1953; de Laubenfels 1965; Gadek et al. 2000). A large study based on two chloroplast loci, the conservative *rbc*L locus and the more variable *mat*K, as well as 45 nonmolecular character traits conducted with representatives of all genera, recognized seven subfamilies (Table 16.1, Fig. 16.4; Gadek et al. 2000), a treatment that has found general support (Mao et al. 2012; Yang et al. 2012) with only minor disagreements, most notably the order in which subfamilies diverged. Yang et al. (2012) place the Athrotaxidoideae after the Sequoioideae, rather than before, as originally proposed by Gadek et al. (2000).

Following the earlier discussion in this chapter related to taxonomic treatments in the Cupressaceae, it should be noted that recent phylogenetic studies (Mao et al. 2012; Yang et al. 2012; Terry and Adams 2015) recognize the genera *Callitropsis, Hesperocyparis*, and *Xanthocyparis*, placing them together in a clade, although there is disagreement among the studies as to how this clade relates to *Cupressus* and *Juniperus*. The variable treatments appear dependent on which molecular data sets and analytical approaches are used in the phylogenetic analysis and the degree of separation among the clades remains small.

Pinaceae

The Pinaceae is the largest of the conifer families, comprised of 11 genera and ca. 230 species (Table 16.1). The family hosts three of the five largest conifer genera with *Pinus* (ca. 113 sp.), *Abies* (ca. 47 sp.), and *Picea* (ca. 38 sp.), all appearing to be the product of recent radiations. Three other genera (*Cathaya, Nothotsuga*, and *Pseudolarix*) are monotypic, and presumably relictual. The genera have been divided into as many as four subfamilies, though the current trend seems to recognize only two: Pinoideae and Abietoideae (Table 16.1; Gernandt et al. 2008; Leslie et al. 2012). Gernandt et al. (2016) inferred phylogenetic relationships among extant genera using both parsimony and Bayesian analyses based on combined morphological character states and plastid and nuclear DNA sequences (Fig. 16.5). The analytical approaches varied only in their placement of the genus *Cedrus.* The parsimony analysis placed *Cedrus* sister to all other Pinaceae genera while the Bayesian analysis placed it sister to the abietoid clade.

Stem ages for the Pinaceae, based on molecular clock estimates, are highly variable and range from ca. 150 mya to ca. 270 mya (e.g., Gernandt et al. 2008; He et al. 2012b; Leslie et al. 2012), while estimates for the divergence of the subfamilies imply the Abietoideae (ca. 200 mya) preceded the Pinoideae (ca. 155 mya; He et al. 2012b). Thus, phylogenetic studies of the Coniferales place the Pinaceae as sister to all other families (Fig. 1.4 this volume), a position which is inconsistent with the fossil record, which dates the family only from the mid-Mesozoic, about 150 mya (Fig. 16.2; Rothwell et al. 2012; Gernandt et al. 2016).

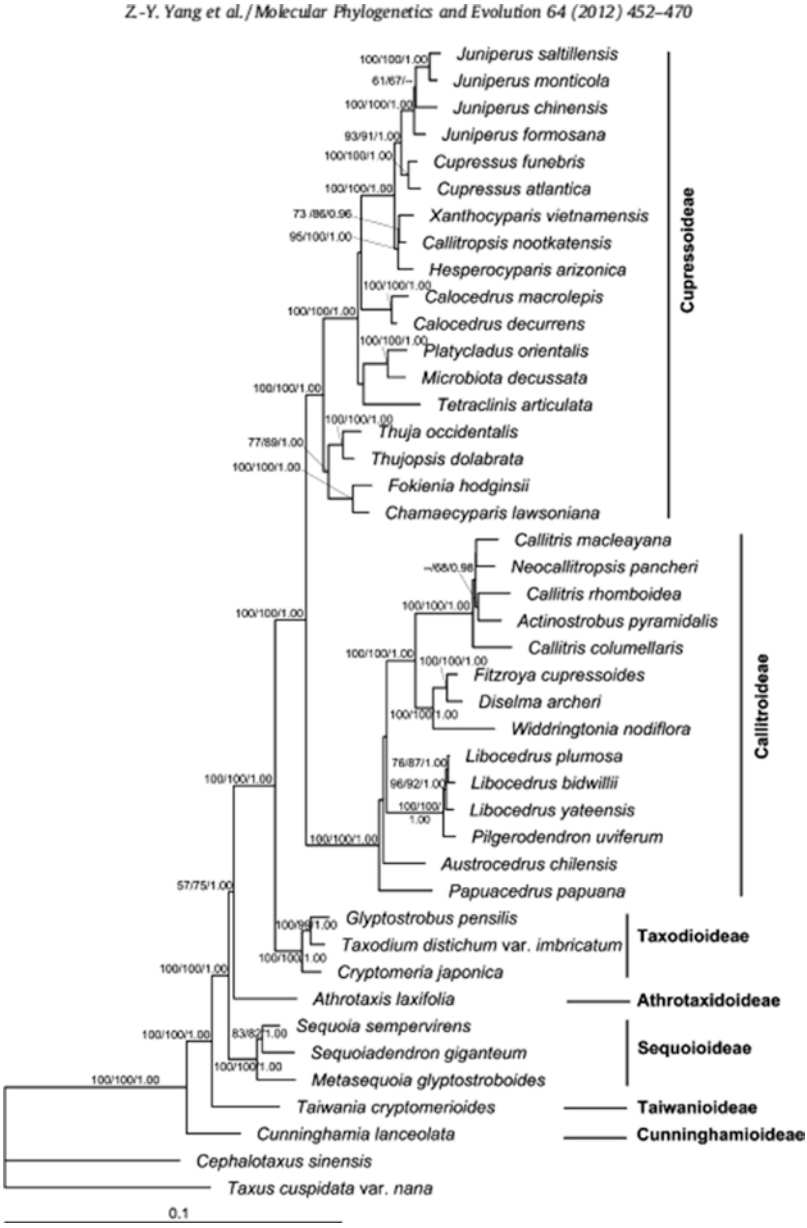


Fig. 16.4 A hypothetical phylogenetic tree of the Cupressaceae sensu *lato* based on DNA sequence variation in four loci (two cp loci and two nuclear loci) delineating six recognized subfamilies (noted in bold lettering to the right of the tree). (Figure is modified from Fig. 5, Yang et al. (2012))

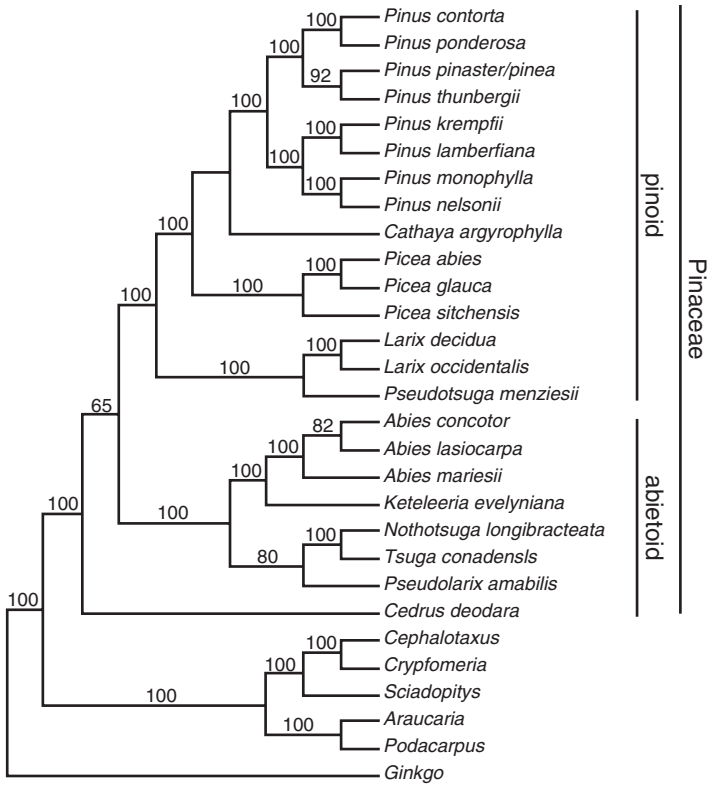


Fig. 16.5 Tree resulting from parsimony analysis of extant species representing 11 genera of the Pinaceae. Tree is based on a single most-parsimonious tree combining data from morphology, nuclear PHYP, and plastid DNA sequence. (From Gernandt et al. (2016))

Pinus The Pinaceae has benefited from a rich phylogenetic literature, a significant majority of which is focused on the genus *Pinus*, based on molecular data analyses. The largest of all conifer genera, *Pinus* has long posed significant challenges for taxonomists and phylogeneticists, and though much has been resolved, in recent years, regarding evolutionary relationships among species, confounding issues remain. The genus is distinguished by a classification system that includes three subgeneric divisions (subgenera, sections, and subsections), the latter two of which have evolved considerably in number and naming priorities over the last 40 years. Early phylogenetic studies based on DNA fragment analyses (Strauss et al. 1990; Krupkin et al. 1996) challenged the long-standing pine classification system proposed by Little and Critchfield in 1969 and substantive revisions were proposed thereafter (Price et al. 1998; Gernandt et al. 2005). Still, questions remained regarding the placement of some species (*P. merkusii* and *P. latteri*) and subsections (*Contortae* and *Krempfianae*) as reflected in a phylogenetic hypothesis offered by Parks et al. (2012; Fig. 16.6).

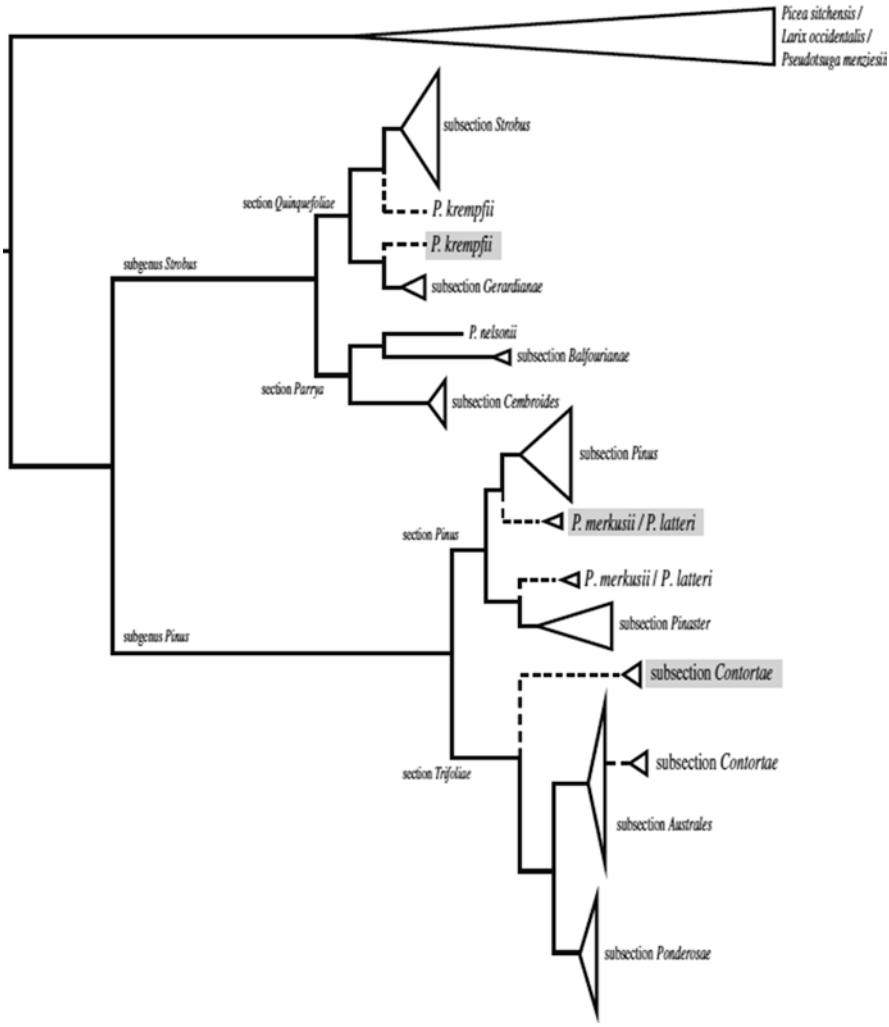


Fig. 16.6 Phylogenetic hypotheses for genus *Pinus*. Alternative placements (indicated by dashed lines) of subsections *Contortae* and *Krempfianae*, as well as the clade consisting of *Pinus merkusii* and *P. latteri* are shown. The most common plastid-based resolution of these groups is indicated by gray shading. (From Parks et al. (2012))

Subsequent studies followed with varied objectives, many seeking ways to refine infrageneric or even infraspecific relationships (Syring et al. 2005, 2007; Parks et al. 2009, 2012; Whittall et al. 2010; Hernandez-Leon et al. 2013; Willyard et al. 2017). One approach suggested was to introduce use of sequence from low-copy nuclear genes. Syring et al. (2005) used four such loci to further explore the evolutionary relationship among *Pinus* subsections. They argued a successful phylogenetic

strategy in *Pinus* would require many low-copy nuclear loci. A later study (Syring et al. 2007), using sequence from a single nDNA locus (LEA-like), uncovered widespread nonmonophyly in species of *Pinus* subgenus *Strobus* and concluded that the absence of allelic coalescence of nuclear loci leads to incomplete lineage sorting in the relatively recently diverged pine taxa, making nuclear loci questionable for phylogenetic studies in pine. Others have advocated for the use of full chloroplast genome sequences to improve resolution at lower levels (Parks et al. 2009, 2012; Whittall et al. 2010). Results of these studies suggest that mutational variability across the cp genome is large in pine and use of full-genome data would preclude errors that might derive from reliance on only a few loci, but steps must be taken to insure whole-genome data are reliable (Parks et al. 2012). The full plastome-based tree of Parks et al. (2012; Appendix 2) is now viewed as the most current for the genus *Pinus* even though the authors concluded that difficulty remains in resolving the phylogenetic relationships of historically recalcitrant nodes, including subsections *Contortae* and *Krempfianae*.

Several studies have looked at employing molecular clock methods to estimate times of divergence of *Pinus* clades (Eckert and Hall 2006; Willyard et al. 2007; Gernandt et al. 2008; Saladin et al. 2017). Estimates vary widely and are influenced greatly by assumptions regarding fossil evidence. Divergence of pine subsections likely occurred during the Miocene (within the last 20 my), with subsect. *Australes*, *Ponderosae*, and *Contortae* radiating within a 5 my time span starting around 18 mya (Willyard et al. 2007).

We conclude this brief section on the genus *Pinus* by drawing notice to issues that have posed significant problems for phylogenetic analyses of recently radiated taxa, all of which lead to discordance between gene genealogies and species phylogenetic trees (Liston et al. 2007; Syring et al. 2007; Willyard et al. 2009; Hernandez-Leon et al. 2013; DeGiorgio et al. 2014). These include introgression and hybridization, gene duplication, and incomplete lineage sorting. Hybridization, relatively common in pines and likely other genera of the Pinaceae (e.g., *Picea*, *Abies*, *Larix*), can lead to reticulated rather than linear evolution which is not well accommodated by current software programs. Use of multiple loci from nuclear and organellar DNA is recommended to detect and resolve some of these issues (Willyard et al. 2009).

Picea A phylogenetic study of *Picea* also recognizes the issues of hybridization, introgression, and incomplete lineage sorting in resolving interspecific relationships (Lockwood et al. 2013). They used plastid, mitochondrial, and nuclear sequences in a taxon-rich study that identified three major clades with strong geographical affinities: Asian, Eurasian, and North American (Fig. 16.7). Placement of *P. breweriana*, an endemic of the western United States and a notoriously difficult taxon to resolve, is listed as the sole North American member of the Asian clade. *Picea* is hypothesized to have originated in Asia and expanded into North America (two events) and Europe (multiple events). The most recent common ancestor of all extant spruce species is estimated to have lived ca. 28 mya. Interestingly, this study found *P. engelmannii* (North America), *P. abies* (Europe), and *P. brachytyla* (China) to be

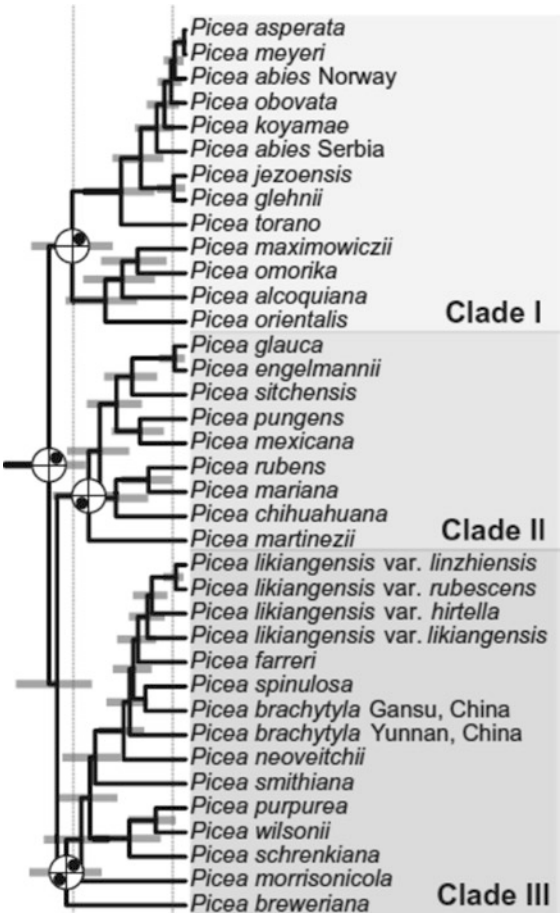


Fig. 16.7 Assignment of 33 of 38 recognized species of *Picea* (Farjon 2010) to three clades, showing strong geographical affinities: Clade I – Eurasia; Clade II – North America; Clade III – Asia. (Derived from Lockwood et al. (2013))

nonmonophyletic; the authors proffer hypotheses as to the nature and causes of these divergent taxa including morphological convergence, interspecific introgression, and incomplete lineage sorting. In the case of *P. abies* (Norway spruce), the northern and southern clades are estimated to have diverged around 6 mya, well before the ice age and its hypothesized environmental refugia, and longer than the split between many species. For now, the evidence is insufficient to provide more than conjecture as to the relationships between northern and southern European populations of this species.

Abies The species of this genus (true firs) are distributed widely across the Northern Hemisphere (25 species in northern and eastern Asia, 16 species in North America, and 7 species in the Mediterranean and southwestern Asia). Farjon and Rushforth (1989) subdivided the genus into ten sections, based on morphological traits, though Eckenwalder (2009) lists only seven sections based on subsequent, though incomplete, DNA-based phylogenetic studies. Recent papers by Semerikova

and Semerikov (2014) and Xiang et al. (2015) have challenged the previous taxonomic treatments. Xiang et al. (2015) conducted a comprehensive phylogenetic study of *Abies* using DNA sequence data of nuclear, chloroplast, and mitochondrial origin to screen ca. 42 species (Fig. 16.8). Their phylogenetic analyses recovered six robust clades largely consistent with some, though not all, earlier taxonomic sections, but results varied based on which genomic sequences were used. The section *Balsamea*, which includes species from both North America (*A. lasiocarpa*, *A. fraseri*, *A. balsamea*) and eastern Asia (*A. veitchii*, *A. sibirica*, *A. sachalinensis*, *A. koreana*, *A. nephrolepis*), was found to be monophyletic based on nuclear internal spacer sequences, but not by cpDNA sequences. Furthermore, the placement of section *Balsamea* was inconsistent between what is indicated by the paternally inherited chloroplast sequences and what is indicated by the maternally inherited mitochondrial sequences. The authors concluded the section *Balsamea* was largely influenced by ancient hybridization events and phylogenetic treatments reflect reticulation and possibly incomplete lineage sorting issues.

The unusual species endemic to the mid-Californian coastal mountains, *A. bracteata*, is treated as a separate section, sister to all other clades. The divergence time between *A. bracteata* and the other *Abies* species was estimated to be about 49 mya. Section *Amabilis* includes two species, *A. amabilis* from western North America and *A. mariesii* from eastern Asia. Xiang et al. (2015) conclude an out-of-North America migration (reverse Bering Land Bridge migration) for the origin of *A. mariesii*, the split between the species occurring some 11 mya.

The Eurasian clades likely separated from the other North American clades around 43 mya and separation of the two large North American clades, not including *A. bracteata*, was estimated to be 28 mya.

Podocarpaceae

The third largest of the conifer families, the Podocarpaceae is currently comprised of 18–19 genera and 157–178 species (Table 16.1). As with the Cupressaceae, many of the genera are species poor and range restricted suggesting a relictual status, though the fossil record indicates many genera were once more robust and widespread (Hill and Brodribb 1999). The current taxonomy of the family reflects in large part the revisions of de Laubenfels (1969, 1985, 1987), based exclusively on morphological traits. The main result of his work was to significantly expand the number of recognized genera, a view broadly supported by subsequent phylogenetic studies, both molecular and morphological (reviewed in Biffin et al. 2011). The widely held view today is that the genus *Phyllocladus* belongs in the Podocarpaceae, but Farjon and Filer (2013) retain it as a monogeneric family (Phyllocladaceae), a position mostly supported by morphological data.

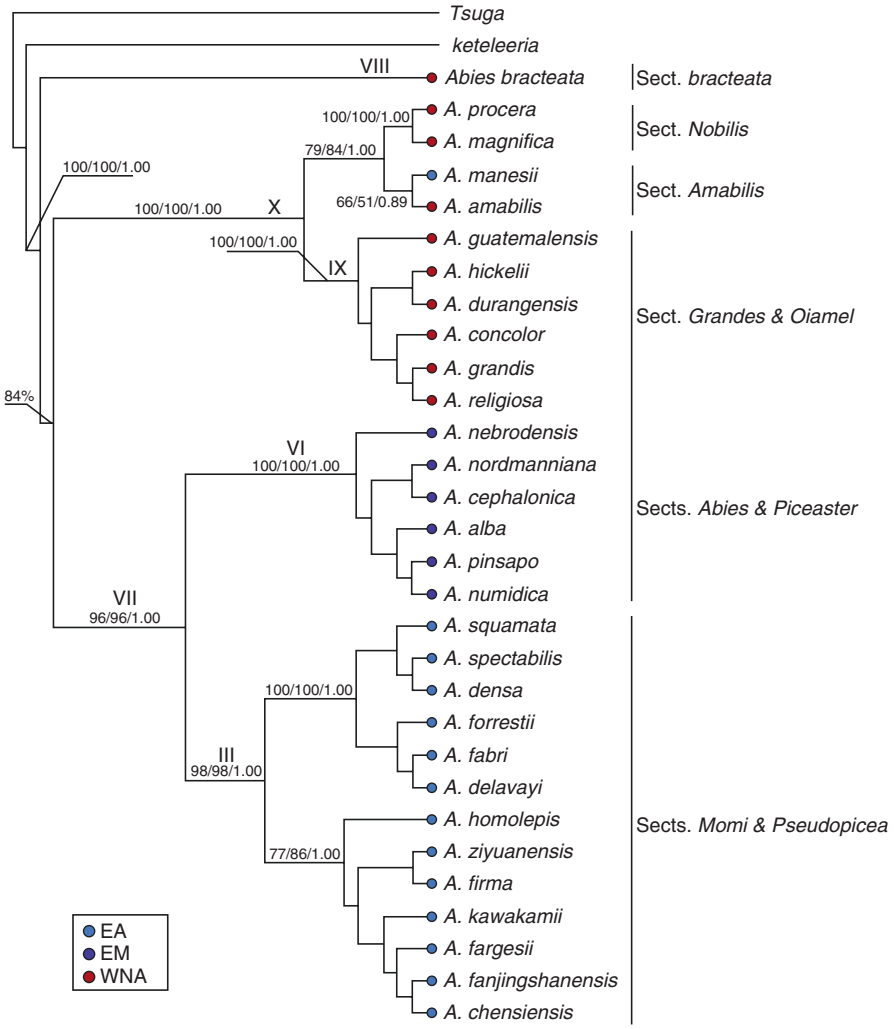


Fig. 16.8 The maximum likelihood cladogram for the genus *Abies*, based on combined data from three chloroplast, mitochondria, and nuclear sequences, with the exclusion of section *Balsamea*. Bootstrap values of MP, ML above 50%, and posterior support values of BI above 0.5 are shown. The species distributions are represented by the circles with different colors (EA eastern Asia, EM eastern Mediterranean, WNA western North America). The cladogram is rooted by the outgroup *Keteleeria.* (From Xiang et al. (2015))

A handful of phylogenetic studies, conducted between 1987 and 2011, have focused primarily on the taxonomic assignments of, and the relationships among, the family's genera (Kelch 1998; Conran et al. 2000; Sinclair et al. 2002; Biffin et al. 2011). Some studies are constrained by low-density taxon sampling, while virtually all molecular-based studies were conducted with few (1 to 3) DNA fragment sequences (cp loci *mat*K and trnL-trnF, nuclear ITS 2). The result is that support for generic relationships is relatively high but species comparisons are weak. The most recent, and complete, phylogenetic study (Biffin et al. 2011) recognizes three primary clades (Prumnopityoid, Dacrydioid, and Podocarpoid), places *Lepidothamnus* and *Plyllocladus* as sister to one another and together as sister to the three primary clades, and identifies *Podocarpus* as the most derived genus (Fig. 16.9).

Many of the family's genera are unique among the conifers, a few of which are briefly noted here. Known as the pigmy pines, two of the three *Lepidothamnus* species are diminutive, seldom reaching a meter in height, and the genus is recognized

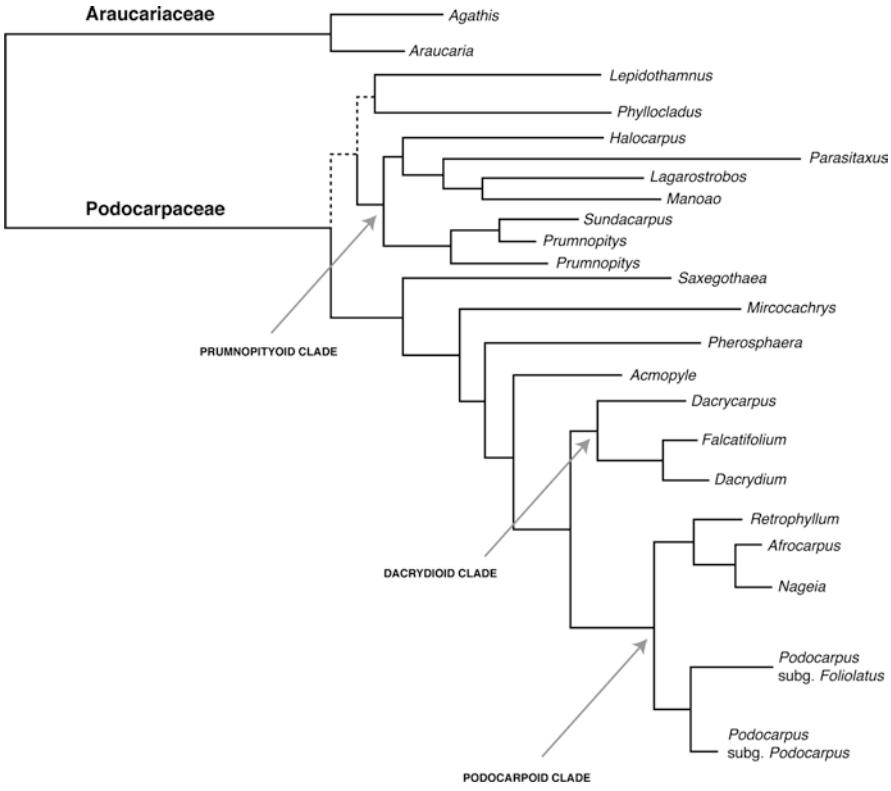


Fig. 16.9 Hypothetical phylogenetic relationships among genera and subgenera of the Podocarpaceae, including the outgroup genera *Agathis* and *Araucaria* from the Araucariaceae. Species placement has been removed to simplify the original figure (Figure modified from Biffin et al. (2011), Fig. 1.2). (Table from Biffin et al. (2011))

for having the smallest genomes among the conifers. The somewhat controversially placed *Phyllocladus* lacks true leaves, but sports photosynthetic, flattened twigs called phylloclades. The next most-basal clade, folded within the Prumnopityoid clade, includes *Parasitaxus*, the only parasitic conifer, and the genus *Retrophyllum*, distinguished by Farjon and Filer (2013) as possessing the only true conifer rheophyte (grows in running water) species (*R. minus*).

Most podocarp species presently occur within angiosperm-dominated, warm, and humid forests, principally in the Southern Hemisphere though *Podocarpus,* easily the largest genus (82–97 or more species) in the family and shown to be the most recently derived (Biffin et al. 2011), has an extensive worldwide distribution, occurring on all continents but Europe and Antarctica. The genus is subdivided into two subgenera, within which several sections are proposed (de Labenfels 1969, 1985). *Podocarpus* is notable, along with other genera in the family (e.g., *Nageia*, *Sundacarpus*), for possessing broad, flattened leaves (Fig. 16.10). A compelling argument has been forwarded (Biffin et al. 2011, 2012) that the flattened leaves, coupled with anatomical changes that facilitate radial transport of solutes from leaf midrib to leaf-margins, have allowed podocarps to compete with angiosperms in low-light tropical forests. These studies have shown, using phylogenetic and relaxed molecular clock analyses, that the radiation of taxa, particularly in *Podocarpus*, is consistent with the development of warm, humid environments and angiosperm predominance in the Southern Hemisphere (node ca. 60 mya; crown taxa within the last 5–20 mya).

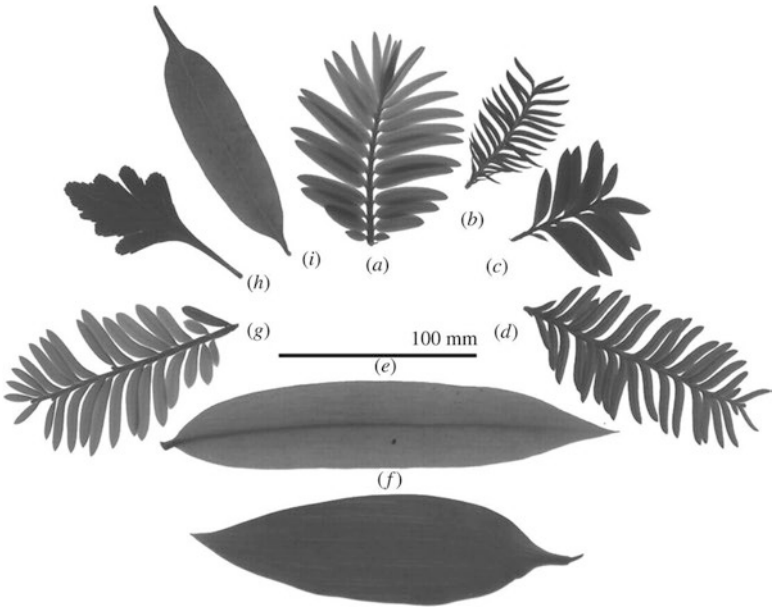


Fig. 16.10 Examples of the diversity of shoot flattening in the nine genera of the Podocarpaceae: (**a**) *Retrophyllum*, (**b**) *Dacrycarpus*, (**c**) *Falcatifolium*, (**d**) *Acmopyle*, (**e**) *Podocarpus*, (**f**) *Nageia*, (**g**) *Prumnopitys*, (**h**) *Phyllocladus*, and (**i**) *Sundacarpus.* (From Biffin et al (2012))

Sciadopityaceae

This family consists of a single genus and species (*Sciadopitys verticillata*), which, until recently, was included in the Taxodiaceae based mostly on the resemblance of its seed cones with those of *Sequoia*. Analyses of DNA, cone developmental, chromosomal, and paleobotanical character traits lent strong support to maintaining the taxon as a separate family (Farjon 2005; Eckenwalder 2009). The family is considered sister to both the Taxaceae and Cupressaceae; together the three families form the order Cupressales (Christenhusz et al. 2011).

Today, *S. verticillata*, or umbrella pine, is found solely in Japan, though it is commonly grown elsewhere as an ornamental. The family is found in the fossil record from the middle to late Cretaceous through the Tertiary (Fig. 16.2) in the Northern Hemisphere (Ohsawa et al. 1991; Stockey et al. as cited in Farjon 2005) suggesting the family was once widespread throughout Eurasia, represented by multiple species, but experienced a severe range contraction during the Pleistocene (Farjon and Filer 2013). A unique feature of umbrella pine is the arrangement (phyllotaxy) and nature of the "needles," of which there are two types (described in Eckenwalder 2009). One type, referred to as "double needles" appear as a pair of needles fused side to side, but technically termed cladodes, or shoots that function as leaves.

Taxaceae

Though small in comparison to other conifer families, the Taxaceae seems to present outsize issues for taxonomists and phylogeneticists. Until Pilger (1926) separated the conifers into seven families, including the Taxaceae, the group was generally considered as a separate order, distinct from other conifers. This treatment lasted well into the twentieth century by some notable taxonomists (Florin 1951). There is little disagreement today regarding the family's position in the Coniferae. In the strict sense, the Taxaceae consists of five genera: *Amentotaxus, Torreya, Austrotaxus, Pseudotaxus,* and *Taxus* with a total of 18–24 species, depending on authority (Table 16.1), though a broader definition, increasingly accepted (Eckenwalder 2009; Christenhusz et al. 2011), includes the genus *Cephalotaxus* (5–8 species) in the family. Farjon and Filer (2013) still recognize the latter genus as belonging to a separate family, the Cephalotaxaceae, a position supported in part by molecular phylogenetic studies (Cheng et al. 2000; Hao et al. 2008b). However, the results of the latter studies, with deep taxon sampling, are dependent upon the array of molecular sequences used. When several chloroplast loci are used, the two taxa appear to be monophyletic within the same clade, but the inclusion of a single nuclear locus appears to separate them (see Fig. 3, p. 99 in Hao et al. 2008b). The most recent study based on analyses of 28 nonmolecular traits clearly places all six genera in a single, monophyletic clade, and recognizes the sole family, Taxaceae (Ghimire and Heo 2014). Clearly this

issue remains unresolved and likely will require deeper sampling of nuclear and mitochondrial sequences for clarification. Phylogenetic treatments for interspecific relationships within the genera *Cephalotaxus* and *Taxus* provide insights into still other problematic taxonomic treatments (Hao et al. 2008a, b) and raise questions about the authors' own phylogenetic hypotheses that include relatively distantly related taxa (Podocarps) as outgroups and an unusually large number of recognized species of *Taxus* (14). With the exceptions of *Austrotaxus spicate* and *Taxus sumatrana*, the family is distributed in the Northern Hemisphere. The oldest representative fossils have been dated to ca. 200 mya.

Summary

Chapter 16 addresses conifer diversity at the taxonomic levels of species, genus, and family, from the evolutionary beginnings of the conifer lineage more than 300 mya to the contemporary hypotheses of their phylogenetic relationships. Our understanding of early conifer evolution derives from a robust but fragmented fossil record. Conifers likely descended from progymnosperms that resembled, in many respects, modern conifers but reproduced by spores. Conifers flourished throughout the Mesozoic Era (250–65 mya), dominating the world's floristic elements with a rich and diverse group of taxa. Extant conifer families likely began to appear about 230–200 mya.

Taxonomic treatments of extinct and extant conifers continue to evolve as new evidence is obtained. Depending on authority, modern conifers may be classified as belonging to 1–3 orders, 6–8 families, 67–71 genera, and 546–670 species. The 5 largest genera (*Pinus*, *Podocarpus*, *Juniperus*, *Abies*, and *Picea*) include ca. 55% of all conifer species, while 29 genera are monotypic. Taxonomic revisions of *Pinus* and *Cupressus*, informed by phylogenetic studies, have received considerable attention in recent decades.

Phylogenetics is the study of the evolutionary relatedness, or genealogical relationships, among groups of organisms. Phylogenetic methods have evolved rapidly with the introduction of molecular markers and modern DNA sequencing technology. Phylogenetic inferences have benefitted from having markers and sequences derived from all three sources of DNA (chloroplast, mitochondria, nuclear). Multiple molecular phylogenetic investigations have been conducted for all conifer families and many of the larger genera. Phylogenetic trees, or cladograms, based on relatively recent studies, are presented for most conifer families and three genera in the Pinaceae.

Comparative Genomics

17

Introduction to Comparative Genomics

"Nothing in biology makes sense except in the light of evolution." Theodosius Dobzhansky

This well-known quote from Dobzhansky, made long before the development of genomic sciences, is the overriding theme of this chapter. We have tried to summarize the study of conifer genetics and genomics in three sections of this volume; Genomes, Variation, and Evolution. In this chapter, we attempt to bring all these sections together to develop a deeper understanding of genomes and variation in the context of the evolution of species of Coniferales. The earliest form of comparative genomics in conifers was the work in comparing chromosome number, genome size, and karyotypes across taxa (Chap. 2). We saw that the 1N chromosome number in conifers varies little; from 11 to 13 with just a few exceptions (Table 2.1). Polyploidy is extremely rare, with just the tetraploids *Fitzroya cupressoides* (Alerce) and *Juniperus chinensis* "Pfitzeriana" and the hexaploid *Sequoia sempervirens* (coast redwood). Genome size in conifers, however, varies over nearly an order of magnitude with the smallest genome being 4067 Mb (*Microcachrys tetragona*) and the largest being 35,084 Mb (*Pinus gerardiana*) (Table 2.1). The variation in genome size can be accounted for by differences in noncoding DNA (Chap. 4); the number of protein coding loci appears to be quite similar among species (Chap. 3). Karyotype analysis using various chromosome banding techniques showed great similarity among chromosomes not only across species but even among chromosomes within species. It was not until FISH techniques were developed and used that karyotype differences among chromosomes were observed and homoeologous chromosomes among species could be determined (Chap. 2). The *classical era* of genomics (pre-2000) established the conservative aspect of conifer chromosome evolution and that the large phenotypic differences among conifers would be due to the allelic differences among species for a similar set of protein-coding loci (Chap. 10) and differences in the expression of these alleles at these loci (Chap. 5). Add to these differences epistatic variation, genotype × environment interaction, and likely epigenetic factors and it becomes easy to account for the large amount of variation in form and function among conifers.

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_17

Comparative Mapping

Genetic mapping in conifers began in the 1970s with allozyme markers (Conkle 1981), but because there were so few markers (< 50), it was not possible to construct genetic maps that spanned entire genomes. Once DNA-based markers (RFLP, RAPD, AFLP, SSR, ESTP, SNP; Chaps. 9 and 10) became available, it was possible to produce complete maps. The first maps were constructed from 50 to 100 markers, whereas maps can now be made from 10s of thousands of SNP markers. The status of genetic maps through 2007 in species of genus *Pinus*, *Picea*, and *Pseudotsuga* and *Cryptomeria japonica* was summarized by Kole (2007). In Chap. 11, we summarized the use of genetic maps to discover QTLs (Table 11.1). There we often made reference to QTL validation within and across species. The cross-species validation was possible due to comparative maps between species; how these maps were made is the focus of this section of this chapter.

Comparative maps between two or more species are constructed based on comparing the map position of orthologous genetic markers. Distinguishing between orthologous and paralogous genetic loci is the necessary condition before comparative maps can be made. Conkle (1981) identified orthologous allozyme loci to show conserved linked loci in related species of *Pinus*, but as noted earlier there were not enough allozyme loci to construct comparative maps. The first DNA-based markers to be used in conifers were RFLPs and it was shown that RFLP markers developed from *Pinus taeda* cDNA probes would hybridize to DNA from other conifers and might be useful for comparative mapping (Ahuja et al. 1994). The first, and ultimately only, study to apply this principle for constructing a comparative map between two species was for *P. radiata* and *P. taeda* (Devey et al. 1999) (Table 17.1; Fig. 17.1). RFLP markers in conifers were technically demanding to use and RFLP probes generally revealed multiple genetic loci so that distinguishing orthologous from paralogous loci was difficult. RAPD and AFLP markers came into broad usage at this time but repeatability issues plagued the application of these markers and distinguishing orthologues from paralogues was nearly impossible. Conifer geneticists thus sought a marker type that could be used in comparative mapping across

Table 17.1 Comparative mapping studies in conifers

Species used for comparative mapping	Reference
Pinus radiata, *P. taeda*	Devey et al. (1999)
Pinus elliottii, *P. taeda*	Brown et al. (2001)
Pinus pinaster, *P. taeda*	Chagné et al. (2003)
Pinus sylvestris, *P. taeda*	Komulainen et al. (2003)
Pseudotsuga menziesii, *Pinus taeda*	Krutovsky et al. (2004)
Picea glauca, *P. mariana*, *P. rubens*, *P. abies*	Pelgas et al. (2006)
Pinus lambertiana, *P. taeda*	Jermstad et al. (2011)
Pinus pinaster, *P. taeda*	Chancerel et al. (2011)
Picea consensus, *Pinus taeda*	Pavy et al. (2012b)

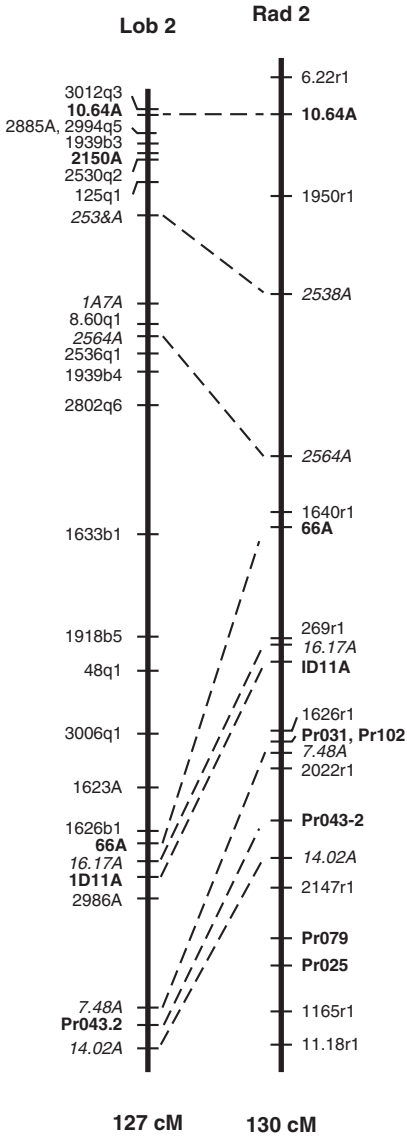


Fig. 17.1 Comparative map for *Pinus radiata* and *P. taeda* at one linkage group. Orthologous RFLP and SSR markers are in bold. (From Devey et al. (1999))

species. The genetic marker that met this requirement was the expressed sequenced tag polymorphism (ESTP) marker (Harry et al. 1998; Temesgen et al. 2001). These markers were developed from ESTs in *P. taeda* (Chap. 3) and polymorphism revealed by denaturing gradient gel electrophoresis (DGGE). This genetic marker type was somewhat technically demanding to use; however, abundant orthologous markers were developed to construct comparative maps within species of *Pinus* and even across species of the Pinaceae (Brown et al. 2001) (Fig. 17.2).

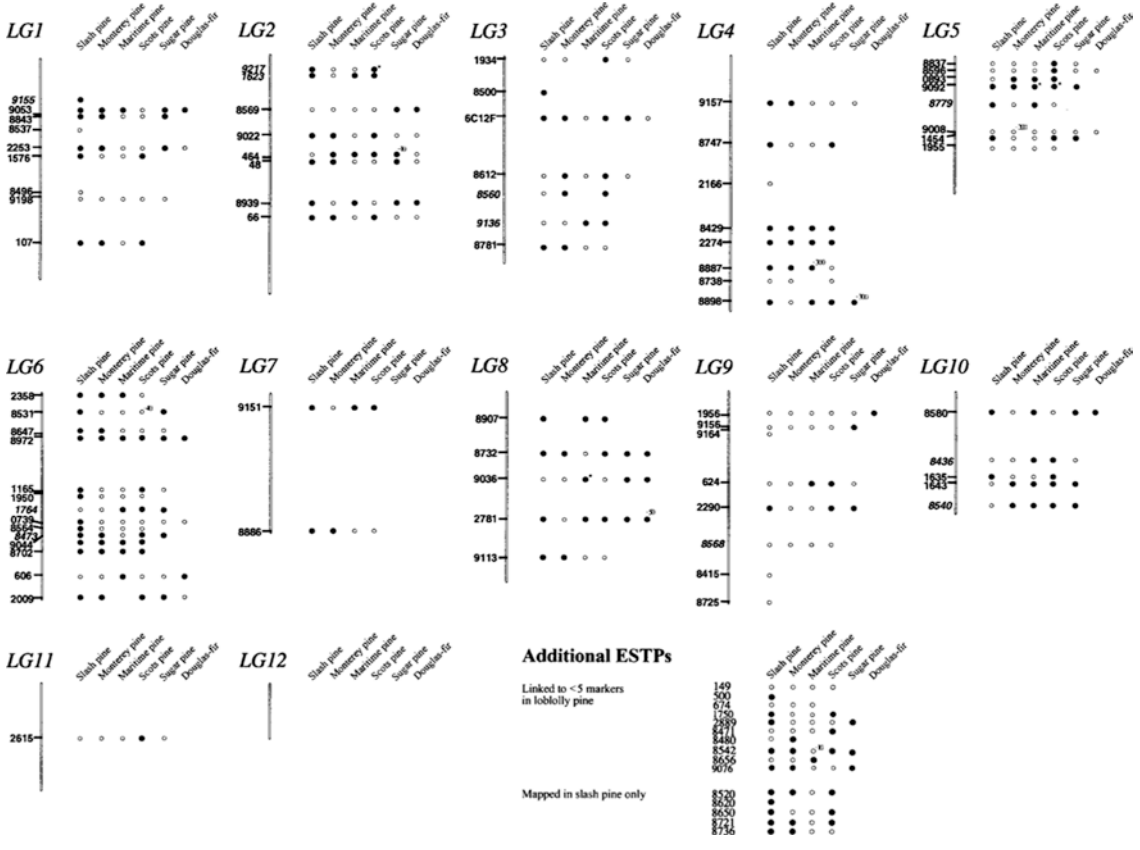


Fig. 17.2 Representation of the potential utility of *Pinus taeda* ESTPs as anchored reference loci. The 12 linkage groups of *P. taeda* are shown (linkage groups (LGs) 1 through 12). Successful PCR amplification in a species is shown by a circle; solid circles denote amplification products that are polymorphic in the pedigree examined; and open circles denote monomorphic products. (From Brown et al. (2001))

The ESTP markers developed from *P. taeda* were then used to develop comparative maps with three pine species, *P. elliottii* (Brown et al. 2001), *P. pinaster* (Chagné et al. 2003), and *P. sylvestris* (Komulainen et al. 2003) and *Pseudotsuga menziesii* (Krutovsky et al. 2004). For example, Fig. 17.3 shows the comparative map at linkage group 6 with all three pine species and *P. menziesii*. The important result obtained from these first comparative genetic maps in conifers is the high degree of macro-synteny among species. This implies that large chromosomal rearrangements such as inversions and translocations have not operated to a large extent over evolutionary time. The extent of micro-synteny will need to await comparative genome sequence analysis.

Comparative mapping in *Picea* has also received significant attention. A consensus map among four *Picea* species (*P. glauca, P. mariana, P. rubens,* and *P. abies*) was first constructed using 805 AFLP, SSR, and ESTP markers (Pelgas et al. 2006). These authors then went on to construct a comparative map with a consensus map from *Pinus* and with *Pseudotsuga menziesii* (Fig. 17.4). This map again shows the high degree of synteny within the Pinaceae, this time with three genera.

By about the year 2010, conifer geneticists began constructing genetic maps from SNP markers versus DNA fragment markers and many species-specific maps were constructed. Even though SNPs were massively abundant in many species for mapping, SNPs are almost entirely species-specific, so they are not orthologous. This limitation was overcome for the construction of a comparative map between *Pinus lambertiana* and *P. taeda* by using the gene annotations from the sequences within which SNPs were found for the two species (Jermstad et al. 2011) to identify orthologous markers. Sixty orthologous loci were identified in this way, and 56 markers were co-linear between species. This comparative map was constructed before the genomes of these two species had been sequenced and resequenced, therefore it should now be possible to produce a very high-density comparative map between the two pines, and also for *Picea abies* and *Pseudotsuga menziesii*. The same approach was also taken to construct a comparative map between *Pinus pinaster* and *P. taeda* (Chancerel et al. 2011).

A very recent comparative map study in conifers aimed to understand the timing of gene duplications in seed plants (Pavy et al. 2012b). In an elegant analysis using orthologous genes from angiosperms and conifers, the authors determined that most gene duplications predated the angiosperm-gymnosperm split and subsequent chromosome rearrangement in the conifers has been very slow. This inference was again well supported by the high degree of synteny between *Picea* and *Pinus* (Fig. 17.5).

In summary, comparative mapping, like earlier karyotype analysis, shows the slow and conservative nature of chromosome evolution in conifers. This situation has made it possible to use comparative maps to facilitate and validate gene and QTL discovery across species; treating conifer as a single genetic system. In the next two sections of this chapter we will begin to explore how comparative transcriptomic and comparative genome sequence analysis informs relationships among conifers at the micro-syntenic level.

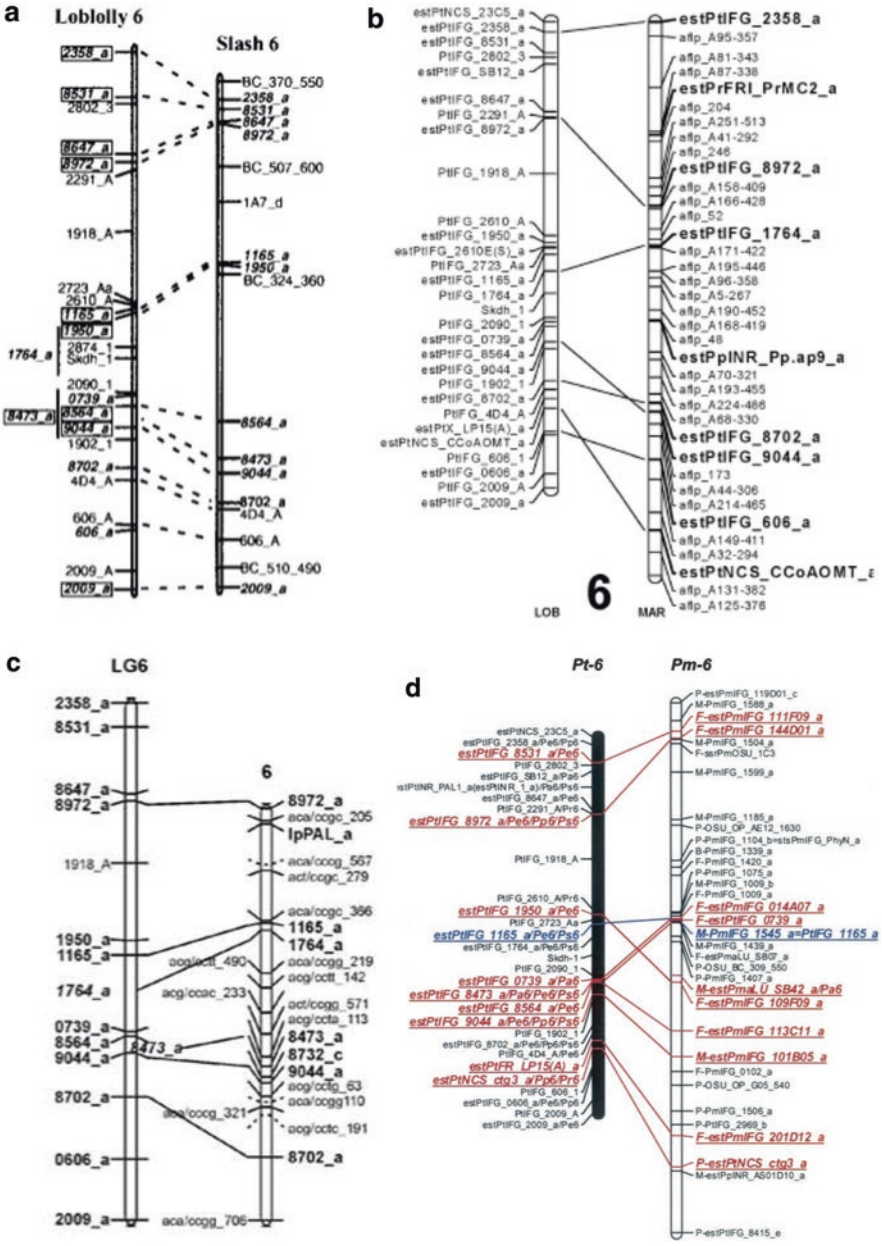


Fig. 17.3 Comparative maps at linkage group 6 of *Pinus taeda* with *P. elliottii*, *P. pinaster*, *P. sylvestris*, and *Pseudotsuga menziesii*. Orthologous genetic markers for linking maps are indicated in bold. (**a**) *Pinus taeda* × *Pinus elliotii*, (**b**) *Pinus taeda* × *Pinus pinaster*, (**c**) *Pinus taeda* × *Pinus sylvestris*, (**d**) *Pinus taeda* × *Pseudotsuga menziesii* (From Brown et al. (2001); Chagné et al. (2003); Komulainen et al. (2003); Krutovsky et al. (2004))

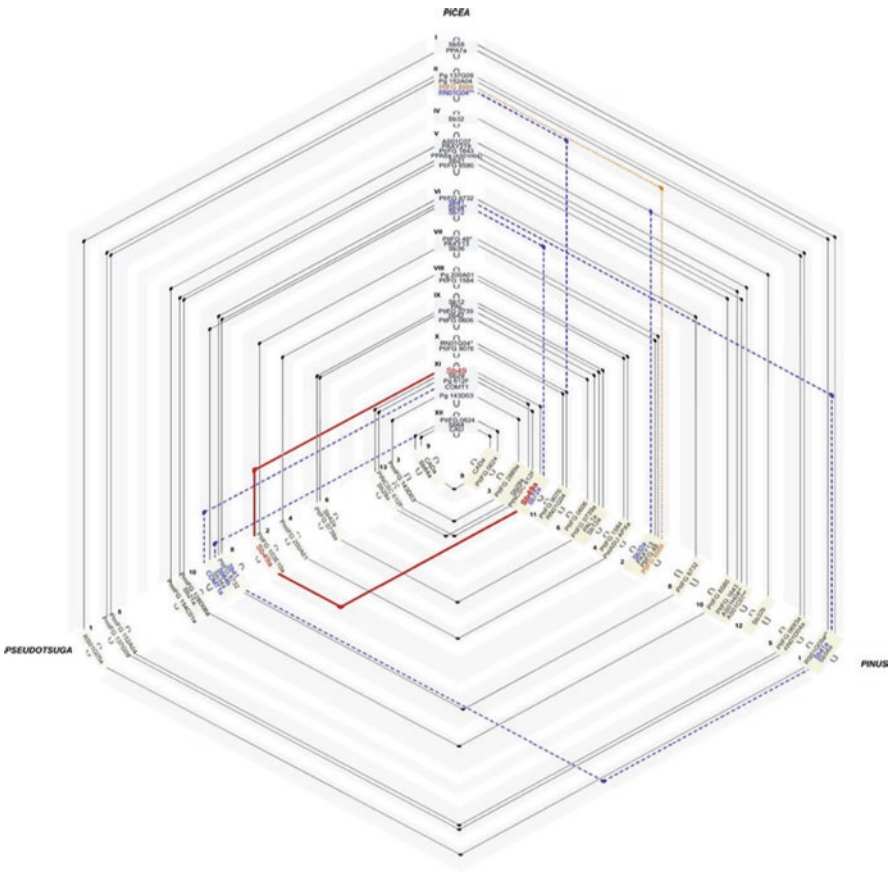


Fig. 17.4 Schematic representation of homoeologous linkage groups among three genera of the Pinaceae: *Pseudotsuga* (represented by *P. menziesii*; background in green on the left), *Picea* (background in blue in the middle), and *Pinus* (background in yellow on the right). Only homologous markers are indicated on each schematic LG. Orthologous markers are connected by a solid black line, except when they are positioned onto nonhomoeologous LGs (red line). Paralogous markers are connected by a dashed blue line, except for the anchor marker PtIFG 8569 (orange dotted line). (From Pelgas et al. (2006))

Comparative Gene Content and Transcriptomics

In Chap. 2 we described the approaches used to discover the expressed gene content in conifers. The early high-throughput approach used *expressed sequenced tags* (ESTs). This approach yielded a large number of ESTs from a small number of species and a limited number of tissue types in each species (Table 3.1). These early studies provided an estimate of the number of expressed protein-coding genes at 50,000 to 100,000. More recently, RNA-seq has been applied to a small number of conifers and a limited number of tissue types (Table 3.2). Estimates of expressed gene number from these studies were a bit lower, averaging around 50,000, which

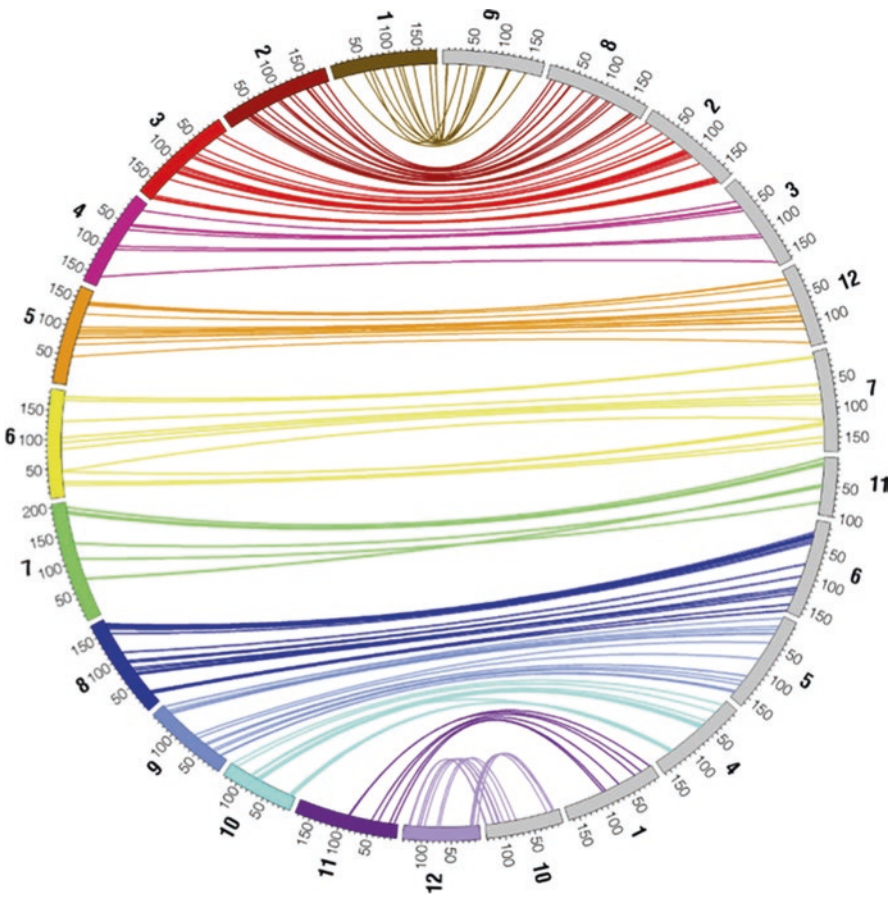


Fig. 17.5 A *Picea*/*Pinus taeda* comparative map. The syntenic positions of the 161 homologous genes mapped on both *Picea* and *P. taeda* genomes were plotted are indicated by color-coded lines connecting the *Picea* (in color) and the *P. taeda* chromosomes (in gray). The chromosome numbers are indicated with bold numbers outside the circle. (From Pavy et al. (2012b))

could be attributed to better bioinformatic analyses to identify redundant transcripts from the same gene. Finally, full-genome sequencing and annotation in five conifer species yielded estimates of expressed gene number from 9000 to 105,000. Collectively, these results make it quite clear that the exact expressed gene number among conifers is still unknown. Until such time that the true number of expressed genes becomes known, it will be impossible to accurately assign orthologous genes across species. Once this is accomplished, then it will become possible to conduct comparative genomic analyses using expressed gene sequence data.

Comparative transcriptome analysis will also soon become possible in conifers. Three different approaches to comparative transcriptome analysis are shown in Fig. 17.6. In Chap. 6 we summarized the extensive literature of conifer

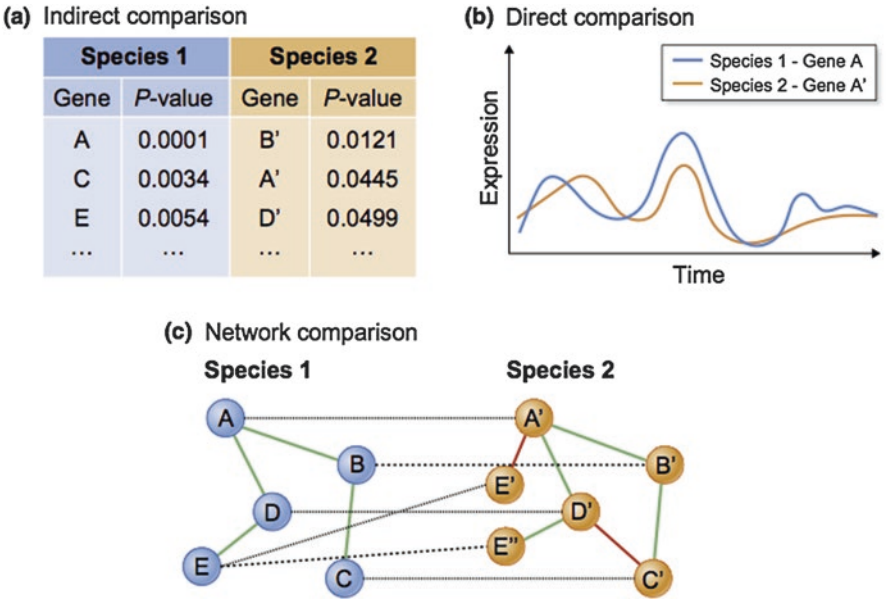


Fig. 17.6 Methods for comparative regulomics. (**a**) Indirect comparison methods analyze expression data in each species separately and compare the results afterward. This is the most common approach to comparative transcriptomics and relies on comparing lists of differentially expressed genes, gene co-expression clusters, or co-expression of gene families. (**b**) Direct comparison methods correlate expression profiles of orthologues directly, and are in many ways the most straightforward method. However, this method requires expression measurements from directly comparable samples across species. (**c**) Network comparison methods rely on aligning co-expression networks across species. This is a flexible method that allows comparisons at different levels from individual co-expression links, via network neighborhood to network modules, as well as analysis of gene properties in a network context including gene centrality and membership in network motifs such as feed-forward loops. (From Ingvarsson et al. (2016))

transcriptomic studies (Table 6.1), but just like EST and RNA-seq data, there are not yet enough studies done with multiple species with the same tissue types, developmental stages, and growing conditions to enable meaningful comparative studies.

Two recent studies reflect on how comparative transcriptomic studies might be conducted in conifers. RNA-seq data were generated from needle tissue from four *Pinus* subgenus *Strobus* species (*P. monticola*, *P. flexilis*, *P. albicaulis*, and *P. lambertiana*) (Baker et al. 2018). Transcriptomes were assembled and annotated for each of the species separately. Then, using specialized bioinformatics tools, orthologous gene families were identified. There were 2025 gene families common to all four species and a lesser number unique to one, two, or three species (Fig. 17.7). It was also shown that the vast majority of these genes families are under purifying selection based on dN/dS ratios (Fig. 17.8, Table 17.2). This comparative transcriptomic analysis sheds light on the genes and pathways responsible for function common to all species.

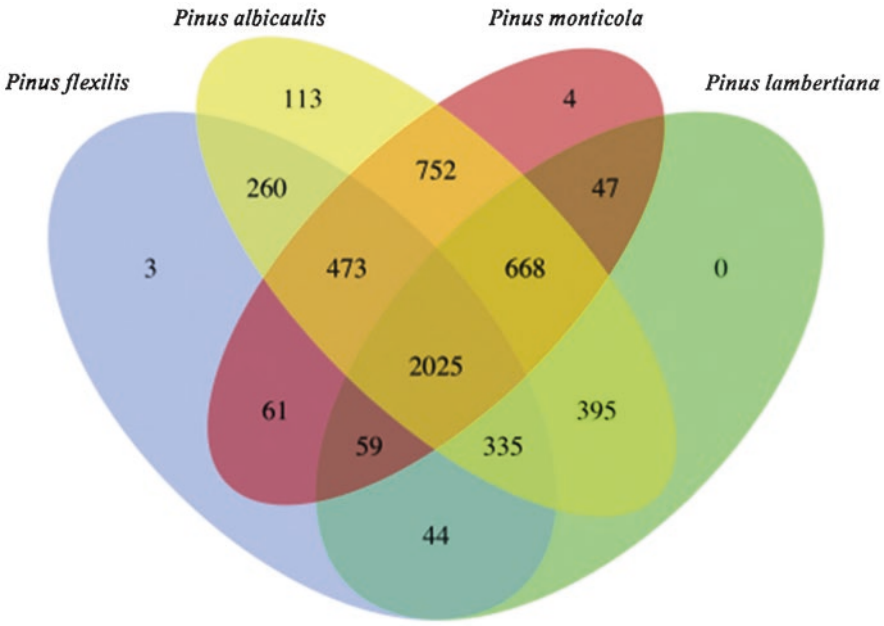


Fig. 17.7 Venn diagram showing the number of orthologous gene families among four species of *Pinus* subgenus *Strobus*, the white pines: *Pinus flexilis*, *P. albicaulis*, *P. monticola*, and *P. lambertiana*. (Modified from Baker et al. 2018)

Tree species co-occurrence was studied in response to drought in a forest population in northern Wisconsin, USA. Twenty-one tree species were found in the population, including five conifers (*Picea glauca*, *Larix laricina*, *Abies balsamea*, *Tsuga canadensis*, and *Thuja occidentalis*) (Swenson et al. 2017). Several functional traits and phylogenetic relatedness were assigned to all species. In addition, transcriptomes were generated for all species using RNA-seq. Results showed that gene expression was a better predictor of co-occurrence than either functional trait or phylogenetic relatedness. This is a nice example of how comparative transcriptomic analyses can inform the ecological distribution of species in forest populations.

Comparative Genome Sequences

The ultimate approach to comparative genome analysis in an evolutionary context will be aligning and comparing complete genome sequences. Differences in genome content and organization might then be correlated with differences in form and function among related species. Such analyses might inform when and how many times a character may have evolved within a related group. Successful comparative genome sequence analysis depends on having multiple and highly contiguous genome sequences. These resources are not yet available in conifers (De La Torre et al. 2014b). In Chap. 3, we listed the five conifer genome sequences

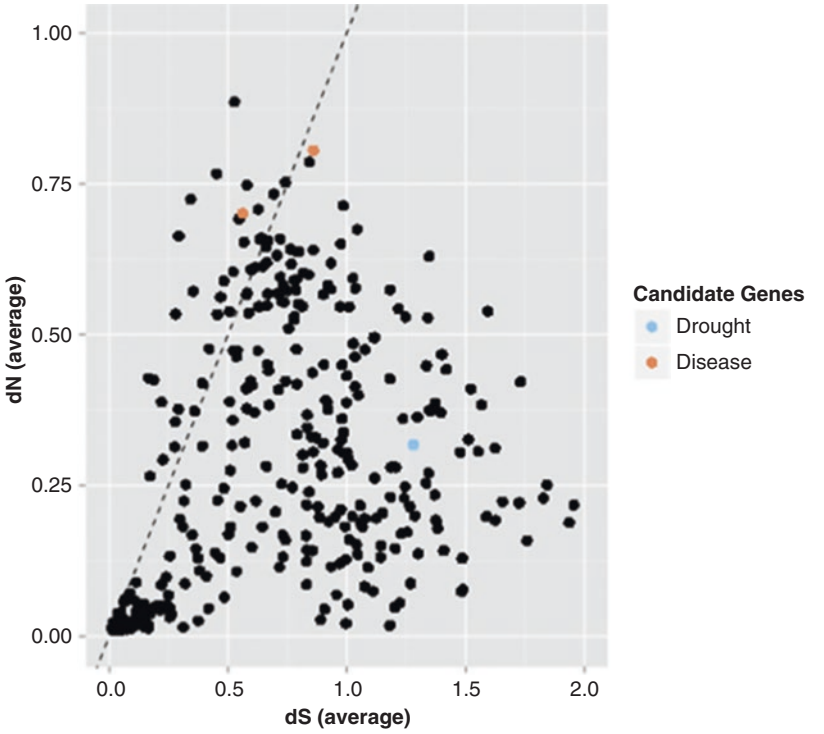


Fig. 17.8 Distribution of averaged values for dN/dS across four species of *Pinus* subgenus *Strobus* for each of the 408 gene families. Candidate genes with previous associations to drought tolerance/aridity and rust resistance are highlighted. (From Baker et al. (2018))

completed as of 2018 and the highly fractured nature of these sequences (Table 3.3). For example, the most contiguous genome sequence is that for *Pinus taeda* which still has 1.5M scaffolds.

Complete genome sequence alignments and comparisons have recently been done for a number of organisms with much smaller genomes with high contiguity. Genome sequences from 12 *Drosophila* species were aligned and compared and relationships with the amount of coding, noncoding, and *cis*-regulatory sequence were shown (Clark et al. 2007). For 10 species of bees, a relationship between genome complexity and eusociality was shown (Kapheim et al. 2015) (Fig. 17.9). Finally, 48 species of birds with complete genome sequences were compared with differences in lifestyle adaptations (Zhang et al. 2014b). These pioneering studies with small and highly contiguous genome sequences are examples of what will ultimately be possible in conifers. Here, comparative genome sequences can be associated with traits of fundamental interest and importance in conifers such as shade tolerance, reproductive cycles, cones versus fleshy fruits such arils, evergreen versus deciduous, mycorrhizal associations, crown form, grass stage, longevity, and serotiny. Comparative genome sequence analysis in conifers will begin to help elucidate when in evolutionary history and how these traits evolved.

Table 17.2 Summary of conserved gene families under positive selection from four white pine species

Gene family annotation	Alignment length (bp)	Gene ontology (molecular function)
Formyltetrahydrofolate deformylase mitochondrial isoform ×1	1053	Formyltetrahydrofolate deformylase activity; amino acid binding; hydroxymethyl-, formyl-, and related transferase activity
F-box kelch-repeat protein skip6-like	1122	Protein degradation tagging activity
Low-quality protein: nitrate reductase	2778	Oxidoreductase activity; metal ion binding; organic cyclic compound binding; heterocyclic compound binding
Arogenate dehydrogenase chloroplastic	1278	Prephenate dehydrogenase activity
Carrier protein chloroplastic	2283	ATP:ADP antiporter activity; ATP binding
Flowering time control protein fpa	3279	
Transcription initiation factor tfiid subunit partial	1818	
e3 ubiquitin-protein ligase keg isoform ×2	4917	Protein degradation tagging activity
Predicted: uncharacterized protein LOC18435046	1524	
Two-component response regulator-like prr37	2856	
Isoamylase chloroplastic	2769	
Probable u3 small nucleolar RNA-associated protein 7	1623	18S ribosomal RNA processing
Predicted: uncharacterized protein LOC103493568	1407	Metal ion binding; sequence-specific DNA binding transcription factor activity
Calcium-transporting ATPase plasma membrane-type-like isoform ×1	3192	Calcium-transporting ATPase activity; calmodulin binding; ATP binding; metal ion binding
Protein notum homolog	1263	
Predicted: uncharacterized protein LOC104607701	1380	
Clathrin assembly protein at5g35200	1644	1-phosphatidylinositol binding; clathrin binding
Predicted: kanadaptin	2274	
Family 18 glycoside hydrolase	1236	Chitinase activity; chitin binding
Dead-box ATP-dependent RNA helicase 13	1176	
Probable inactive purple acid phosphatase 27	1977	Acid phosphatase activity; metal ion binding; dephosphorylation
Arginine decarboxylase	2283	Carboxy-lyase activity
Predicted: uncharacterized protein LOC104591536	1626	
Erythronate-4-phosphate dehydrogenase-like protein	975	

(continued)

Table 17.2 (continued)

Gene family annotation	Alignment length (bp)	Gene ontology (molecular function)
Interferon-induced guanylate-binding protein 2-like	3207	GTPase activity; GTP binding
Nf-×1-type zinc finger protein nfxl1	4290	Metal ion binding
Transmembrane protein 87b-like	1560	
Fructokinase-like chloroplastic	1644	Kinase activity; phosphotransferase activity, alcohol group as acceptor
Bel1-like homeodomain protein 1	2517	DNA binding
cbs domain-containing protein cbsx6	1311	
duf21 domain-containing protein at4g14240	1623	
Probable wrky transcription factor 14	1431	
Myeloid leukemia factor 1-like isoform ×2	1059	
Mannose-1-phosphate guanylyltransferase 1	948	
Phytoene synthase chloroplastic	1308	Geranylgeranyl-diphosphate geranylgeranyltransferase activity; phytoene synthase activity
Unknown	336	
Predicted: myosin-10-like	1803	
Universal stress protein a-like protein	366	
Predicted: uncharacterized protein LOC104602728	966	

From Baker et al. (2018)

Summary

The power of comparative genomic analyses to understand the evolution and genetic basis of form and function is well established throughout biology. Likewise, comparative genomic analyses in conifers will have much to contribute toward understanding biological processes fundamentally important or unique to conifers. Comparative genomic analyses began with determining chromosome number, genome size, and karyotypes. This work established the conservative nature of chromosome evolution in conifers. Comparative mapping using orthologous genetic markers established the high degree of synteny in conifers. This was expected based on the earlier karyotype analysis. In the modern genomic era, comparative transcriptomic and comparative genome sequence studies done with multiple and related species will help to reveal the genetic basis of traits important to conifers and when and how often these events have occurred in evolutionary time.

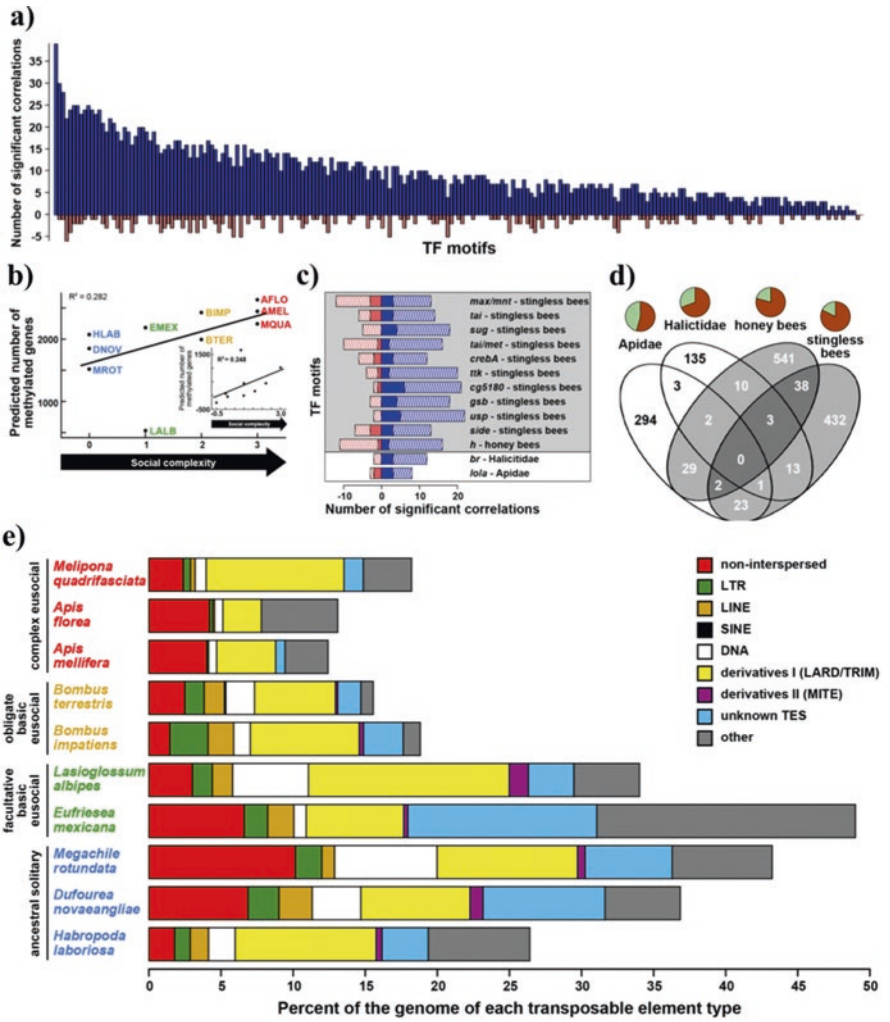


Fig. 17.9 Genomic signatures of evolutionary transitions from solitary to group life in bees. (**a**) Increasing social complexity is associated with increasing presence of *cis*-regulatory TFBS in promoter regions. Each bar represents a TFBS for which presence correlates significantly with social complexity (blue: positive; red: negative). (**b**) Relationship between predicted number of methylated genes and social complexity before and after (inset) phylogenetic correction (**c**) TFBS motifs showing a relationship between social complexity and evolutionary rate of coding and non-coding sequences in different lineages. Bar length indicates the number of significant correlations (blue: positive; red: negative) between each motif score and social complexity among genes evolving faster (solid) or slower (hatched) in lineages with different levels of social complexity. (**d**) Number of genes for which evolutionary rate is faster or slower in lineages with higher compared to lower social complexity. Pie charts represent the proportion of genes evolving slower (light green) or faster (dark orange) with increased social complexity. (**e**) Complex eusocial species have a reduced proportion of repetitive DNA compared to other bees. (From Kapheim et al. (2015))

18 Historical Perspective and Future Directions in Forest Genetics and Genomics

Historical Perspective

Before speaking to the future of conifer genetics and genomics research, a brief overview of the history of the discipline may be valuable for placing the discussion in perspective. While previous chapters in this book review the scientific approaches and findings of studies seeking to characterize genetic variation in conifers, a concomitant look at the economic, social, and policy factors that guided this research may inform how the discipline will proceed in the future. Much of the historical discussion follows from a review piece we coauthored on the evolution of our field of study in the United States (Wheeler et al. 2015). Parallel stories, though certainly not exacting in detail, can be told for other regions of the world.

Forest genetics research in the United States started soon after the rediscovery of Mendel's studies and for the next 40 years was dominated by species, racial, and provenance trials of economically important conifer species and artificial hybridization studies. This work was conducted largely by employees of the US Forest Service. Along with applied tree breeding and testing, forest genetic research enjoyed rapid and expansive growth for the next 30 years (1950s to 1980s). Abundant support for research came from Federal, State, and corporate programs, and university/industrial cooperatives became important centers of expertise. By the early 1980s in the United States alone, 65 private companies, 22 state agencies, all Forest Service regions and Research Stations, and most university forestry programs employed someone engaged in the study of forest genetics and/or tree improvement. Graduate student programs flourished. The development of new technologies, such as allozyme analyses, tissue culture, recombinant DNA technology, genetic transformation, and later, genomic technologies, drew new resources and people to a highly regarded field of study. Importantly, competitive grant programs emerged for forest genetic research, primarily funding single-scientist projects.

The subsequent 30-year period (mid-1980s to present) witnessed an equally dramatic contraction and transformation of the forest genetics and tree improvement communities resulting from the confluence of several economic, social, and

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5_18

policy factors. The forest industrial landscape was dramatically altered by corporate mergers, acquisitions, and the move to convert land holdings to real estate investment trusts (REITs) and timber investment management organizations (TIMOs). The result was a discontinuation of most in-house genetic and tree improvement programs. A policy change in the US Forest Service shifted emphasis from forest genetic and tree improvement research to an ecological/conservation focus. Coupled with tightening budgets, this has led a significant reduction in forest genetics expertise in the agency. With the development of modern genomic technologies, competitive grant programs shifted toward large, multi-institutional and disciplinary projects, the result of which was to infuse significant resources into forest genetics research, but to fewer, genomics-specialized groups at a smaller number of institutions. Many universities lost their forest genetics expertise to retirement or shifts to administration as funding for small-scale or traditional genetics research projects dried up. Ironically, during the latter phases of this erosive transition for the community, the genomics era has developed the technical capacity to produce reference genome sequences in a rapid, cost-effective manner. Already reference genome sequences exist for a handful of species and the results are being used broadly throughout the existing community. It is not unreasonable to assume this gold standard of genetic resources could be developed for virtually every conifer species within a decade.

Current Situation

Today, the conifer genetics community is confronted by new and evolving challenges, the most notable of which are climate change and increasingly frequent introductions of alien pests and diseases. Historically, conifer genetic research has concentrated on relatively few commercially important species worldwide, with the modest exceptions of common garden trials (Chap. 8) and population genetic studies (Chap. 9). The twin threats of climate change and introduced pests and diseases pose ecosystem-wide dangers, and encompass most or all conifer species, common and rare. Our vision for the future direction of conifer genetics and genomics research is therefore guided by the idea that our community will need to address a broad array of species with as much attention to adaptation and survival as to productivity. Consequently, we have divided our discussion around groups of species: (a) primary commercial species, (b) regionally important species, (c) noncommercial species of ecological significance, and (d) rare and endangered species.

Future Directions

The following discussion on the future direction of conifer genetics and genomics research is not so much predictive (glass ball) in nature but rather is based on our subjective opinions regarding research priorities. Our treatment embraces and promotes the use of both traditional genetic and modern genomic technology

platforms, since species group needs will surely vary. We anticipate most genomic technology platforms will continue to evolve and become more cost-effective, though this may not be true in all cases. High-throughput genotyping, for instance, remains costly for large sample sizes and constitutes a hurdle to some of the approaches we will recommend. In addition to the following discussion readers may wish to review the comprehensive discussion presented in a previous treatment on conifer genetics (Dean 2011).

Primary Commercial Species (Group A)

Historical investments in conifer genetic and genomic research have been directed, in very large part, to a handful of species, a trend unlikely to change in the near term. This short list might arguably include the pines *Pinus taeda*, *P. radiata*, and *P. pinaster*, the spruces *Picea glauca* and *P. abies*, a Douglas-fir (*Pseudotsuga menziesii*), and sugi (*Cryptomeria japonica*). For each, long-term investments have been made in applied tree improvement programs, supported by significant basic research efforts. Future efforts for these vital commercial species will likely include continued support of tree improvement activities including the use of molecular markers to facilitate program management as well as traditional breeding, testing, and selection activities (Table 18.1). Research on marker breeding, including genomic selection and association genetics for finite populations, is best suited for this group. These efforts should be facilitated by the development of good reference sequences, transcriptomes, genetic maps, and marker libraries. Published reference genome sequences currently exist for both spruces, *P. taeda* and *P. menziesii.* Future efforts should ensure improved sequences exist for each of these key commercial species.

The value of the gene editing tool CRISPER would be best leveraged for Group A species, especially for those amenable to cloning of plants (e.g., *C. japonica*, *P. radiata*). While cloning of select genotypes has great potential for both productivity and adaptive response to abiotic and biotic challenges, it seems unlikely the tool will enjoy significant developmental funding for other Group A species soon. Despite significant investments in recent decades, efforts to produce cost-effective clonal propagules for desired genotypes have been largely unsuccessful.

Group A species are typically well represented by large and extensive genetic tests but may not be adequately represented in range-wide or outside the native species range environments. Planning for and establishing such trials may anticipate changing climates and adaptive windows and inform decisions on assisted migration to insure adaption to future environmental conditions.

Regionally Important Species (Group B)

Regionally important species have significant commercial value and, in most cases, are likely supported by applied tree improvement programs that include breeding and testing programs and production seed orchards. Examples of species in this group include *Pinus contorta*, *Pinus monticola*, *Pinus ponderosa*, *Pinus elliottii*, and *Pinus lambertiana* in North America, *Pinus patula* in Central America, and

Table 18.1 A proposed list of activities in support of forest genetic and genomic research with suggested priorities, listed by species groups: (a) primary commercial species, (b) regionally important species, (c) noncommercial species of ecological significance, and (d) rare and endangered species

	Species groups			
Activity	a	b	c	d
Genomic resources				
Reference genome sequence	1	2	2	2
Low-density genome sequence	4	1	1	1
Transcriptome/RNAseq	1	2	2	2
Markers/genotyping	1	1	1	2
Genetic maps	1	2	2	2
Proteomics	1	3	4	4
Metabolomics	1	3	4	4
Database management	1	1	1	1
Tree improvement activities				
Traditional breeding, testing, and selection	1	2	4	4
Marker Breeding (i.e., Association, Genome Selection)	1	2	4	4
Marker Assisted Program Management (i.e., Fingerprinting, Paternity analysis, Orchard Research)	1	2	4	4
Cloning	1	2	3	1
Disease and Insect Screening	1	1	3	3
CRISPER Gene Editing	1	4	4	4
Genecology				
Geo-referenced Germplasm Collections	2	1	1	1
Reciprocal Transplant Trials (Common Garden, Provenance)	2	1	1	3
Landscape Genomics (G X E Associations)	2	1	1	3
Population Genetics				
Measures of Diversity	2	1	1	1
Conservation and Restoration	3	2	1	1
Assisted Migration	1	1	1	1
Ex situ Collections	2	2	1	1
In situ conservation set asides	2	1	1	1

Priority: 1 = high, 2 = medium, 3 = low, 4 = none

Abies alba, *Larix decidua*, *Pinus halepensis*, *Pinus koraiensis*, and *Pinus roxburghii* in Eurasia. At present, the cost of developing genomic resources for Group B species has been prohibitive or the funding simply lacking, though for species threatened by alien pests, such as the white pines in North America, exceptions may be found. Support for such exceptions should receive immediate consideration. At the very least, development of an array of genomic resources to facilitate tree improvement program management activities is encouraged.

Continued support of traditional tree improvement activities is important though some programs may choose to shift selection emphasis from productivity to adaptation. As with Group A species, we feel that Group B species will benefit from the design and installation of long-term species, provenance, and/or family trials over broad environmental ranges much like those developed in British Columbia,

Canada, and surrounding states (USA) (Ukrainetz et al. 2011; Marris 2009; O'Neill et al. 2014). Support of research directed at elucidating the potential issues associated with assisted migration in response to changing climates will be desirable.

Noncommercial Species of Ecological Significance (Group C)

The preponderance of extant conifer genetic diversity resides in the large collection of Group C species. Many, if not all, of these species serve important ecological functions in their respective floras and some may be economically important in niche markets. Most of the Podocarpaceae, Araucariaceae, and Taxaceae species probably belong in this group, as do the majority of Pinaceae and Cupressaceae. A significant number of these species have suffered from over-harvesting and habitat loss. In general, species in this group have received little to no genetic/genomic research attention.

In view of the existential threat of climate change for many of these species we believe collaborative efforts to genetically characterize as many of these species as possible is an important priority. These efforts might include collection and curation of large, geo-referenced germplasm samples (DNA, seed, RNA), RNAseq analyses to characterize known adaptive gene space, and research aimed at elucidating important genotype by environment associations (landscape genomics, Chap. 13) that may guide future species management efforts like assisted migration or conservation.

For some members of this group special attention may be required due to imminent threats from introduced pests and diseases. For instance, the North American white pines, as detailed in Chap. 15, are under duress from the combined effects of the white pine blister rust (*Cronartium ribicola*), bark beetles, and climate change. Coordinated efforts by Federal and State agencies that employ disease screening, large-scale geo-referenced germplasm collections, and development of key genomic resources have been successful in identifying resistance and guiding deployment decisions. This program should serve as a model for other issues that exist or are likely to arise.

Rare and Endangered Species (Group D)

This is not a short list. As noted in Chap. 1, nearly a third of all conifer species were listed as vulnerable, endangered, or critically endangered in 2015, according to a review by the International Union for Conservation of Nature Red List of Threatened Species (IUCN 2017). Thirty-nine species are considered relicts, known to exist in a single locality, often confined to a few km^2. Surely, some of these species are doomed to extinction in the wild and may not warrant investment. Others may benefit from timely intervention, which could take the form of ex situ collection and curation of seed or plantations or in situ conservation and restoration efforts.

Moving Forward: Policy and Management Changes Needed

Our universe of suggested future research activities obviously far outstrips existing worldwide forest genetic expertise, available resources, and institutional infrastructure for implementation. Even given priority setting, as suggested in Table 18.1**,** the hurdle remains high. The goals are, nevertheless, worthy of serious attention.

An ever-increasing world population will drive demand for forest products while simultaneously contributing to the erosion of forest ecosystem health through climate change and movement of exotic pests and diseases around the globe. We believe forest genetics and ecosystem management are key and indispensable elements of any integrated plan to manage forest health and/or enhance forest productivity and deserve renewed support from international organizations (i.e., FAO, IUFRO), federal and regional governments, and private companies, worldwide. In the United States, for instance, this would likely take the form of a dramatically increased leadership role of the US Forest Service, renewed state agency and university support, and corporate responsibility, all working closely together and interacting with parallel groups in Canada and Mexico.

History has provided lessons about how science policy, economics and social factors have shaped the forest genetics discipline, as they have for forest sciences in general. Perhaps most critical to forest genetics (FG) and tree improvement (TI) is the inevitable shifting of research priorities and policy that occurs over time, sometimes over very short periods. Almost without exception, successful FG/TI endeavors have depended upon uninterrupted, long-term fiscal and policy support, coupled with a market for the fruits of the research. By its very nature, tree breeding and forest genetics research is conducted over extended periods of time. In the United States, the loss of established FG/TI programs in the USDA USFS, state forestry programs, private industry, and many universities essentially reduced the discipline to reliance on University/State/Industrial Cooperative programs and relatively short-term grants, both of which favor a very select group of species and scientists. While grants have been enormously beneficial in advancing basic research objectives, particularly in genomics, their relatively short funding cycles are not well designed to deal with sustained development and application of genetic resources for the multitude of species that make up the bulk of our natural forest ecosystems. To quote a passage from our paper on the history of forest genetics research in the United States (Wheeler et al. 2015):

> Beginning in the 1980s, it became clear that the "improvement" focus of most FG/TI work was becoming unfashionable as broader societal and policy trends began to strongly favor ecosystem values over forest commodities. Although this shift disfavored tree improvement, it did not render obsolete the need for forest genetics research. Challenged by biotic and abiotic agents, and pervasive human influences, our forests must still be managed if species and populations are to be retained as ecosystem components. Resource management, even of natural forests, requires an understanding of their genetic structure and, sometimes, how to manipulate it in the face of threats. For this purpose, genomics and related technologies are very well suited and, indeed, indispensable.

So how might the priorities noted in Table 18.1 be addressed? Worldwide, efforts should be forged that transcend political and economic boundaries, requiring multiagency, NGOs, and government's support, in addition to traditional sources of support within individual countries. For the United States, we have advocated for a multi-pronged approach (Wheeler et al. 2015) that includes the following recommendations:

- Funding – Development of a balanced and broad-based funding model that includes
  - *Long-term* support for applied tree improvement research in support of improved productivity and wood quality
  - *Long-term* support for forest and ecosystem health
  - *Short-term* support for biotechnology, genomic, and ecosystem management
- Research Coordination – Support and development of national coordinating boards, akin to the National Plant Board in the United States, which would work in a strategic manner with stakeholders, government agencies, and grant funding programs to insure investments made by long-term and short-term funders are made in a balanced and coordinated manner, addressing strategic tree improvement needs and forest health threats.
- Education
  - Increased training of scientists and technical staff in forest genetics and tree improvement, including traditional and genomic technologies and applications
  - Continuing education for K-12 teachers and the public on the importance of production and natural forests, the threats they face, and the technologies used to address their management.
- Leadership – By its very nature, forest genetics and tree improvement are long-term propositions that require long-term institutional support and memory. Historically, in the United States, this was provided by the US Forest Service and, to a lesser degree, industry and universities. Today, it resides almost entirely within a few university/industrial cooperatives, and then for only a handful of species. We advocate for renewed leadership by the US Forest Service that includes
  - Renewed support of forest genetics in regional experimental stations
  - Creation of regional centers of research emphasis in universities that address developing forest health issues

While much of the language noted above is US-centric, we believe other regions of the world would benefit equally from parallel efforts focused on their native conifers.

Appendix 1

Genus species	Primary reference
Agathis alba	(Lam.) Foxw.
Agathis atropurpurea	B. Hyland
Agathis australis	(D. Don) R. A. Salisbury
Agathis borneensis	O. Warburg
Agathis dammara	(A. Lambert) L. Richard
Agathis lanceolata	O. Warburg
Agathis macrophylla	(Lindley) M. T. Masters
Agathis microstachya	J. F. Bailey & C. White
Agathis montana	de Laubenfels
Agathis moorei	(Lindley) M. T. Masters
Agathis ovata	(C. Moore ex Vieillard) O. Warburg
Agathis robusta	(C. Moore ex F. J. Mueller) F. M. Bailey
Araucaria angustifolia	(A. Bertoloni) O. Kuntze
Araucaria araucana	(G. Molina) K. Koch
Araucaria bernieri	J. Buchholz
Araucaria bidwillii	W. J. Hooker
Araucaria biramulata	J. Biramulata
Araucaria columnaris	J. R. Forster
Araucaria cunninghamii	D. Don
Araucaria heterophylla	(R. A. Salisbury) Franco
Araucaria humboldtensis	J. Buchholz
Araucaria hunsteinii	K. Schumann
Araucaria laubenfelsii	Corbasson
Araucaria luxurians	(Ad. Brongniart & Gris) de Laubenfels
Araucaria montana	Ad. Brongniart & Gris
Araucaria muelleri	(Carrière) Ad. Brongniart & Gris
Araucaria nemorosa	de Laubenfels
Araucaria rulei	F. J. Mueller ex Lindley
Araucaria schmidii	de Laubenfels
Araucaria scopulorum	de Laubenfels
Araucaria subulata	Vieillard

(continued)

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5

Genus species	Primary reference
Wollemia nobilis	W. G. Jones, K. Hill & J. M. Allen
Actinostrobus pyramidalis	Miquel
Athrotaxis cupressoides	D. Don
Athrotaxis x laxifolia	W. J. Hooker
Athrotaxis selaginoides	D. Don
Austrocedrus chilensis	(D. Don) Pichi Sermolli & Bizzarri
Callitris canescens	(Parlatore) S.T. Blake
Callitris columellaris	F. J. Mueller
Callitris endlicheri	(Parlatore) F. M. Bailey
Callitris macleayana	(F. J. Mueller) F. J. Mueller
Callitris preissii	Miquel in J. G. Lehman
Callitris rhomboidea	R. Brown ex A. Richard & L. Richard
Callitris verrucosa	(A. Cunningham ex Endlicher) F. J. Mueller
Calocedrus decurrens	(J. Torrey) Florin
Calocedrus formosana	(Florin) Florin
Calocedrus macrolepis	W. Kurz
Chamaecyparis formosensis	J. Matsumura
Chamaecyparis lawsoniana	(A. Murray bis) Parlatore
Chamaecyparis nootkatensis	(D. Don) Spach
Chamaecyparis obtusa	(P. Siebold & Zuccarini) Endlicher
Chamaecyparis pisifera	(P. Siebold & Zuccarini) Endlicher
Chamaecyparis thyoides	(Linnaeus) N. Britton, Sterns, and Poggenburg
Cryptomeria japonica	(Linnaeus fil.) D. Don
Cunninghamia konishii	Hayata
Cunninghamia lanceolata	(A. Lambert) W. J. Hooker
Cupressus arizonica	E. Greene
Cupressus bakeri	Jepson
Cupressus cashmeriana	Royle ex Carrière
Cupressus chengiana	S. Y. Hu
Cupressus duclouxiana	P. Hickel ex A. Camus
Cupressus dupreziana	A. Camus
Cupressus funebris	Endlicher
Cupressus gigantea	W. C. Cheng & L. K. Fu
Cupressus goveniana	G. Gordon
Cupressus guadalupensis	S. Watson
Cupressus x leylandii	A. B. Jackson & Dallimore
Cupressus lusitanica	P. Miller
Cupressus macnabiana	A. Murray bis
Cupressus macrocarpa	K. Hartweg ex. G. Gordon
Cupressus nootkatensis	D. Don in A. Lambert
Cupressus sargentii	Jepson
Cupressus sempervirens	Linnaeus
Cupressus torulosa	D. Don in A. Lambert
Cupressus vietnamensis	(Farjon & T. H. Nguyễn) Silba
Diselma archeri	J. Hooker
Fitzroya cupressoides	(G. Molina) I. Johnston

Genus species	Primary reference
Fokienia hodginsii	(S. Dunn) A. Henry & H. H. Thomas
Glyptostrobus pensilis	(Staunton) K. Koch
Juniperus bermudiana	Linnaeus
Juniperus cedrus	P. Webb & Berthelot
Juniperus chinensis	Linnaeus
Juniperus communis	Linnaeus
Juniperus convallium	Rehder & E. H. Wilson
Juniperus foetidissima	Willdenow
Juniperus formosana	Hayata
Juniperus horizontalis	Moench
Juniperus maritima	R. P. Adams
Juniperus monosperma	(Engelmann) C. Sargent
Juniperus occidentalis	J. Hooker
Juniperus osteoperma	(J. Torrey) E. Little
Juniperus oxycedrus	Linnaeus
Juniperus phoenicea	Linnaeus
Juniperus pingii	W. C. Cheng in de Ferré
Juniperus procera	C. F. Hochstetter ex Endlicher
Juniperus przewalskii	Kom.
Juniperus rigida	P. Siebold & Zuccarini
Juniperus sabina	Linnaeus
Juniperus saltillensis	M. Hall
Juniperus saltuaria	Rehder & E. H. Wilson
Juniperus scopulorum	C. Sargent
Juniperus squamata	Buchanan-Hamilton ex D. Don
Juniperus tibetica	V. Komarov
Juniperus virginiana	Linnaeus
Libocedrus bidwillii	J. Hooker
Libocedrus plumosa	(D. Don) C. Sargent
Libocedrus yateensis	Guillaumin
Metasequoia glyptostroboides	H. H. Hu & W. C. Cheng
Microbiota decussata	V. Komarov
Neocallitropsis pancheri	(Carrière) de Laubenfels
Papuacedrus papuana	(F. J. Mueller) H. L. Li
Pilgerodendron uviferum	(D. Don) Florin
Platycladus orientalis	(Linnaeus) Franco
Sequoia sempervirens	(D. Don) Endlicher
Sequoiadendron giganteum	(Lindley) J. Buchholz
Taiwania cryptomerioides	Hayata
Taxodium distichum	(Linnaeus) L. Richard
Taxodium huegelii	C. Lawson
Tetraclinis articulata	(M. H. Vahl) M. T. Masters
Thuja koraiensis	T. Nakai
Thuja occidentalis	Linnaeus
Thuja plicata	D. Don

(continued)

Genus species	Primary reference
Thuja standishii	(G. Gordon) Carrière
Thuja sutchuenensis	Franchet
Thujopsis dolabrata	(Linnaeus) P. Siebold & Zuccarini
Widdringtonia nodiflora	(Linnaeus) Powrie
Widdringtonia schwarzii	(Marloth) M. T. Masters
Widdringtonia wallichii	Endlicher ex Carrière
Abies alba	P. Miller
Abies balsamea	(Linnaeus) P. Miller
Abies bracteata	(D. Don) P. Poiteau
Abies cephalonica	J. C. Loudon
Abies concolor	(G. Gordon & Glendinning) F. G. Hildebrand
Abies fargesii	Franchet
Abies grandis	(D. Douglas ex D. Don in A. Lambert) Lindley
Abies holophylla	Maximowicz
Abies homolepis	P. Siebold & Zuccarini
Abies koreana	E. H. Wilson
Abies lasiocarpa	(W. J. Hooker) T. Nuttall
Abies magnifica	A. Murray bis
Abies nephrolepis	(Trautvetter) Maximowicz
Abies nordmanniana	(Steven) Spach
Abies numidica	de Lannoy ex Carrière
Abies pindrow	(Royle ex D. Don) Royle
Abies pinsapo	Boissier
Abies procera	Rehder
Abies sachalinensis	(Friedr. Schmidt) M. T. Masters
Abies sibirica	Ledebour
Abies veitchii	Lindley
Cathaya argyrophylla	W. Y. Chun & Kuang
Cedrus atlantica	(Endl.) Manetti ex Carrière
Cedrus deodara	(D. Don) G. Don
Cedrus libani	A. Richard
Keteleeria evelyniana	Mast.
Larix decidua	P. Miller
Larix x eurolepis	A. Henry
Larix gmelinii	(F. Ruprecht) Kuzeneva
Larix griffithii	J. Hooker
Larix kaempferi	(A. Lambert) Carrière
Larix laricina	(Du Roi) K. Koch
Larix mastersiana	Rehder & E. H. Wilson
Larix occidentalis	T. Nuttall
Larix x polonica	Racib.
Larix potaninii	Batalin
Larix sibirica	Ledebour
Picea abies	(Linnaeus) H. Karsten
Picea alcoquiana	(J. G. Veitch ex Lindley) Carrière
Picea asperata	M. T. Masters

Genus species	Primary reference
Picea brachytyla	(Franchet) E. Pritzel
Picea breweriana	S. Watson
Picea crassifolia	Kom.
Picea engelmannii	C. Parry ex Engelmann
Picea x fennica	(Regel) Kom.
Picea glauca	(Moench) A. Voss
Picea glehnii	(Friedr. Schmidt) M. T. Masters
Picea jezoensis	(P. Siebold & Zuccarini) Carrière
Picea koraiensis	Nakai
Picea koyamae	Shirasawa
Picea likiangensis	(Franchet) E. Pritzel
Picea mariana	(P. Miller) N. Britton, Sterns & Poggenburg
Picea maximowiczii	E. Regel ex M. T. Masters
Picea meyeri	Rehder & E. H. Wilson
Picea morrisonicola	Hayata
Picea neoveitchii	M. T. Masters
Picea obovata	Ladeb.
Picea omorika	(Pančić) Purkyně
Picea orientalis	(Linnaeus) J. Link
Picea pungens	Engelmann
Picea purpurea	M. T. Masters
Picea rubens	C. Sargent
Picea schrenkiana	F. E. L. Fischer & C. Meyer
Picea x shirasawae	Yasaka Hayashi
Picea sitchensis	(Bongard) Carrière
Picea smithiana	(N. Wallich) Boissier
Picea wilsonii	M. T. Masters
Pinus albicaulis	Engelmann
Pinus aristata	Engelmann
Pinus arizonica	Engelmann
Pinus armandii	Franchet
Pinus attenuata	J. Lemmon
Pinus ayacahuite	C. G. Ehrenberg ex D. F. L. Schlechtendal
Pinus balfouriana	Greville & J. Balfour
Pinus banksiana	A. Lambert
Pinus brutia	Ten.
Pinus bungeana	Zuccarini ex Endlicher
Pinus canariensis	R. Sweet ex K. Sprengel
Pinus caribaea	R. Sweet ex K. Sprengel
Pinus cembra	Linnaeus
Pinus cembriodes	Zuccarini
Pinus clausa	(Chapm. ex Engelm.) Vasey ex Sarg.
Pinus contorta	D. Douglas ex J. C. Loudon
Pinus coulteri	D. Don
Pinus cubensis	Grisebach

(continued)

Genus species	Primary reference
Pinus culminicola	Andresen & J. Beaman
Pinus densata	M. T. Masters
Pinus densiflora	P. Siebold & Zuccarini
Pinus devoniana	Lindley
Pinus douglasiana	M. Martínez
Pinus durangensis	M. Martínez
Pinus echinata	P. Miller
Pinus edulis	Engelmann
Pinus elliottii	Engelmann
Pinus engelmannii	Carrière
Pinus flexilis	E. James
Pinus gerardiana	N. Wallich ex D. Don
Pinus greggii	Engelmann ex Parlatore
Pinus halepensis	P. Miller
Pinus hartwegii	Lindley
Pinus heldreichii	H. Christ
Pinus henryi	Mast.
Pinus herrerae	M. Martínez
Pinus hwangshanensis	W. Y. Hsia
PInus jeffreyi	Greville & J. Balfour
Pinus kesiya	Royle ex G. Gordon
Pinus koraiensis	P. Siebold & Zuccarini
Pinus krempfii	P. Lecomte
Pinus lambertiana	D. Douglas
Pinus lawsonii	Roezl ex G. Gordon
Pinus leiophylla	Schiede & Deppe ex D. F. L. Schlechtendal & Chamisso
Pinus longaeva	D. Bailey
Pinus luchuensis	H. Mayr
Pinus lumholtzii	B. Robinson & Fernald
Pinus massoniana	A. Lambert
Pinus maximartinezii	Rzedowski
Pinus maximinoi	H. Moore
Pinus merkusii	Junghuhn & de Vriese
Pinus monophylla	J. Torrey & Frémont
Pinus montezumae	A. Lambert
Pinus monticola	D. Douglas ex D. Don
Pinus mugo	Turra
Pinus muricata	D. Don
Pinus neilreichiana	H. Reichardt
Pinus nelsonii	G. R. Shaw
Pinus nigra	J. F. Arnold
Pinus occidentalis	Swartz
Pinus oocarpa	Schiede ex D. F. L. Schlechtendal
Pinus palustris	P. Miller
Pinus parviflora	P. Siebold & Zuccarini

Genus species	Primary reference
Pinus patula	Schiede & Deppe ex D. F. L. Schlechtendal & Chamisso
Pinus peuce	Grisebach
Pinus pinaster	W. Aiton
Pinus pinceana	G. Gordon
Pinus pinea	Linnaeus
Pinus ponderosa	D. Douglas ex P. Lawson & C. Lawson
Pinus pringlei	G. R. Shaw
Pinus pseudostrobus	Lindley
Pinus pumila	(P. Pallas) E. Regel
Pinus pungens	A. Lambert
Pinus radiata	D. Don
Pinus resinosa	W. Aiton
Pinus rhaetica	Brügger
Pinus rigida	P. Miller
Pinus roxburghii	C. Sargent
Pinus sabiniana	D. Douglas ex D. Don
Pinus serotina	Michx.
Pinus sibirica	Du Tour
Pinus strobiformis	Engelmann
Pinus strobus	Linnaeus
Pinus sylvestris	Linnaeus
Pinus tabluliformis	Carrière
Pinus taeda	Linnaeus
Pinus taiwanensis	Hayata
Pinus tecunumanii	Eguiluz & J. P. Perry
Pinus teocote	Schiede & Deppe ex Chamisso & D. F. L. Schlechtendal
Pinus thunbergii	Parlatore
Pinus torreyana	C. Parry ex Carrière
Pinus tropicalis	P. Morelet
Pinus unicinata	Ramond ex D. C.
Pinus virginiana	P. Miller
Pinus wallichiana	A. B. Jackson
Pinus yunnanensis	Franchet
Pseudolarix amabilis	(J. Nelson) Rehder
Pseudotsuga menziesii	(Mirbel) Franco
Tsuga canadensi	(Linnaeus) Carrière
Tsuga carolinian	Engelmann
Tsuga chinenesis	(Franchet) E. Pritzel
Tsuga diversifolia	(Maximowicz) M. T. Masters
Tsuga dumosa	(D. Don) A. Eichler
Tsuga heterophylla	(Rafinesque) C. Sargent
Tsuga x jeffreyi	(A. Henry) A. Henry
Tsuga mertensiana	(Bongard) Carrière

(continued)

Genus species	Primary reference
Tsuga sieboldii	Carrière
Acmopyle pancheri	(Ad. Brongniart & Gris) Pilger
Acmopyle sahniana	J. Buchholz
Afrocarpus falcatus	(Thunberg) C. Page
Afrocarpus gracilior	(Pilg.) C. N. Page
Afrocarpus mannii	(J. Hooker) C. Page
Dacrycarpus cinctus	(Pilger) de Laubenfels
Dacrycarpus compactus	(Wasscher) de Laubenfels
Dacrycarpus dacrydioides	(A. Richard) de Laubenfels
Dacrycarpus imbricatus	(Blume) de Laubenfels
Dacrycarpus vieillardii	(Parlatore) de Laubenfels
Dacrydium araucarioides	Ad. Brongniart & Gris
Dacrydium balansae	Ad. Brongniart & Gris
Dacrydium cupressinum	Solander ex J. G. Forster
Dacrydium elatum	(Roxburgh) N. Wallich ex W. J. Hooker
Dacrydium gracile	de Laubenfels
Dacrydium guillauminii	J. Buchholz
Dacrydium lycopodioides	Ad. Brongniart & Gris
Dacrydium nausoriense	de Laubenfels
Dacrydium nidulum	de Laubenfels
Falcatifolium falciforme	(Parlatore) de Laubenfels
Falcatifolium gruezoi	de Laubenfels
Falcatifolium taxoides	(Ad. Brongniart & Gris) de Laubenfels
Halocarpus bidwillii	(J. Hooker ex T. Kirk) Quinn
Halocarpus biformis	(W. J. Hooker) Quinn
Halocarpus kirkii	(F. J. Mueller ex Parlatore) Quinn
Lagarostrobos franklinii	(J. Hooker) Quinn
Lepidothamnus fonkii	R. Philippi
Lepidothamnus laxifolius	(J. Hooker) Quinn
Manoao colensai	(W. J. Hooker) Molloy
Microcachrys tetragona	J. Hooker
Nageia fleuryi	(P. Hickel) de Laubenfels
Nageia formosensis	(Dummer) C. N. Page
Nageia nagi	(Thunberg) O. Kuntze
Nageia wallichiana	(C. Presl) O. Kuntze
Parasitaxus ustus	(Vieillard) de Laubenfels
Pherosphaera fitzgeraldii	(F. Muell.) Hook. F.
Pherosphaera hookeriana	W. Archer bis
Phyllocladus aspleniifolius	(Labillardière) J. Hooker
Phyllocladus hypophyllus	J. Hooker
Phyllocladus toatoa	Molloy
Phyllocladus trichomanoides	D. Don
Podocarpus acutifolius	T. Kirk
Podocarpus affinis	B. Seemann
Podocarpus annamiensis	N. Gary
Podocarpus aristulatus	Parl.

Genus species	Primary reference
Podocarpus brassii	Pilger
Podocarpus celatus	de Laubenfels
Podocarpus chinensis	Wall. ex J. Forbes
Podocarpus costalis	C. Presl
Podocarpus cunninghamii	Colenso
Podocarpus dispermus	C. White
Podocarpus elongatus	(W. Aiton) L'Héritier ex Persoon
Podocarpus gnidioides	Carrière
Podocarpus guatemalensis	P. Standley
Podocarpus henkelii	Stapf ex Dallimore & A. B. Jackson
Podocarpus lambertii	Klotzsch ex Endlicher
Podocarpus latifolius	(Thunberg) R. Brown ex Mirbel
Podocarpus lawrencei	J. Hooker
Podocarpus longifoliolatus	Pilger
Podocarpus macrophyllus	(Thunberg) R. Sweet
Podocarpus matudae	C. Lundell
Podocarpus neriifolius	D. Don
Podocarpus nivalis	W. J. Hooker
Podocarpus nubigenus	Lindley
Podocarpus polystachyus	R. Brown ex Endlicher
Podocarpus rumphii	Blume
Podocarpus salignus	D. Don
Podocarpus smithii	de Laubenfels
Podocarpus sylvestris	J. Buchholz
Podocarpus totara	G. Bennett ex D. Don
Prumnopitys andina	(Poeppig ex Endlicher) de Laubenfels
Prumnopitys ferruginea	(G. Bennett ex D. Don) de Laubenfels
Prumnopitys ferruginoides	(R. Compton) de Laubenfels
Prumnopitys ladei	(F. M. Bailey) de Laubenfels
Prumnopitys taxifolia	(J. Banks & Solander ex D. Don) de Laubenfels
Retrophyllum comptonii	(J. Buchholz) C. Page
Retrophyllum minus	(Carrière) C. N. Page
Retrophyllum rospigliosii	(Pilger) C. Page
Retrophyllum vitiense	(B. Seemann) C. Page
Saxegothaea conspicua	Lindley
Sundacarpus amarus	(Blume) C. N. Page
Sciadopitys verticillata	(Thunberg) P. Siebold & Zuccarini
Amentotaxus argotaenia	(Hance) Pilger
Amentotaxus formosana	H. L. Li
Amentotaxus yunnanensis	H. L. Li
Austrotaxus spicata	R. Compton
Cephalotaxus fortunei	W. J. Hooker
Cephalotaxus hainanensis	H. L. Li
Cephalotaxus harringtonii	(J. Knight ex Jas. Forbes) K. Koch

(continued)

Genus species	Primary reference
Cephalotaxus lanceolata	K. M. Feng ex C. Y. Cheng W. C. Cheng & L. K. Fu
Cephalotaxus latifolia	W. C. Cheng & L. K. Fu ex L. K. Fu & R. R. Mill
Cephalotaxus mannii	J. Hooker
Cephalotaxus oliveri	M. T. Masters
Cephalotaxus sinensis	(Rehder & E. H. Wilson) H. L. Li
Pseudotaxus chienii	(W. C. Cheng) W. C. Cheng
Taxus baccata	Linnaeus
Taxus brevifolia	T. Nuttall
Taxus canadensis	H. Marshall
Taxus cuspidata	P. Siebold & Zuccarini
Taxus floridana	A. W. Chapman
Taxus fuana	Nan Li & R. R. Mill
Taxus globosa	D. F. L. Schlechtendal
Taxus x hunnewelliana	Rehder
Taxus x media	Rehder
Taxus sumatrana	(Miquel) de Laubenfels
Taxus wallichiana	Zuccarini
Torreya californica	J. Torrey
Torreya fargesii	Franchet
Torreya grandis	Fortune ex Lindley
Torreya jackii	W. Y. Chun
Torreya nucifera	(Linnaeus) P. Siebold & Zuccarini
Torreya taxifolia	G. Arnott

Appendix 2

Phylogenetic relationships within genus *Pinus* as determined from full plastome alignment. Cladogram based on maximum likelihood (ML) topology, showing support values below branches as ML bootstrap support/Bayesian posterior probability/parsimony bootstrap support. Support values are shown only for nodes with less than 100% bootstrap support, posterior probabilities less than 1.0, or both; single values indicate either ML bootstrap support or Bayesian posterior probability. Branches not supported in Bayesian or parsimony analysis are indicated with *.

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5

From Parks et al. (2012)

References

Aagaard, J. E., Krutovskii, K. V., & Strauss, S. H. (1998a). RAPDs and allozymes exhibit similar levels of diversity and differentiation among populations and races of Douglas-fir. *Heredity, 81*(1), 69–78.

Aagaard, J. E., Krutovskii, K. V., & Strauss, S. H. (1998b). RAPD markers of mitochondrial origin exhibit lower population diversity and higher differentiation than RAPDs of nuclear origin in Douglas fir. *Molecular Ecology, 7*(7), 801–812.

Aagaard, J. E., Vollmer, S. S., Sorensen, F. C., & Strauss, S. H. (1995). Mitochondrial DNA products among RAPD profiles are frequent and strongly differentiated between races of Douglas-fir. *Molecular Ecology, 4*(4), 441–446.

Abrahamsson, S., Nilsson, J. E., Wu, H., García Gil, M. R., & Andersson, B. (2012). Inheritance of height growth and autumn cold hardiness based on two generations of full-sib and half-sib families of *Pinus sylvestris*. *Scandinavian Journal of Forest Research, 27*(5), 405–413.

Abril, N., Gion, J. M., Kerner, R., Müller-Starck, G., Cerrillo, R. M. N., Plomion, C., et al. (2011). Proteomics research on forest trees, the most recalcitrant and orphan plant species. *Phytochemistry, 72*(10), 1219–1242.

Adams, M. D., & Kelley, J. M. (1991). Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science, 252*(5013), 1651–1656.

Adams, R. P. (2015). Allopatric hybridization and introgression between *Juniperus maritima* and *J. scopulorum*. *Evidence from nuclear and cpDNA and leaf terpenoids*. *Phytologia, 97*, 55–66.

Adams, R. P., Bartel, J. A., & Price, R. A. (2009). A new genus, *Hesperocyparis*, for the cypresses of the western hemisphere. *Phytologia, 91*(1), 160–185.

Adams, R. P., Bartel, J. A., Terry, R., Callahan, F., & Bisbee, J. (2014). Taxonomy of *Hesperocyparis montana, H. revealiana,* and *H. stephensonii:* Evidence from leaf essential oils analyses and DNA sequences. *Phytologia, 96*(2), 71–83.

Adams, W. T. (1983). *Application of isozymes in tree breeding. Developments in plant genetics and breeding* (pp. 381–400). Amsterdam: Elsevier Science Publishers BV.

Adams, W. T., & Burczyk, J. (2000). Magnitude and implications of gene flow in gene conservation reserves. In A. Young, D. Boshier, & T. Boyle (Eds.), *Forest conservation genetics: Principles and practice* (pp. 215–244). Oxon/Collingwood: Commonwealth Scientific and Industrial Research Organization (CSIRO) Publishing/CABI Publishing.

Adams, W. T., Hipkins, V. D., Burczyk, J., & Randall, W. K. (1997). Pollen contamination trends in a maturing Douglas-fir seed orchard. *Canadian Journal of Forest Research, 27*(1), 131–134.

Adams, W. T., Zuo, J., Shimizu, J. Y., & Tappeiner, J. C. (1998). Impact of alternative regeneration methods on genetic diversity in coastal Douglas-fir. *Forest Science, 44*(3), 390–396.

Adie, H., & Lawes, M. J. (2011). Podocarps in Africa: Temperate zone relicts or rainforest survivors? In B. L. Turner & L. A. Cernusak (Eds.), *Ecology of the Podocarpaceae in tropical forests* (Smithsonian Contributions to Botany, No. 95) (pp. 79–100). Washington, D.C.: Smithsonian Institution Scholarly Press.

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5

Adomas, A., Heller, G., Olson, A., Osborne, J., Karlsson, M., Zyl, J. L. V., et al. (2008). Comparative analysis of transcript abundance in *Pinus sylvestris* after challenge with a saprotrophic, pathogenic or mutualistic fungus. *Tree Physiology, 28*(6), 885.

Aegerter, B. J., & Gordon, T. R. (2006). Rates of pitch canker induced seedling mortality among Pinus radiata families varying in levels of genetic resistance to *Gibberella circinata* (anamorph *Fusarium circinatum*). *Forest Ecology and Management, 235*(1), 14–17.

Aguirre-Planter, E., Furnier, G. R., & Eguiarte, L. E. (2000). Low levels of genetic variation within and high levels of genetic differentiation among populations of species of *Abies* from southern Mexico and Guatemala. *American Journal of Botany, 87*(3), 362–371.

Agúndez, D., Degen, B., Von Wuehlisch, G., & Alía, R. (1999). Multilocus analysis of *Pinus halepensis* Mill. From Spain: Genetic diversity and clinal variation. *Silvae Genetica, 48*, 173–178.

Ahuja, M. R. (2005). Polyploidy in gymnosperms: revisited. *Silvae Genetica, 54*(2), 59–69.

Ahuja, M. R., Devey, M. E., Groover, A. T., Jermstad, K. D., & Neale, D. B. (1994). Mapped DNA probes from loblolly pine can be used for restriction fragment length polymorphism mapping in other conifers. *Theoretical and Applied Genetics, 88*(3–4), 279–282.

Ahuja, M. R., & Neale, D. B. (2002). Origins of polyploidy in coast redwood (*Sequoia sempervirens* (D. Don) Endl.) and relationship of coast redwood to other genera of Taxodiaceae. *Silvae Genetica, 51*(2–3), 93–99.

Ahuja, M. R., & Neale, D. B. (2005). Evolution of genome size in conifers. *Silvae Genetica, 54*(3), 126–137.

Aitken, S. N., & Adams, W. T. (1997). Spring cold hardiness under strong genetic control in Oregon populations of *Pseudotsuga menziesii* var. *menziesii*. *Canadian Journal of Forest Research, 27*(11), 1773–1780.

Aitken, S. N., & Libby, W. J. (1994). Evolution of the pygmy-forest edaphic subspecies of *Pinus contorta* across an ecological staircase. *Evolution, 48*, 1009–1019.

Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., & Curtis-McLane, S. (2008). Adaptation, migration or extirpation: Climate change outcomes for tree populations. *Evolutionary Applications, 1*(1), 95–111.

Alberto, F. J., Aitken, S. N., Alia, R., Gonzalez-Martinez, S. C., Hanninen, H., Kremer, A., Lefevre, F., Lenormand, T., Yeaman, S., Whetten, R., & Savolainen, O. (2013). Potential for evolutionary responses to climate change - evidence from tree populations. *Global Change Biology, 19*, 1645–1661. https://doi.org/10.1111/gcb.12181.

Alfaro, R., John, I., King, N., El-Kassaby, Y. A., Brown, G., & Lewis, K. (2000). Use of artificial infestations in the selection of Sitka spruce genotypes for resistance to the white pine weevil in British Columbia. In R. Alfaro, et al. (Eds.), *Protection of World Forests from Insect Pests: Advances in Research*. Papers presented at the XXI IUFRO World Congress (7–12 August 2000, Kuala Lumpur, Malaysia), IUFRO World Series Vol. 11.

Alfaro, R. I., Fady, B., Vendramin, G. G., Dawson, I. K., Fleming, R. A., Sáenz-Romero, C., et al. (2014). The role of forest genetic resources in responding to biotic and abiotic factors in the context of anthropogenic climate change. *Forest Ecology and Management, 333*, 76–87.

Ali, I. F., Neale, D. B., & Marshall, K. A. (1991). Chloroplast DNA restriction fragment length polymorphism in *Sequoia sempervirens* D. Don Endl., *Pseudotsuga menziesii* (Mirb.) Franco, *Calocedrus decurrens* (Torr.), and *Pinus taeda* L. *Theoretical and Applied Genetics, 81*(1), 83–89.

Allona, I., Quinn, M., Shoop, E., Swope, K., Cyr, S. S., Carlis, J., et al. (1998). Analysis of xylem formation in pine by cDNA sequencing. *Proceedings of the National Academy of Sciences, 95*(16), 9693–9698.

Alosi, M. C., & Neale, D. B. (1992). Light-and phytochrome-mediated gene expression in Douglas-fir seedlings. *Physiologia Plantarum, 86*(1), 71–76.

Alosi, M. C., Neale, D. B., & Kinlaw, C. S. (1990). Expression of cab genes in Douglas-fir is not strongly regulated by light. *Plant Physiology, 93*(2), 829–832.

Al-Rabab'ah, M. A., & Williams, C. G. (2004). An ancient bottleneck in the Lost Pines of central Texas. *Molecular Ecology, 13*(5), 1075–1084.

Alvarez, M. (2007). *The state of America's forests*. DIANE Publishing.
Alverez, I., & Wendel, J. F. (2003). Ribosomal ITS sequences and plant phylogenetic inference. *Molecular Phylogenetics and Evolution, 29*, 417–434.
Amarasinghe, V., & Carlson, J. E. (1998). Physical mapping and characterization of 5S rRNA genes in Douglas-fir. *Journal of Heredity, 89*(6), 495–500.
Amerson, H. V., Nelson, C. D., Kubisiak, T. L., Kuhlman, E. G., & Garcia, S. A. (2015). Identification of nine pathotype-specific genes conferring resistance to fusiform rust in loblolly pine (*Pinus taeda* L.). *Forests, 6*(8), 2739–2761.
Anderson, E. (1948). Hybridization of the habitat. *Evolution, 2*, 1–9.
Anderson, E. (1949). *Introgressive Hybridization*. New York, NY: Wiley.
Anderson, L. L., Hu, F. S., Nelson, D. M., Petit, R. J., & Paige, K. N. (2006). Ice-age endurance: DNA evidence of a white spruce refugium in Alaska. *Proceedings of the National Academy of Sciences, 103*(33), 12447–12450.
Anderson, L. L., Hu, F. S., & Paige, K. N. (2011). Phylogeographic history of white spruce during the last glacial maximum: Uncovering cryptic refugia. *The Journal of Heredity, 102*(2), 207–216.
Anekonda, T. S., Adams, W. T., Aitken, S. N., Neale, D. B., Jermstad, K. D., & Wheeler, N. C. (2000). Genetics of cold hardiness in a cloned full-sib family of coastal Douglas-fir. *Canadian Journal of Forest Research, 30*(5), 837–840.
Anekonda, T. S., Lomas, M. C., Adams, W. T., Kavanagh, K. L., & Aitken, S. N. (2002). Genetic variation in drought hardiness of coastal Douglas-fir seedlings from British Columbia. *Canadian Journal of Forest Research, 32*(10), 1701–1716.
Aravanopoulos, F. A. (2011). Genetic monitoring in natural perennial plant populations. *Botany, 89*(2), 75–81.
Aravanopoulos, F. A. (2016). Conservation and monitoring of tree genetic resources in temperate forests. *Current Forestry Reports, 2*(2), 119–129.
Aravanopoulos, F. A., Tollefsrud, M. M., Graudal, L., Koskela, J., Kätzel, R., Soto, A., et al. (2015). Development of genetic monitoring methods for genetic conservation units of forest trees in Europe. European Forest Genetic Resources Programme (EUFORGEN), *Bioversity International, Rome, Italy. xvi*, p.3.
Arcade, A., Faivre-Rampant, P., Pâques, L. E., & Prat, D. (2002). Localisation of genomic regions controlling microdensitometric parameters of wood characteristics in hybrid larches. *Annals of Forest Science, 59*(5–6), 607–615.
Argus, G. W. (1966). Botanical investigations in northeastern Saskatchewan: The subarctic Patterson-Hasbala Lakes region. *Canadian Field-Naturalist, 80*, 119–143.
Arregui, A., Espinel, S., Aragones, A., & Sierra, R. (1999). Estimate of genetic parameters in a progeny test of *Pinus radiata* in the Basque Country. *Investigacion Agraria Sistemas y Recursos Forestales, 8*(1), 119–128.
Atzmon, N., Moshe, Y., & Schiller, G. (2004). Ecophysiological response to severe drought in *Pinus halepensis* Mill. trees of two provenances. *Plant Ecology, 171*(1), 15–22.
Avise, J. C. (2009). Phylogeography: Retrospect and prospect. *Journal of Biogeography, 36*(1), 3–15.
Avise, J. C., Arnold, J., Ball, R. M., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A., & Saunders, N. C. (1987). Intraspecific phylogeography: The mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics, 18*(1), 489–522.
Avramidou, E. V., Doulis, A. G., & Aravanopoulos, F. A. (2017). Linkage and QTL mapping in *Cupressus sempervirens* L. provides the first detailed genetic map of the species and identifies a QTL associated with crown form. *Tree Genetics & Genomes, 13*(1), 2.
Axelrod, D. I. (1959). Late tertiary evolution of the Sierran big-tree forest. *Evolution, 13*, 9–23.
Axelrod, D. I. (1964). The Miocene Trapper Creek flora of southern Idaho: University of California Publications in Geological. *Science, 51*, 1–148.

Axelrod, D. I. (1980). *History of the maritime closed-cone pines, Alta and Baja California* (Vol. 120, p. 143). Berkeley: Univ. Calif. Publ. Geol. Sci.

Bagnoli, F., Fady, B., Fineschi, S., Oddou-Muratorio, S., Piotti, A., Sebastiani, F., & Vendramin, G. G. (2011). Neutral patterns of genetic variation and applications to conservation in conifer species. In *Genetics, genomics and breeding of conifers* (pp. 141–195). Boca Raton: CRC Press.

Bahnweg, G., Schubert, R., Kehr, R. D., Müller-Starck, G., Heller, W., Langebartels, C., & Sandermann, H., Jr. (2000). Controlled inoculation of Norway spruce (*Picea abies*) with *Sirococcus conigenus*: PCR-based quantification of the pathogen in host tissue and infection-related increase of phenolic metabolites. *Trees, 14*(8), 435–441.

Bahrman, N., & Petit, R. J. (1995). Genetic polymorphism in maritime pine (*Pinus pinaster* Ait.) assessed by two-dimensional gel electrophoresis of needle, bud, and pollen proteins. *Journal of Molecular Evolution, 41*(2), 231–237.

Bahrman, N., Zivy, M., Damerval, C., & Baradat, P. H. (1994). Organisation of the variability of abundant proteins in seven geographical origins of maritime pine (*Pinus pinaster* Ait.). *Theoretical and Applied Genetics, 88*(3–4), 407–411.

Baker, E. A., Wegrzyn, J. L., Sezen, U. U., Falk, T., Maloney, P. E., Vogler, D. R., et al. (2018). Comparative transcriptomics among four white pine species. *G3: Genes, Genomes Genetics, 8*(5), 1461–1474.

Ballian, D., Longauer, R., Mikić, T., Paule, L., Kajba, D., & Gömöry, D. (2006). Genetic structure of a rare European conifer, Serbian spruce (*Picea omorika* (Panč.) Purk.). *Plant Systematics and Evolution, 260*(1), 53–63.

Bannister, M. (1965). Forest Research Institute, Rotorua, New Zealand. In *The genetics of colonizing species: proceedings* (p. 353). New York: Academic Press.

Bansal, S., Harrington, C. A., Gould, P. J., & St Clair, J. B. (2015a). Climate-related genetic variation in drought-resistance of Douglas-fir (*Pseudotsuga menziesii*). *Global Change Biology, 21*, 947–958.

Bansal, S., Harrington, C. A., & St Clair, J. B. (2016). Tolerance to multiple climate stressors: A case study of Douglas-fir drought and cold hardiness. *Ecology and Evolution, 6*, 2074. https://doi.org/10.1002/ece3.2007.

Bansal, S., St Clair, J. B., Harrington, C. A., & Gould, P. J. (2015b). Impact of climate change on cold hardiness of Douglas-fir (*Pseudotsuga menziesii*): Environmental and genetic considerations. *Global Change Biology, 21*, 3814–3826.

Bao, W., O'Malley, D. M., Whetten, R., & Sederoff, R. R. (1993). A laccase associated with lignification in loblolly pine xylem. *Science, 260*(5108), 672–672.

Barrett, J. W., Beech, R. N., Dancik, B. P., & Strobeck, C. (1994). A genomic clone of a type I cab gene encoding a light harvesting chlorophyll a/b binding protein of photosystem II identified from lodgepole pine. *Genome, 37*(1), 166–172.

Bartholomé, J., Bink, M. C., van Heerwaarden, J., Chancerel, E., Boury, C., Lesur, I., et al. (2016). Linkage and association mapping for two major traits used in the maritime pine breeding program: Height growth and stem straightness. *PLoS One, 11*(11), e0165323.

Bashalkhanov, S., Eckert, A. J., & Rajora, O. P. (2013). Genetic signatures of natural selection in response to air pollution in red spruce (*Picea rubens*, Pinaceae). *Molecular Ecology, 22*(23), 5877–5889.

Bates, C. G. (1930). The frost hardiness of geographic strains of Norway pine. *Journal of Forestry, 28*, 327–333.

Beaulieu, J., Doerksen, T., Boyle, B., Clément, S., Deslauriers, M., Beauseigle, S., et al. (2011). Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression. *Genetics, 188*(1), 197–214.

Beaulieu, J., & Rainville, A. (2005). Adaptation to climate change: Genetic variation is both a short- and a long-term solution. *The Forestry Chronicle, 81*(5), 704–709.

Beaumont, M. A., & Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London B: Biological Sciences, 263*(1377), 1619–1626.

Beck, C. B. (1960). The identity of *Archaeopteris* and *Callixylon*. *Brittonia, 12*, 351–368.
Beck, C. B. (1988). *Origin and evolution of gymnosperms*. New York: Columbia University Press.
Beck, C. B., & Wight, D. C. (1988). Progymnosperms. In C. B. Beck (Ed.), *Origin and evolution of gymnosperms*. New York: Columbia University Press.
Bedon, F., Bomal, C., Caron, S., Levasseur, C., Boyle, B., Mansfield, S. D., et al. (2010). Subgroup 4 R2R3-MYBs in conifer trees: Gene family expansion and contribution to the isoprenoid-and flavonoid-oriented responses. *Journal of Experimental Botany, 61*(14), 3847–3864.
Bedon, F., Grima-Pettenati, J., & Mackay, J. (2007). Conifer R2R3-MYB transcription factors: Sequence analyses and gene expression in wood-forming tissues of white spruce (*Picea glauca*). *BMC Plant Biology, 7*(1), 17.
Bedon, F., Levasseur, C., Grima-Pettenati, J., Séguin, A., & MacKay, J. (2009). Sequence analysis and functional characterization of the promoter of the *Picea glauca* Cinnamyl Alcohol Dehydrogenase gene in transgenic white spruce plants. *Plant Cell Reports, 28*(5), 787–800.
Beier-Petersen, B., Esbjerg, P., & Soeegaard, B. (1974). The response of *Dreyfusia (Adelges) nordmannianae* Eckst. (Hom., Adelgidae) and of *Oligonychus ununguis* Jac. (Acarina, tetranychidae) to individual clones of *Abies alba* (Mill.) and to light and water. *Forest Tree Improvement, Arboretet-Hoersholm, 8*, 33–52.
Bennuah, S. Y., Wang, T., & Aitken, S. N. (2004). Genetic analysis of the *Picea sitchensis* × *P. glauca* introgression zone in British Columbia. *Forest Ecology and Management, 197*(1), 65–77.
Bergin, D. O., & Kimberley, M. O. (1992). Provenance variation in *Podocarpus totara*. *New Zealand Journal of Ecology, 16*(1), 5–13.
Bergin, D. O., Kimberley, M. O., & Low, C. B. (2008). Provenance variation in *Podocarpus totara* (D. Don): Growth, tree form and wood density on a coastal site in the north of natural range, New Zealand. *Forest Ecology and Management, 255*, 1367–1378.
Bergmann, F., Gregorius, H. R., & Larsen, J. B. (1990). Levels of genetic variation in European silver fir (*Abies alba*). *Genetica, 82*(1), 1–10.
Bi, H., Li, R., Wu, Z., Huang, Q., Liu, Q., Zhou, Y., & Li, Y. (2013). Early performance of *Pinus radiata* provenances in the earthquake-ravaged dry river valley area of Sichuan, southwest China. *Journal of Forestry Research, 24*(4), 619–632.
Biffin, E., Brodribb, T. J., Hill, R. S., Thomas, P., & Lowe, A. J. (2012). Leaf evolution in Southern Hemisphere conifers tracks the angiosperm ecological radiation. *Proceedings of the Royal Society B, 279*, 341–348.
Biffin, E., Conran, J., & Lowe, A. (2011). Podocarp evolution: A molecular phylogenetic perspective. *Ecology of the Podocarpaceae in Tropical Forests, 95*, 1–20.
Bigras, F. J., Ryyppö, A., Lindström, A., & Stattin, E. (2001). Cold acclimation and deacclimation of shoots and roots of conifer seedlings. In *Conifer cold hardiness* (pp. 57–88). Dordrecht: Springer.
Bilgen, B., & Kaya, N. (2007). Allozyme variations in six natural populations of scots pine (*Pinus sylvestris*) in Turkey. *Biologia, 62*(6), 697–703.
Bingham, R. T. (1983). Blister rust resistant western white pine for the Inland Empire: The story of the first 25 years of the research and development program. USDA Forest Service General Technical Report INT-146.
Bingham, R. T., Olson, R. J., Backer, W. A., & Marsden, M. A. (1969). Breeding blister rust resistant white pine. V. Estimates of heritability, combining ability, and genetic advance based on tester matings. *Silvae Genetica, 18*, 23–38.
Birol, I., Raymond, A., Jackman, S. D., Pleasance, S., Coope, R., Taylor, G. A., et al. (2013). Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data. *Bioinformatics, 29*(12), 1492–1497.
Blada, I. (2000). Genetic variation in blister rust resistance and growth traits in *Pinus strobus* x *P. peuce* hybrid at age 17 (experiment 1). *Forest Genetics, 7*(2), 109–120.
Blada, I., & Popescu, F. (2007). Swiss stone pine provenance experiment in Romania: II Variation in growth and branching traits to age 14. *Silvae Genetica, 56*(3–4), 148–158.

Blodgett, J. T., Eyles, A., & Bonello, P. (2007). Organ-dependent induction of systemic resistance and systemic susceptibility in *Pinus nigra* inoculated with *Sphaeropsis sapinea* and *Diplodia scrobiculata*. *Tree Physiology, 27*(4), 511–517.

Blodgett, J. T., & Stanosz, G. R. (1997). Differential inhibition of *Sphaeropsis sapinea* morphotypes by a phenolic compound and several monoterpenes of red pine. *Phytopathology, 87*(6), 606–609.

Blödner, C., Majcherczyk, A., Kües, U., & Polle, A. (2007). Early drought-induced changes to the needle proteome of Norway spruce. *Tree Physiology, 27*(10), 1423–1431.

Blouin, D., Beaulieu, J., Daoust, G., & Poliquin, J. (1994). Wood quality of Norway spruce grown in plantations in Quebec. *Wood and Fiber Science, 26*(3), 342–353.

Bobola, M. S., Hillenberg, K. A., Gendreau, S. B., Eckert, R. T., Klein, A. S., & Stapelfeldt, K. (1996). Hybridization between *Picea rubens* and *Picea mariana*: Differences observed between montane and coastal island populations. *Canadian Journal of Forest Research, 26*(3), 444–452.

Bobola, M. S., Smith, D. E., & Klein, A. S. (1992). Five major nuclear ribosomal repeats represent a large and variable fraction of the genomic DNA of *Picea rubens* and *P. mariana*. *Molecular Biology and Evolution, 9*(1), 125–137.

Bobowicz, M. A., & Danielewicz, W. (2000). Isoenzymatic variability in progeny of *Pinus mugo* Turra x *Pinus sylvestris* L. hybrids from Bór na Czerwonem, in experimental culture. *Acta Societatis Botanicorum Poloniae, 69*(2), 137–144.

Bodare, S., Stocks, M., Yang, J. C., & Lascoux, M. (2013). Origin and demographic history of the endemic Taiwan spruce (*Picea morrisonicola*). *Ecology and Evolution, 3*(10), 3320–3333.

Boerjan, W., Ralph, J., & Baucher, M. (2003). Lignin biosynthesis. *Annual Review of Plant Biology, 54*(1), 519–549.

Bogunic, F., Muratovic, E., & Siljak-Yakovlev, S. (2006). Chromosomal differentiation between *Pinus heldreichii* and *Pinus nigra*. *Annals of Forest Science, 63*(3), 267–274.

Bohlmann, J., Crock, J., Jetter, R., & Croteau, R. (1998a). cDNA cloning, characterization, and functional expression of wound-inducible (*E*)-α-bisabolene synthase from grand fir (*Abies grandis*). *Proceedings of the National Academy of Sciences USA, 95*, 6756–6761.

Bohlmann, J., Meyer-Gauen, G., & Croteau, R. (1998b). Plant terpenoid synthases: Molecular biology and phylogenetic analysis. *Proceedings of the National Academy of Sciences, 95*(8), 4126–4133.

Bohlmann, J., Phillips, M., Ramachandiran, V., Katoh, S., & Croteau, R. (1999). cDNA cloning, characterization, and functional expression of four new monoterpene synthase members of the Tpsd gene family from grand fir (*Abies grandis*). *Archives of Biochemistry and Biophysics, 368*(2), 232–243.

Bohlmann, J., Steele, C. L., & Croteau, R. (1997). Monoterpene synthases from grand fir (*Abies grandis*) cDNA isolation, characterization, and functional expression of myrcene synthase, (−)-(4S)-limonene synthase, and (−)-(1S, 5S)-pinene synthase. *Journal of Biological Chemistry, 272*(35), 21784–21792.

Bonello, P., Gordon, T. R., & Storer, A. J. (2001). Systemic induced resistance in Monterey pine. *Forest Pathology, 31*, 99–106.

Bongarten, B. C., Wheeler, N. C., & Jech, K. S. (1985). Isozyme heterozygosity as a selection criterion for yield improvement in Douglas-fir. In *New ways in forest genetics* (pp. 19–22). Proceedings of the Canadian Tree Improvement Association, Quebec City, Quebec.

Boscherini, G., Morgante, M., Rossi, P., & Vendramin, G. G. (1994). Allozyme and chloroplast DNA variation in Italian and Greek populations of *Pinus leucodermis*. *Heredity, 73*(3), 284–290.

Bouillé, M., Senneville, S., & Bousquet, J. (2011). Discordant mtDNA and cpDNA phylogenies indicate geographic speciation and reticulation as driving factors for the diversification of the genus *Picea*. *Tree Genetics & Genomes, 7*(3), 469–484.

Bowe, L. M., Coat, G., & dePamphilis, C. W. (2000). Phylogewny of seed plants based on all three genomic compartments: Extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. *Proceedings of the National Academy of Sciences, 97*(8), 4092–4097.

Bowen, M. R., & Whitmore, T. C. (1980). *Agathis* - a genus of fast growing rain forest conifers. *The Commonwealth Forestry Review, 59*(3), 307–310.

Bower, A. D., & Aitken, S. N. (2006). Geographic and seasonal variation in cold hardiness of whitebark pine. *Canadian Journal of Forest Research, 36*(7), 1842–1850.

Bower, A. D., & Aitken, S. N. (2007). Mating system and inbreeding depression in whitebark pine (*Pinus albicaulis* Engelm.). *Tree Genetics & Genomes, 3*(4), 379–388.

Bower, A. D., & Aitken, S. N. (2008). Ecological genetics and seed transfer guidelines for *Pinus albicaulis* (Pinaceae). *American Journal of Botany, 95*(1), 66–76.

Boyer, M. G. (1964). Studies on white pine phenols in relation to blister rust. *Canadian Journal of Botany, 42*(8), 979–987.

Boyer, M. G., & Isaac, P. K. (1964). Some observations on white pine blister rust as compared by light and electron microscopy. *Canadian Journal of Botany, 42*(9), 1305–1309.

Boys, J., Cherry, M., & Dayanandan, S. (2005). Microsatellite analysis reveals genetically distinct populations of red pine (*Pinus resinosa*, Pinaceae). *American Journal of Botany, 92*(5), 833–841.

Bragg, J. G., Supple, M. A., Andrew, R. L., & Borevitz, J. O. (2015). Genomic variation across landscapes: Insights and applications. *New Phytologist, 207*(4), 953–967.

Brandes, A., Heslop-Harrison, J. S., Kamm, A., Kubis, S., Doudrick, R. L., & Schmidt, T. (1997). Comparative analysis of the chromosomal and genomic organization of Ty1-copia-like retrotransposons in pteridophytes, gymnosperms and angiosperms. *Plant Molecular Biology, 33*(1), 11–21.

Braukmann, T. W. A., Kuzmina, M., & Stefanovic, S. (2009). Loss of all plastid nhd genes in Gnetales and conifers: Extent and evolutionary significance for the seed plant phylogeny. *Current Genetics, 55*(3), 323–337.

Breidenstein, J., Bastien, J.-C., & Roman-Amat, B. (1990, August 20–24). Douglas-fir range-wide variation results from the IUFRO data base. P. 2.13-1.26. In *Proceedings of joint meeting Western Forest Genetics Association and IUFRO Working Parties S2.02-05, 06,12 and 14 – Douglas-fir, contorta pine, Sitka spruce, and Abies breeding and genetic resources, Olympia, WA*. Weyerhaeuser Co., Federal Way, WA.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.

Brendel, O., Pot, D., Plomion, C., Rozenberg, P., & Guehl, J. M. (2002). Genetic parameters and QTL analysis of δ13C and ring width in maritime pine. *Plant, Cell & Environment, 25*(8), 945–953.

Bridgwater, F. E., Nelson, C. D., Byram, T. D., Kubisiak, T. L., Young, C. H., & Doudrick, R. L. (2003). Development of differential screening panels for slash pine-fusiform rust reaction types. In C. R. McKinley (Ed.), *Proceedings 27th Southern Forest Tree Improvement Conference*, Stillwater, Oklahoma, USA, June 24–27, 2003, p. 214.

Britten, R. J., & Kohne, D. E. (1968). Repeated sequences in DNA. *Science, 161*, 529–540.

Brown, G. R., Amarasinghe, V., Kiss, G., & Carlson, J. E. (1993). Preliminary karyotype and chromosomal localization of ribosomal DNA sites in white spruce using fluorescence in situ hybridization. *Genome, 36*(2), 310–316.

Brown, G. R., Bassoni, D. L., Gill, G. P., Fontana, J. R., Wheeler, N. C., Megraw, R. A., et al. (2003). Identification of quantitative trait loci influencing wood property traits in loblolly pine (*Pinus taeda* L.). III. QTL verification and candidate gene mapping. *Genetics, 164*(4), 1537–1546.

Brown, G. R., & Carlson, J. E. (1997). Molecular cytogenetics of the genes encoding 18S-5.8 S-26S rRNA and 5S rRNA in two species of spruce (*Picea*). *Theoretical and Applied Genetics, 95*(1–2), 1–9.

Brown, G. R., Gill, G. P., Kuntz, R. J., Langley, C. H., & Neale, D. B. (2004). Nucleotide diversity and linkage disequilibrium in loblolly pine. *Proceedings of the National Academy of Sciences of the United States of America, 101*(42), 15255–15260.

Brown, G. R., Kadel, E. E., Bassoni, D. L., Kiehne, K. L., Temesgen, B., Van Buijtenen, J. P., et al. (2001). Anchored reference loci in loblolly pine (*Pinus taeda* L.) for integrating pine genomics. *Genetics, 159*(2), 799–809.

Brown, P. M. (2013). Rocky Mountain tree-ring research, OldList, a database of old trees. http://www.rmtrr.org/oldlist.htm. Accessed 21 Oct 2017.

Bruederle, L. P., Tomback, D. F., Kelly, K. K., & Hardwick, R. C. (1998). Population genetic structure in a bird-dispersed pine, *Pinus albicaulis* (Pinaceae). *Canadian Journal of Botany, 76*(1), 83–90.

Brundrett, M. C. (2008). Evolution of mycorrhizas. In *Mycorrhizal associations: The web resource.* http://mycorrhizas.info/evol.html. Accessed 21 Oct 2017.

Brunsfeld, S. J., Soltis, P. S., Soltis, D. E., Gadek, P. A., Quinn, C. J., Strenge, D. D., & Ranker, T. A. (1994). Phylogenetic relationships among the genera of Taxodiaceae and Cupressaceae: Evidence from *rbcL* sequences. *Systematic Botany, 19*(2), 253–262.

Bucci, G., Vendramin, G. G., Lelli, L., & Vicario, F. (1997). Assessing the genetic divergence of *Pinus leucodermis* Ant. endangered populations: Use of molecular markers for conservation purposes. *Theoretical and Applied Genetics, 95*(7), 1138–1146.

Buchert, G. P., Rajora, O. P., Hood, J. V., & Dancik, B. P. (1997). Effects of harvesting on genetic diversity in old-growth eastern white pine in Ontario, Canada. *Conservation Biology, 11*(3), 747–758.

Budde, K. B., Heuertz, M., Hernández-Serrano, A., Pausas, J. G., Vendramin, G. G., Verdú, M., & González-Martínez, S. C. (2014). In situ genetic association for serotiny, a fire-related trait, in Mediterranean maritime pine (*Pinus pinaster*). *New Phytologist, 201*(1), 230–241.

Burczyk, J., DiFazio, S. P., & Adams, W. T. (2004b). Gene flow in forest trees: How far do genes really travel? *Forest Genetics, 11*(3/4), 179.

Burczyk, J., Lewandowski, A., & Chalupka, W. (2004a). Local pollen dispersal and distant gene flow in Norway spruce (Picea abies [L.] Karst.). *Forest Ecology and Management, 197*(1), 39–48.

Burdon, R. D. (1998). Multi-site provenance trials of *Pinus radiata*. *Forest Genetic Resources, 26*, 3–8.

Burdon, R. D., & Bannister, M. H. (1973). Provenances of *Pinus radiata*: Their early performance and silvicultural potential. *New Zealand Journal of Forestry, 18*(2), 217–232.

Burdon, R. D., Currie, D., & Chou, C. K. S. (1982). Responses to inoculation with *Diplodia pinea* in progenies of apparently resistant trees of Pinus radiata. *Australasian Plant Pathology, 11*(4), 37–39.

Burdon, R. D., Firth, A., Low, C. B., & Miller, M. A. (1997). Native provenances of *Pinus radiata* in New Zealand: Performance and potential. *New Zealand Journal of Forestry, 41*, 32–36.

Burdon, R. D., & Low, C. B. (1991). Performance of *Pinus ponderosa* and *Pinus jeffreyi* provenances in New Zealand. *Canadian Journal of Forest Research, 21*, 1401–1414.

Burleigh, J. G., & Mathews, S. (2007). Assessing among-locus variation in the inference of seed plant phylogeny. *International Journal of Plant Sciences, 168(2)*, 111–124.

Burley, J., & Nikles, D. G. (1973). Tropical provenance and progeny research and international cooperation. In *Proceedings of joint workshop IUFRO in Nairobi, Kenya*. Oxford, UK: Commonwealth Forestry Institute.

Businge, E., Brackmann, K., Moritz, T., & Egertsdotter, U. (2012). Metabolite profiling reveals clear metabolic changes during somatic embryo development of Norway spruce (*Picea abies*). *Tree Physiology, 32*(2), 232–244.

Businge, E., Bygdell, J., Wingsle, G., Moritz, T., & Egertsdotter, U. (2013). The effect of carbohydrates and osmoticum on storage reserve accumulation and germination of Norway spruce somatic embryos. *Physiologia Plantarum, 149*(2), 273–285.

Byun, K. O., Kim, M. Z. S., Shim, S. Y., Hong, S. H., & Sohn, S. I. (1989). Review of pitch-loblolly hybrid pine (*Pinus rigida* and *P. taeda*) breeding researches in Korea and future strategy. *Res Rep, For Genet Res Inst, Korea, 25*, 204–211.

Cai, Q., Zhang, D., Liu, Z. L., & Wang, X. R. (2006). Chromosomal localization of 5S and 18S rDNA in five species of subgenus *Strobus* and their implications for genome evolution of *Pinus*. *Annals of Botany, 97*(5), 715–722.

Cairney, J., Zheng, L., Cowels, A., Hsiao, J., Zismann, V., Liu, J., et al. (2006). Expressed sequence tags from loblolly pine embryos reveal similarities with angiosperm embryogenesis. *Plant Molecular Biology, 62*(4–5), 485–501.

Ćalić, I., Bussotti, F., Martínez-García, P. J., & Neale, D. B. (2016). Recent landscape genomics studies in forest trees—What can we believe? *Tree Genetics & Genomes, 12*(1), 3.

Camcore. (2006). *2006 Camcore annual report. Camcore international tree conservative and domestication*. Raleigh, NC: Department of Forestry and Environmental Resources, North Carolina State University.

CAMCORE Cooperative. (2000). *Conservation and testing of tropical and subtropical forest tree species by the CAMCORE Cooperative*. Raleigh: College of Natural Resources, NCSU. http://www.camcore.org/index.php.

Campbell, F. T., & Schlarbaum, S. E. (1994). Fading forests. North American trees and the threat of exotic pests. Natural Resources Defense Council report, 47 pp. Available online at http://treeimprovement.utk.edu/FFI.htm. Last accessed 16 Sept 2014.

Campbell, F. T., & Schlarbaum S. E. (2002). Fading forests II. Trading Away North America's Natural Heritage? Healing Stones Foundation Publ., 128 pp. Print and Compact Disc. Available online at http://treeimprovement.utk.edu/FFII.htm. Last accessed 16 Sept 2014.

Campbell, F. T., & Schlarbaum S. E. (2014). Fading Forests III. American Forests. What choice will we make? The nature conservancy and the University of Tennessee, 155 pp. Available online at http://treeimprovement.utk.edu/FFII.htm. Last accessed 16 Sept 2014.

Campbell, R. K. (1979). Genecology of Douglas-fir in a watershed in the Oregon Cascades. *Ecology, 60*(5), 1036–1050.

Campbell, R. K. (1986). Mapped genetic variation of Douglas-fir to guide seed transfer in southwest Oregon. *Silvae Genetica, 35*(2–3), 85–96.

Campbell, R. K. (1991). Soils, seed-zone maps, and physiography: Guidelines for seed transfer of Douglas-fir in Southwestern Oregon. *Forest Science, 37*(4), 973–986.

Campbell, R. K., & Sorensen, F. C. (1978). Effect of test environment on expression of clines and on delineation of seed zones in Douglas-fir. *Theoretical and Applied Genetics, 51*, 223–246.

Campbell, R. K., & Sugano, A. I. (1979). Genecology of bud-burst phenology in Douglas-fir: Response to flushing temperature and chilling. *Botanical Gazette, 140*, 223–231.

Campbell, R. K., & Sugano, A. I. (1993). Genetic variation and seed zones of Douglas-fir in the Siskiyou National Forest. USDA Forest Service research paper PNW-RP-461.

Canales, J., Bautista, R., Label, P., Gómez-Maldonado, J., Lesur, I., Fernández-Pozo, N., et al. (2014). De novo assembly of maritime pine transcriptome: Implications for forest breeding and biotechnology. *Plant Biotechnology Journal, 12*(3), 286–299.

Cannell, M. G. R., & Sheppard, L. J. (1982). Seasonal changes in the frost hardiness of provenances of *Picea sitchensis* in Scotland. *Forestry, 55*(2), 137–153.

Carlsbecker, A., Sundström, J., Tandre, K., Englund, M., Kvarnheden, A., Johanson, U., & Engström, P. (2003). The DAL10 gene from Norway spruce (*Picea abies*) belongs to a potentially gymnosperm-specific subclass of MADS-box genes and is specifically active in seed cones and pollen cones. *Evolution & Development, 5*(6), 551–561.

Carson, S. D. (1990). A breed of radiata pine resistant to *Dothistroma* needle blight - methods used for development and expected gains. *Phytopathology, 80*, 1007. (abstr).

Carson, S. D., & Carson, M. J. (1989). Breeding for resistance in forest trees-a quantitative genetic approach. *Annual Review of Phytopathology, 27*(1), 373–395.

Carson, S. D., & Carson, M. J. (1991). Realising gains in resistance to *Dothistroma*. In C. Dean & C. Hanel. *Proceedings of the 11th Research Working Group 1 (Forest Genetics)*, Australian Forestry Council, 11–15 March, 1991, Coonawarra, South Australia, pp. 71–75.

Castellanos-Acuna, D., Saenz-Romero, C., Lindig-Cisneros, R. A., Sanchez-Vargas, N. M., Lobbit, P., & Montero-Castro, J. C. (2013). Revista Chapingo Serie Ciencias Forestales y del Ambiente. doi: https://doi.org/10.5154/r.rchscfa.2013.01.002. http://www.chapingo.mx/revistas

Castresana, J. (2007). Topological variation in single-gene phylogenetic trees. *Genome Biology., 8*(6), 216. https://doi.org/10.1186/gb-2007-8-6-216.

Chagné, D., Brown, G., Lalanne, C., Madur, D., Pot, D., Neale, D., & Plomion, C. (2003). Comparative genome and QTL mapping between maritime and loblolly pines. *Molecular Breeding, 12*(3), 185–195.

Chagné, D., Lalanne, C., Madur, D., Kumar, S., Frigério, J. M., Krier, C., et al. (2002). A high density genetic map of maritime pine based on AFLPs. *Annals of Forest Science, 59*(5–6), 627–636.

Chaisurisri, K., & El-Kassaby, Y. A. (1994). Genetic diversity in a seed production population vs. natural populations of Sitka spruce. *Biodiversity and Conservation, 3*(6), 512–523.

Chalupka, W., Mejnartowicz, L., & Lewandoswski, A. (2008). Reconstitution of a lost forest tree population: A case study of Norway spruce *(Picea abies* [L.] Karst.). *Forest Ecology and Management, 255*, 2103–2108.

Chancerel, E., Lamy, J. B., Lesur, I., Noirot, C., Klopp, C., Ehrenmann, F., et al. (2013). High-density linkage mapping in a pine tree reveals a genomic region associated with inbreeding depression and provides clues to the extent and distribution of meiotic recombination. *BMC Biology, 11*(1), 50.

Chancerel, E., Lepoittevin, C., Le Provost, G., Lin, Y. C., Jaramillo-Correa, J. P., Eckert, A. J., et al. (2011). Development and implementation of a highly-multiplexed SNP array for genetic mapping in maritime pine and comparative mapping with loblolly pine. *BMC Genomics, 12*(1), 368.

Chang, S., Puryear, J., Funkhouser, E. A., Newton, R. J., & Cairney, J. (1996). Cloning of a cDNA for a chitinase homologue which lacks chitin-binding sites and is down-regulated by water stress and wounding. *Plant Molecular Biology, 31*(3), 693–699.

Chase, M. W., Soltis, D. E., Olmstead, R. G., Morgan, D., Les, D. H., Mishler, B. D., et al. (1993). Phylogenetics of seed plants: An analysis of nucleic sequences from the plastid gene *rbcL. Annals of the Missouri Botanical Garden, 80*, 528–580.

Chaw, S.-M., Parkinson, C. L., Cheng, Y., Vincent, T. M., & Palmer, J. D. (2000). Seed plant phylogeny inferred from all three plant genomes: Monophyly of extant gymnosperms and origin of Gnetales from conifers. *Proceedings- National Academy of Sciences USA, 97*(8), 4086–4091.

Chaw, S. M., Zharkikh, A., Sung, H. M., Lau, T. C., & Li, W. H. (1997). Molecular phylogeny of extant gymnosperms and seed plant evolution: Analysis of nuclear 18S rRNA sequences. *Molecular biology and evolution, 14*, 56–68.

Chen, C., Durand, E., Forbes, F., & François, O. (2007). Bayesian clustering algorithms ascertaining spatial population structure: A new computer program and a comparison study. *Molecular Ecology Notes, 7*(5), 747–756.

Chen, J., Hao, Z., Xu, H., Yang, L., Liu, G., Sheng, Y., Zheng, C., et al. (2015). The complete chloroplast genome sequence of the relict woody plant *Metasequoia glyptostroboides* Hu et Cheng. *Frontiers in Plant Science, 6*, 447. https://doi.org/10.3389/fpls.2015.00447.

Chen, J., Källman, T., Ma, X., Gyllenstrand, N., Zaina, G., Morgante, M., et al. (2012b). Disentangling the roles of history and local selection in shaping clinal variation of allele frequencies and gene expression in Norway spruce (*Picea abies*). *Genetics, 191*(3), 865–881.

Chen, J., Tsuda, Y., Stocks, M., Källman, T., Xu, N., Kärkkäinen, K., et al. (2014). Clinal variation at phenology-related genes in spruce: Parallel evolution in FTL2 and *Gigantea*? *Genetics, 197*(3), 1025–1038.

Chen, J., Uebbing, S., Gyllenstrand, N., Lagercrantz, U., Lascoux, M., & Källman, T. (2012a). Sequencing of the needle transcriptome from Norway spruce (*Picea abies* Karst L.) reveals lower substitution rates, but similar selective constraints in gymnosperms and angiosperms. *BMC Genomics, 13*(1), 589.

Chen, T., Wu, X., Chen, Y., Li, X., Huang, M., Zheng, M., et al. (2009). Combined proteomic and cytological analysis of Ca^{2+}−calmodulin regulation in *Picea meyeri* pollen tube growth. *Plant Physiology, 149*(2), 1111–1126.

Chen, Y., Chen, T., Shen, S., Zheng, M., Guo, Y., Lin, J., et al. (2006). Differential display proteomic analysis of *Picea meyeri* pollen germination and pollen-tube growth after inhibition of actin polymerization by latrunculin B. *The Plant Journal, 47*(2), 174–195.

Cheng, Y., Nicolson, R. G., Tripp, K., & Chaw, S.-J. (2000). Phylogeny of Taxaceae and Cephalotaxaceae genera inferred from chloroplast *mat*K gene and nuclear rDNA ITS region. *Molecular Phylogenetics and Evolution., 14*(3), 353–365.

Chhatre, V. E., Byram, T. D., Neale, D. B., Wegrzyn, J. L., & Krutovsky, K. V. (2013). Genetic structure and association mapping of adaptive and selective traits in the East Texas loblolly pine (*Pinus taeda* L.) breeding populations. *Tree Genetics & Genomes, 9*(5), 1161–1178.

Chiba, S. (1963). Studies on the breeding of *Larix* species 1. Variation between species on the gnawing by the mountain hares and the wild mice. Technical note, Institute for Forest Tree Improvement, Oji Paper Co., Ltd., No. 29, Reprinted from Trans. Mtg. Hokkaido Br. Jpn. For. Soc. 12th 1963.

Chiba, S., & Nagata, Y. (1963). Studies on the breeding of Larix species 2. Variation between species in susceptibility by *Mycosphaerella larici-leptolepis* and *Pysalospora laricina*. Technical note, Institute for Forest Tree Improvement, Oji Paper Co., Ltd., No. 30, Reprinted from Trans. Mtg. Hokkaido Br. Jpn. For. Soc. 12th 1963.

Chiba, S., & Nagata, Y. (1972). Studies on the breeding of *Larix* species 4. Selection of resistant trees for needle cast and shoot blight in Japanese larch, and observation on its growth. Technical note, Institute for Forest Tree Improvement, Oji Paper Co. Ltd., No. 116, Reprinted from Forest Tree Breeding of Hokkaido 15(1) 1972.

Chiba, S., & Nagata, Y. (1978). Clonal testing of selected trees, resistant for needle cast and shoot blight in *Larix* species. Technical note, Institute for Forest Tree Improvement, Oji Paper Co., Ltd. No. 174, Reprinted from Forest Tree Breeding of Hokkaido 21 1978.

Chiba, S., Nagata, Y., & Tomaki, K. (1982). Variation of vole resistance in Larix leptolepis, and comparison with the hybrid of *L. gmelini and L. leptolepis.* Technical note, Institute for Forest Tree Improvement, Oji Paper Co., Ltd., No. 209, Reprinted from Forest Tree Breeding of Hokkaido 25(2) 1982.

Chiba, S., Nagata, Y., & Tomaki, K. (1989). Palatability test of vole from several localities using same clonal materials in *Larix* species. Technical note, Institute for Forest Tree Improvement, Oji Paper Co., Ltd., No. 268, Reprinted from Forest Tree Breeding of Hokkaido 32(2) 1989.

Ching, K. K. (1965). *Early growth of Douglas-fir in a reciprocal planting* (Forest Research Laboratory research paper 3). Corvallis: Oregon State University.

Christenhusz, M. J. M., Reveal, J. L., Farjon, A., Gardner, M. F., Mill, R. R., & Chase, M. W. (2011). A new classification and linear sequence of extant gymnosperms. *Phytotaxa, 19*, 55–70.

Christensen, K. I. (1987). A morphometric study of the *Pinus mugo* Turra complex and its natural hybridization with *P. sylvestris* L.(Pinaceae). *Feddes Repertorium, 98*(11–12), 623–635.

Chung, M. G., Oh, G. S., & Chung, J. M. (1999). Allozyme variation in Korean populations of *Taxus cuspidata* (Taxaceae). *Scandinavian Journal of Forest Research, 14*(2), 103–110.

Chybicki, I. J., Oleksa, A., & Kowalkowska, K. (2012). Variable rates of random genetic drift in protected populations of English yew: Implications for gene pool conservation. *Conservation Genetics, 13*(4), 899–911.

Cinget, B., Lafontaine, G., Gérardi, S., & Bousquet, J. (2015). Integrating phylogeography and paleoecology to investigate the origin and dynamics of hybrid zones: Insights from two widespread North American firs. *Molecular Ecology, 24*(11), 2856–2870.

Clapham, D. H., Kolukisaoglu, H. Ü., Larsson, C. T., Qamaruddin, M., Ekberg, I., Wiegmann-Eirund, C., et al. (1999). Phytochrome types in *Picea* and *Pinus*. Expression patterns of PHYA-related types. *Plant Molecular Biology, 40*(4), 669–678.

Clark, A. G., et al. (2007). Evolution of genes and genomes on the Drosophila phylogeny. *Nature, 450*(7167), 203–218.

Clausen, J., Keck, D. D., & Hiesey, W. M. (1940). *Experimental studies on the nature of species. 1. The effect of varied environments on western North American Plants* (Carnegie Institution of Washington publication no. 520). Washington, D.C.: Carnegie Institution of Washington.

Cobb, F. W., & Libby, W. J. (1968). Susceptibility of Monterey, Guadalupe Island, Cedros Island, and bishop pines to Scirrhia (*Dothistroma*) pini, the cause of red band needle blight. *Phytophathology, 58*(1), 88–90.

Conkle, M. T. (1981, April). Isozyme variation and linkage in six conifer species. In Conkle MT, technical coordinator. Proc. Symp. Isozymes North Am. Forest Trees and Forest Inspects. Gen. Tech. REP. PSW-48. Berkeley, California: Pacific SW Forest and Range Exp. Sta (pp. 11–17).

Conkle, M. T., & Critchfield, W. B. (1988). Genetic variation and hybridization of ponderosa pine. In D. M. Baumgarner & J. E. Lotan (Eds.), *Proceedings of ponderosa pine – The Species and its management symposium* (pp. 27–43). Pullman, WA: Washington State University.

Conran, J. G., Wood, G. G., Martin, P. G., Dowd, J. M., Quinn, C. J., Gadek, P. A., & Price, R. A. (2000). Generic relationships within and between the gymnosperm families Podocarpacaeae and Phyllocladaceae based on an analysis of the chloroplast gene *rbc*L. *Australian Journal of Botany, 48*, 715–724.

Coomes, D. A., Allen, R. B., Bentley, W. A., Burrows, L. E., Canham, C. D., Fagan, L., Forsyth, D. M., Gaxiola-Alcantar, A., Parfitt, R. L., Ruscoe, W. A., Wardle, D. A., Wilson, D. J., & Wright, E. F. (2005). The hare, the tortoise and the crocodile: The ecology of angiosperm dominance, conifer persistence and fern filtering. *Journal of Ecology, 93*, 918–935.

Coop, G., Witonsky, D., Di Rienzo, A., & Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics, 185*(4), 1411–1423.

Cordoba-Rodriguez, D., Vargas-Hernandez, J. J., Lopez-Upton, J., & Munoz-Orozco, A. (2011). Root growth in young plants of *Pinus pinceana* Gordon in response to soil moisture. *Agrociencia, 45*, 493–506.

Correla, I., Alia, R., Yan, W., David, T., Aguiar, A., & Almeida, M. H. (2010). Genotype environment interactions in *Pinus pinaster* at age 10 in a multi-environment trial in Portugal: A maximum likelihood approach. *Annals of Forest Science, 67*(6), 612. https://doi.org/10.1051/forest/2010025. https://hal.archives-ouvertes.fr/hal-00883587.

Costa e Silva, J., & Graudal, L. (2008). Evaluation of an international series of *Pinus kesiya* provenance trials for growth and wood quality traits. *Forest Ecology and Management, 255*, 3477–3488.

Costa, P., Bahrman, N., Frigerio, J. M., Kremer, A., & Plomion, C. (1998). Water-deficit-responsive proteins in maritime pine. *Plant Molecular Biology, 38*(4), 587–596.

Costa, P., Pionneau, C., Bauw, G., Dubos, C., Bahrmann, N., Kremer, A., et al. (1999). Separation and characterization of needle and xylem maritime pine proteins. *Electrophoresis, 20*, 1098–1108.

Costa, P., & Plomion, C. (1999). Genetic analysis of needle proteins in maritime pine 2. Variation of protein accumulation. *Silvae Genetica, 48*(3), 146–150.

Costa, P., Pot, D., Dubos, C., Frigerio, J. M., Pionneau, C., Bodenes, C., et al. (2000). A genetic map of Maritime pine based on AFLP, RAPD and protein markers. *Theoretical and Applied Genetics, 100*(1), 39–48.

Côté, C. L., Boileau, F., Roy, V., Ouellet, M., Levasseur, C., Morency, M. J., et al. (2010). Gene family structure, expression and functional analysis of HD-Zip III genes in angiosperm and gymnosperm forest trees. *BMC Plant Biology, 10*(1), 273.

Court-Picon, M., Gadbin-Henry, C., Guibal, F., & Roux, M. (2004). Dendrometry and morphometry of *Pinus pinea* L. in lower Provence (France): Adaptability and variability of provenances. *Forest Ecology and Management, 194*, 319–333.

Cregg, B. M. (1994). Carbon allocation, gas exchange, and needle morphology of *Pinus ponderosa* genotypes known to differ in growth and survival under imposed drought. *Tree Physiology, 14*(7), 883–898.

Cregg, B. M., Olivas-Garcia, J. M., & Hennessey, T. C. (2000). Provenance variation in carbon isotope discrimination of mature ponderosa pine trees at two locations in the Great Plains. *Canadian Journal of Forest Research, 30*, 428–439.

Critchfield, W. B. (1957). *Geographic variation in Pinus contorta* (Maria Moors Cabot Foundation, Publication Number 3). Cambridge, MA: Harvard University. 118 p.

Critchfield, W. B. (1967). Crossability and relationships of the closed-cone pines. *Silvae Genetica, 16*(3), 89–97.

Critchfield, W. B. (1977). Hybridization of foxtail and bristlecone pines. *Madrono, 24*(4), 193–212.

Critchfield, W. B. (1980). *The genetics of lodgepole pine*. Research Paper, USDA Forest Service, Washington, DC, (WO-37).

Critchfield, W. B. (1984a, August 30). Impact of the Pleistocene on the genetic structure of North American conifers. In 8th *North American forest biology workshop*. Logan, Utah (Edited and Compiled by R. Lanner. pp. 70–118).

Critchfield, W. B. (1984b). Crossability and relationships of Washoe pine. *Madroño, 31*(3), 144–170.

Critchfield, W. B. (1985). The late quaternary history of lodgepole and jack pines. *Canadian Journal of Forest Research, 15*(5), 749–772.

Critchfield, W. B. (1986). Hybridization and classification of the white pines (*Pinus* section *Strobus*). *Taxon, 35*, 647–656.

Critchfield, W. B. (1988). Hybridization of the California firs. *Forest Science, 34*(1), 139–151.

Critchfield, W. B. & Krugman, S. L. (1967). Crossing the western pines at Placerville, California. University of Washington Arboretum Bulletin, Seattle, Volume XXX No. 4, Winter 1967, pp. 78–81.

Critchfield, W. B., & Little, E. L., Jr. (1966). *Geographic distribution of the pines of the world (no. 991)*. Washington, D.C.: US Department of Agriculture, Forest Service.

Cronn, R., Dolan, P. C., Jogdeo, S., Wegrzyn, J. L., Neale, D. B., Clair, J. S., & Denver, D. R. (2017). Transcription through the eye of a needle: Daily and annual cycles of gene expression variation in Douglas-fir needles. *bioRxiv*, 117374.

Cronn, R., Liston, A., Parks, M., Gernandt, D. S., Shen, R., & Mockler, T. (2008). Multiplex sequencing of plant chloroplast genomes using Solexa sequencing-by-synthesis technology. *Nucleic Acids Research, 36*(19), e122.

Cronn, R., & Wendel, J. F. (2004). Cryptic trysts, genomic mergers, and plant speciation. *New Phytologist, 161*(1), 133–142.

Cronquist, A., Takhtajan, A., & Zimmermann, W. (1966). On the higher taxa of Embryobionta. *Taxon, 15*(4), 129–134.

Cruickshank, M. G., Lejour, D., & Morrison, D. J. (2006). Traumatic resin canals as markers of infection events in Douglas-fir roots infected with *Armillaria* root disease. *Forest Pathology, 36*(5), 372–384.

Cruz-Nicolás, J., Vargas-Hernández, J. J., Ramírez-Vallejo, P., & López-Upton, J. (2011). Diversidad genética y diferenciación de las poblaciones de *Pseudotsuga menziesii* (Mirb.) Franco en México. *Revista Fitotecnia Mexicana, 34*(4), 233–240.

Cubbage, F. W., Pye, J. M., Holmes, T. P., & Wagner, J. E. (2000). An economic evaluation of fusiform rust research. *Southern Journal of Applied Forestry, 24*, 77–85.

Cuenca, A., Escalante, A. E., & Piñero, D. (2003). Long-distance colonization, isolation by distance, and historical demography in a relictual Mexican pinyon pine (*Pinus nelsonii* Shaw) as revealed by paternally inherited genetic markers (cpSSRs). *Molecular Ecology, 12*(8), 2087–2097.

Cullingham, C. I., James, P., Cooke, J. E., & Coltman, D. W. (2012). Characterizing the physical and genetic structure of the lodgepole pine × jack pine hybrid zone: Mosaic structure and differential introgression. *Evolutionary Applications, 5*(8), 879–891.

Cumbie, W. P., Eckert, A., Wegrzyn, J., Whetten, R., Neale, D., & Goldfarb, B. (2011). Association genetics of carbon isotope discrimination, height and foliar nitrogen in a natural population of *Pinus taeda* L. *Heredity, 107*(2), 105.

Currat, M., Ruedi, M., Petit, R. J., & Excoffier, L. (2008). The hidden side of invasions: Massive introgression by local genes. *Evolution, 62*(8), 1908–1920.

Cwynar, L. C., & MacDonald, G. M. (1987). Geographical variation of lodgepole pine in relation to population history. *The American Naturalist, 129*(3), 463–469.

Dagher-Kharrat, M. B., Grenier, G., Bariteau, M., Brown, S., Siljak-Yakovlev, S., & Savouré, A. (2001). Karyotype analysis reveals interspecific differentiation in the genus *Cedrus* despite genome size and base composition constancy. *Theoretical and Applied Genetics, 103*(6–7), 846–854.

Dallimore, W., Jackson, A. B., & Harrison, S. G. (1967). *A handbook of Coniferae and Ginkgoaceae* (4th ed.). New York: St. Martin's Press.

Dantas, L. G., Esposito, T., de Sousa, A. C. B., Félix, L., Amorim, L. L., Benko-Iseppon, A. M., et al. (2015). Low genetic diversity and high differentiation among relict populations of the neotropical gymnosperm Podocarpus sellowii (Klotz.) in the Atlantic Forest. *Genetica, 143*(1), 21–30.

Daoust, G., & Beaulieu, J. (2004). Genetics, breeding, improvement and conservation of *Pinus strobus* in Canada. In R. A. Sniezko, S. Samman, S. E. Schlarbaum, & H. B. Kriebel (Eds.), *Breeding and genetic resources of five-needle pines: Growth, adaptability, and pest resistance. Proceedings RMRS-P-32* (pp. 3–11). Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Darychuk, N., Hawkins, B. J., & Stoehr, M. (2012). Trade-offs between growth and cold and drought hardiness in submaritime Douglas-fir. *Canadian Journal of Forest Research, 42*(8), 1530–1541.

Daubenmire, R. (1968). *Plant communities: A textbook of plant synecology*. New York: Harper and Row, Publishers.

Dauwe, R., Holliday, J. A., Aitken, S. N., & Mansfield, S. D. (2012). Metabolic dynamics during autumn cold acclimation within and among populations of Sitka spruce (*Picea sitchensis*). *New Phytologist, 194*(1), 192–205.

De Diego, N., Pérez-Alfocea, F., Cantero, E., Lacuesta, M., & Moncaleán, P. (2012). Physiological response to drought in radiata pine: Phytohormone implication at leaf level. *Tree Physiology, 32*(4), 435–449.

De La Mata, R., & Zas, R. (2010). Performance of maritime pine Spanish Mediterranean provenances at young ages in transitional region between Atlantic and Mediterranean climates in NW Spain. *Silvae Genetica, 59*(1), 8–17.

De La Torre, A. R., Roberts, D. R., & Aitken, S. N. (2014a). Genome-wide admixture and ecological niche modelling reveal the maintenance of species boundaries despite long history of interspecific gene flow. *Molecular Ecology, 23*(8), 2046–2059.

De La Torre, A. R., Wang, T., Jaquish, B., & Aitken, S. N. (2014c). Adaptation and exogenous selection in a *Picea glauca* × *Picea engelmannii* hybrid zone: Implications for forest management under climate change. *New Phytologist, 201*(2), 687–699.

De La Torre, A. R., Puiu, D., Crepeau, M. W., Stevens, K., Salzberg, S. L., Langley, C. H., & Neale, D. B. (2018). Genomic architecture of complex traits in loblolly pine. *New Phytologist*.

De La Torre, A. R., Birol, I., Bousquet, J., Ingvarsson, P. K., Jansson, S., Jones, S. J., et al. (2014b). Insights into conifer giga-genomes. *Plant Physiology, 166*(4), 1724–1732.

De La Torre, J. E., Egan, M. G., Katari, M. S., Brenner, E. D., Stevenson, D. W., Coruzzi, G. M., & DeSalle, R. (2006). ESTimating plant phylogeny: Lessons from partitioning. *BMC Evolutionary Biology, 6*(1), 48.

De Lafontaine, G., Prunier, J., Gérardi, S., & Bousquet, J. (2015). Tracking the progression of speciation: Variable patterns of introgression across the genome provide insights on the species delimitation between progenitor–derivative spruces (*Picea mariana* × *P. rubens*). *Molecular Ecology, 24*(20), 5229–5247.

De Laubenfels, D. J. (1965). The relationships of *Fitzroya cupressoides* (Molina) Johnstone and *Diselma archeri* J.D. Hooker based on morphological considfeations. *Phytomorphology, 15*, 414–419.

De Laubenfels, D. J. (1969). A revision of the Malesian and Pacific rainforest conifers, I. Podocarpaceae, in part. *Journal of the Arnold Arboretum, 50*, 274–369.

De Laubenfels, D. J. (1972). Podocarpaceae, pp. 9-22 in J.-F. Leroy, ed., Flore de Madagascar et des Comores, Gymnospermes. Paris: Museum National d'Histoire Naturelle. (in French).

De Laubenfels, D. J. (1985). A taxonomic revision of the genus *Podocarpus*. *Blumea, 30*, 251–278.

De Laubenfels, D. J. (1987). Revision of the genus *Nageia* Podocarpaceae. *Blumea, 12*, 209–211.

De Luc, A., Adams, R. A., & Zhang, M. (1999). Using random amplification of polymorphic DNA for taxonomic evaluation of Pfitzer junipers. *Hortscience, 34*, 1123–1125.

De Miguel, M., Cabezas, J. A., de María, N., Sánchez-Gómez, D., Guevara, M. Á., Vélez, M. D., et al. (2014). Genetic control of functional traits related to photosynthesis and water use effi-

ciency in *Pinus pinaster* Ait. drought response: integration of genome annotation, allele association and QTL detection for candidate gene identification. *BMC Genomics, 15*(1), 464.
Dean, J. (2011). Future prospects. In C. Plomion, J. Bousquet, & C. Kole (Eds.), *Genetics, genomics and breeding of conifers*. New Hampshire: Science Publishers.
Debazac, E. F. (1964). *Manuel des Coniferes*. Nancy: Ecole Nationale des Eaus et Forets.
DeGiorgio, M., Syring, J., Eckert, A. J., Liston, A., Cronn, R., Neale, D. B., & Rosenberg, N. A. (2014). An empirical evaluation of two-stage species tree inference strategies using a multilocus dataset from North American pines. *BMC Evolutionary Biology, 14*(1), 67.
Delgado, P., Chaos, A., & Alvarez-Buylla, E. R. (1999). High population differentiation and genetic variation in the endangered Mexican pine *Pinus rzedowskii* (Pinaceae). *American Journal of Botany, 86*(5), 669–676.
Delgado, P., Cuenca, A., Escalante, A. E., Molina-Freaner, F., & Piñero, D. (2002). Comparative genetic structure in pines: Evolutionary and conservation consequences. *Revista Chilena de Historia Natural, 75*, 27–37.
Delgado, P., Piñero, D., Rebolledo, V., Jardón, L., & Chi, F. (2011). Genetic variation and demographic contraction of the remnant populations of Mexican Caribbean pine (*Pinus caribaea* var. *hondurensis*: Pinaceae). *Annals of Forest Science, 68*(1), 121–128.
Delgado, P., Salas-Lizana, R., Vázquez-Lobo, A., Wegier, A., Anzidei, M., Alvarez-Buylla, E. R., et al. (2007). Introgressive hybridization in *Pinus montezumae* Lamb and *Pinus pseudostrobus* Lindl.(Pinaceae): Morphological and molecular (cpSSR) evidence. *International Journal of Plant Sciences, 168*(6), 861–875.
Deng, H. S., Zhang, D. M., Fu, C. X., & Hong, D. Y. (2008). Behavior of meiotic chromosomes in *Pinus wallichiana*, *P. strobus* and their hybrid and nrDNA localization in pollen mother cells of the hybrid by using FISH. *Journal of Integrative Plant Biology, 50*(3), 360–367.
DeVerno, L. L., Charest, P. J., & Bonen, L. (1993). Inheritance of mitochondrial DNA in the conifer Larix. *Theoretical and Applied Genetics, 86*(2–3), 383–388.
Devey, M. E., Carson, S. D., Nolan, M. F., Matheson, A. C., Te Riini, C., & Hohepa, J. (2004a). QTL associations for density and diameter in *Pinus radiata* and the potential for marker-aided selection. *Theoretical and Applied Genetics, 108*(3), 516–524.
Devey, M. E., Delfino-Mix, A., Kinloch, B. B., Jr., & Neale, D. B. (1995). Random amplified polymorphic DNA markers tightly linked to a gene for resistance to white pine blister rust in sugar pine. *Proceedings of the National Academy of Sciences of the United States of America, 92*(6), 2066–2070.
Devey, M. E., Groom, K. A., Nolan, M. F., Bell, J. C., Dudzinski, M. J., Old, K. M., et al. (2004b). Detection and verification of quantitative trait loci for resistance to Dothistroma needle blight in *Pinus radiata*. *Theoretical and Applied Genetics, 108*(6), 1056–1063.
Devey, M. E., Matheson, A. C., & Carson, M. (2001). An international project to study pitch canker in radiata pine. *New Zealand Journal of Forestry, 46*(2), 38–41.
Devey, M. E., Sewell, M. M., Uren, T. L., & Neale, D. B. (1999). Comparative mapping in loblolly and radiata pine using RFLP and microsatellite markers. *Theoretical and Applied Genetics, 99*(3–4), 656–662.
DeWald, L. E., & Kolanoski, K. M. (2017). Conserving genetic diversity in ecological restoration: A case study with ponderosa pine in northern Arizona, USA. *New Forests, 48*(2), 337–361.
Dhillon, S. S., Berlyn, G. P., & Miksche, J. P. (1978). Nuclear DNA content in populations of *Pinus rigida*. *American Journal of Botany, 65*, 192–196.
Dhir, N. K., & Miksche, J. P. (1974). Intraspecific variation of nuclear DNA content in *Pinus resinosa* Ait. *Canadian Journal of Genetics and Cytology, 16*(1), 77–83.
Di Pierro, E. A., Mosca, E., Rocchini, D., Binelli, G., Neale, D. B., & La Porta, N. (2016). Climate-related adaptive genetic variation and population structure in natural stands of Norway spruce in the South-Eastern Alps. *Tree Genetics & Genomes, 12*(2), 16.
Díaz, V., Muniz, L. M., & Ferrer, E. (2001). Random amplified polymorphic DNA and amplified fragment length polymorphism assessment of genetic variation in Nicaraguan populations of *Pinus oocarpa*. *Molecular Ecology, 10*(11), 2593–2603.
Di-Giovanni, F., & Kevan, P. (1991). Factors affecting pollen dynamics and its importance to pollen contamination: A review. *Canadian Journal of Forest Research, 21*, 1155–1170.

Dillon, S. K., Nolan, M., Li, W., Bell, C., Wu, H. X., & Southerton, S. G. (2010). Allelic variation in cell wall candidate genes affecting solid wood properties in natural populations and land races of *Pinus radiata*. *Genetics, 185*(4), 1477–1487.

Dillon, S. K., Nolan, M. F., Matter, P., Gapare, W. J., Bragg, J. G., & Southerton, S. G. (2013). Signatures of adaptation and genetic structure among the mainland populations of *Pinus radiata* (D. Don) inferred from SNP loci. *Tree Genetics & Genomes, 9*(6), 1447–1463.

Distribution map of *Abies alba* EUFORGEN 2009. www.euforgen.org.

Dolgosheina, E. V., Morin, R. D., Aksay, G., Sahinalp, S. C., Magrini, V., Mardis, E. R., et al. (2008). Conifers have a unique small RNA silencing signature. *RNA, 14*(8), 1508–1515.

Dong, J., & Wagner, D. B. (1993). Taxonomic and population differentiation of mitochondrial diversity in *Pinus banksiana* and *Pinus contorta*. *Theoretical and Applied Genetics, 86*(5), 573–578.

Dong, J., Wagner, D. B., Yanchuk, A. D., Carlson, M. R., Magnussen, S., Wang, X. R., & Szmidt, A. E. (1992). Paternal chloroplast DNA inheritance in *Pinus consora* and *Pinus banksiana*: independence of parenetal species or cross direction. *Journal of Heredity, 83*(6), 419–422.

Donoghue, M. J., & Doyle, J. A. (2000). Seed plant phylogeny: demise of the anthophyte hypothesis? *Current Biology, 10*(3), R106–R109.

Donoso, A., Rodriguez, V., Carrasco, A., Ahumada, R., Sanfuentes, E., & Valenzuela, S. (2015). Relative expression of seven candidate genes for pathogen resistance on *Pinus radiata* infected with *Fusarium circinatum*. *Physiological and Molecular Plant Pathology, 92*, 42–50.

Dorman, K. W. (1975). *The genetics and breeding of Southern pines* (USDA Forest Service Agriculture Handbook No. 471). Washington, D.C.: U.S. Department of Agriculture, Forest Service.

Doudrick, R. L., Heslop-Harrison, J. S., Nelson, C. D., Schmidt, T., Nance, W. L., & Schwarzacher, T. (1995). Karyotype of slash pine (*Pinus elliottii var. elliottii*) using patterns of fluorescence in situ hybridization and fluorochrome banding. *Journal of Heredity, 86*(4), 289–296.

Dowlatabadi, R., Weljie, A. M., Thorpe, T. A., Yeung, E. C., & Vogel, H. J. (2009). Metabolic footprinting study of white spruce somatic embryogenesis using NMR spectroscopy. *Plant Physiology and Biochemistry, 47*(5), 343–350.

Du, F. K., Peng, X. L., Liu, J. Q., Lascoux, M., Hu, F. S., & Petit, R. J. (2011). Direction and extent of organelle DNA introgression between two spruce species in the Qinghai-Tibetan Plateau. *New Phytologist, 192*(4), 1024–1033.

Du, F. K., Petit, R. J., & Liu, J. Q. (2009). More introgression with less gene flow: Chloroplast vs. mitochondrial DNA in the *Picea asperata* complex in China, and comparison with other conifers. *Molecular Ecology, 18*(7), 1396–1407.

Dubouzet, J. G., Donaldson, L., Black, M. A., McNoe, L., Liu, V., & Lloyd-Jones, G. (2014). Heterologous hybridisation to a *Pinus* microarray: Profiling of gene expression in *Pinus radiata* saplings exposed to ethephon. *New Zealand Journal of Forestry Science, 44*(1), 21.

Dubreuil, M., Riba, M., González-Martínez, S. C., Vendramin, G. G., Sebastiani, F., & Mayol, M. (2010). Genetic effects of chronic habitat fragmentation revisited: Strong genetic structure in a temperate tree, *Taxus baccata* (Taxaceae), with great dispersal capability. *American Journal of Botany, 97*(2), 303–310.

Ducci, F., Proietti, R., & Favre, J. M. (1999). Allozyme assessment of genetic diversity within the relic Sicilian fir *Abies nebrodensis* (Lojac.) Mattei. *Annals of Forest Science, 56*(4), 345–355.

Duffield, J. W., & Righter, F. I. (1953). *Annotated list of pine hybrids made at the Institute of Forest Genetics* (*Research note (California Forest and Range Experiment Station); no. 86*) (pp. 1–9). Berkeley, CA: U.S. Department of Agriculture, Forest Service, California Forest and Range Experiment Station.

Dumroese, R. K., Williams, M. I., Stanturf, J. A., & Clair, J. B. S. (2015). Considerations for restoring temperate forests of tomorrow: Forest restoration, assisted migration, and bioengineering. *New Forests, 46*(5–6), 947–964.

Dungey, H. S. (2001). Pine hybrids—A review of their use performance and genetics. *Forest Ecology and Management, 148*(1), 243–258.

Dungey, H. S., Low, C. B., Lee, J., Millier, M. A., Fleet, K., & Yanchuk, A. (2012). Developing breeding and deployment options for Douglas fir in New Zealand: Breeding and future forest conditions. *Silvae Genetica, 61*, 104–115.

Duran, C., Appleby, N., Edwards, D., & Batley, J. (2009). Molecular genetic markers: Discovery, applications, data storage and visualisation. *Current Bioinformatics, 4*(1), 16–27.

Dvorak, W. S., Hodge, G. R., & Kietzka, J. E. (2007). Genetic variation in survival, growth, and stem form of *Pinus leiophylla* in Brazil and South Africa and provenance resistance to pitch canker. *Southern Hemisphere Forestry Journal, 69*(3), 125–135.

Dvornyk, V., Sirviö, A., Mikkonen, M., & Savolainen, O. (2002). Low nucleotide diversity at the pal1 locus in the widely distributed *Pinus sylvestris*. *Molecular Biology and Evolution, 19*(2), 179–188.

Dzialuk, A., Chybicki, I., Gout, R., Mączka, T., Fleischer, P., Konrad, H., Curtu, A. L., Sofletea, N., & Valadon, A. (2014). No reduction in genetic diversity of Swiss stone pine (*Pinus cembra* L) in Tatra Mountains despite high fragmentation and small population size. *Conservation Genetics, 15*(6), 1433–1445.

Earle, C. J. (Ed.). (1997–2017). Gymnosperm database. http://www.conifers.org/index.php. Accessed 21 Oct 2017.

Earle, C. J. (2011). Conifer longevity. In C. J. Earle (Ed.) (1997–2017) Gymnosperm database. http://www.conifers.org/topics/longevity.htm. Accessed 21 Oct 2017.

Echt, C. S., DeVerno, L. L., Anzidei, M., & Vendramin, G. G. (1998). Chloroplast microsatellites reveal population genetic diversity in red pine, *Pinus resinosa* Ait. *Molecular Ecology, 7*(3), 307–316.

Echt, C. S., May-Marquardt, P., Hseih, M., & Zahorchak, R. (1996). Characterization of microsatellite markers in eastern white pine. *Genome, 39*(6), 1102–1108.

Eckenwalder, J. E. (1976). Re-evaluation of Cupressaceae and Taxodiaceae: A proposed merger. *Madrono, 23*, 237–256.

Eckenwalder, J. E. (2009). *Conifers of the world*. Portland: Timber Press.

Eckert, A. J., Bower, A. D., González-Martínez, S. C., Wegrzyn, J. L., Coop, G., & Neale, D. B. (2010a). Back to nature: Ecological genomics of loblolly pine (*Pinus taeda*, *Pinaceae*). *Molecular Ecology, 19*(17), 3789–3805.

Eckert, A. J., Bower, A. D., Jermstad, K. D., Wegrzyn, J. L., Knaus, B. J., Syring, J. V., & Neale, D. B. (2013a). Multilocus analyses reveal little evidence for lineage-wide adaptive evolution within major clades of soft pines (*Pinus* subgenus *Strobus*). *Molecular Ecology, 22*(22), 5635–5650.

Eckert, A. J., Bower, A. D., Wegrzyn, J. L., Pande, B., Jermstad, K. D., Krutovsky, K. V., et al. (2009a). Association genetics of coastal Douglas fir (*Pseudotsuga menziesii* var. *menziesii*, Pinaceae). I. Cold-hardiness related traits. *Genetics, 182*(4), 1289–1302.

Eckert, A. J., & Hall, B. D. (2006). Phylogeny, historical biogeography, and patterns of diversification for *Pinus* (Pinaceae): Phylogenetic tests of fossil-based hypotheses. *Molecular Phylogenetics and Evolution, 40*, 166–182.

Eckert, A. J., Maloney, P. E., Vogler, D. R., Jensen, C. E., Mix, A. D., & Neale, D. B. (2015). Local adaptation at fine spatial scales: An example from sugar pine (*Pinus lambertiana*, *Pinaceae*). *Tree Genetics & Genomes, 11*(3), 42.

Eckert, A. J., Shahi, H., Datwyler, S. L., & Neale, D. B. (2012b). Spatially variable natural selection and the divergence between parapatric subspecies of lodgepole pine (*Pinus contorta*, Pinaceae). *American Journal of Botany, 99*(8), 1323–1334.

Eckert, A. J., Tearse, B. R., & Hall, B. D. (2008). A phylogeographical analysis of the range disjunction for foxtail pine (*Pinus balfouriana*, Pinaceae): The role of Pleistocene glaciation. *Molecular Ecology, 17*(8), 1983–1997.

Eckert, A. J., van Heerwaarden, J., Wegrzyn, J. L., Nelson, C. D., Ross-Ibarra, J., González-Martínez, S. C., & Neale, D. B. (2010b). Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., *Pinaceae*). *Genetics, 185*(3), 969–982.

Eckert, A. J., Wegrzyn, J. L., Cumbie, W. P., Goldfarb, B., Huber, D. A., Tolstikov, V., et al. (2012a). Association genetics of the loblolly pine (*Pinus taeda,* Pinaceae) metabolome. *New Phytologist, 193*(4), 890–902.

Eckert, A. J., Wegrzyn, J. L., Liechty, J. D., Lee, J. M., Cumbie, W. P., Davis, J. M., et al. (2013b). The evolutionary genetics of the genes underlying phenotypic associations for loblolly pine (*Pinus taeda*, Pinaceae). *Genetics, 195*(4), 1353–1372.

Eckert, A. J., Wegrzyn, J. L., Pande, B., Jermstad, K. D., Lee, J. M., Liechty, J. D., et al. (2009b). Multilocus patterns of nucleotide diversity and divergence reveal positive selection at candidate genes related to cold hardiness in coastal Douglas fir (*Pseudotsuga menziesii* var. *menziesii*). *Genetics, 183*(1), 289–298.

Edwards, S., Jesson, L. K., Quiring, D., Weng, Y., Johns, R., & Park, Y. S. (2016). Genetically-based resistance of balsam fir (Pinaceae) to damage from the balsam twig aphid (Hemiptera: Aphididae). *The Canadian Entomologist, 148*(04), 426–433.

Edwards-Burke, M., Hamrick, J., & Price, R. (1997). Frequency and direction of hybridization in sympatric populations of *Pinus taeda* and *P. echinata* (Pinaceae). *American Journal of Botany, 84*(7), 879–879.

Egertsdotter, U., Van Zyl, L. M., MacKay, J., Peter, G., Kirst, M., Clark, C., et al. (2004). Gene expression during formation of earlywood and latewood in loblolly pine: Expression profiles of 350 genes. *Plant Biology, 6*(6), 654–663.

Ekramoddoullah, A. K. M., Liu, J. J., & Zamani, A. (2006). Cloning and characterization of a putative antifungal peptide gene (Pm-AMP1) in *Pinus monticola*. *Phytopathology, 96*(2), 164–170.

Ekramoddoullah, A. K. M., Yu, X., Sturrock, R., Zamani, A., & Taylor, D. (2000). Detection and seasonal expression pattern of a pathogenesis-related protein (PR-10) in Douglas-fir (*Pseudotsuga menziesii*) tissues. *Physiologia Plantarum, 110*(2), 240–247.

El-Kassaby, Y. A., & Benowicz, A. (2000). Effects of commercial thinning on genetic, plant species and structural diversity in second growth Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) stands. *Forest Genetics, 7*, 193–203.

El-Kassaby, Y. A., Dunsworth, B. G., & Krakowski, J. (2003). Genetic evaluation of alternative silvicultural systems in coastal montane forests: Western hemlock and amabilis fir. *Theoretical and Applied Genetics, 107*(4), 598–610.

El-Kassaby, Y. A., & Lstibůrek, M. (2009). Breeding without breeding. *Genetics Research, 91*(02), 111–120.

El-Kassaby, Y. A., & Ritland, K. (1996a). Genetic variation in low elevation Douglas-fir of British Columbia and its relevance to gene conservation. *Biodiversity and Conservation, 5*(6), 779–794.

El-Kassaby, Y. A., & Ritland, K. (1996b). Impact of selection and breeding on the genetic diversity in Douglas-fir. *Biodiversity and Conservation, 5*(6), 795–813.

El-Kassaby, Y. A., & Sziklai, O. (1982). Genetic variation of allozyme and quantitative traits in a selected Douglas-fir [*Pseudotsuga menziesii* var. *menziesii* (Mirb.) Franco] population. *Forest Ecology and Management, 4*(2), 115–126.

El-Kassaby, Y. A., & Yanchuk, A. D. (1994). Genetic diversity, differentiation, and inbreeding in Pacific yew from British Columbia. *Journal of Heredity, 85*(2), 112–117.

El-Lakany, M. H., & Sziklai, O. (1971). Intraspecific variation in nuclear characteristics of Douglas-fir. *Advancing Frontiers of Plant Science, 28*, 363–378.

Elmlinger, M. W., Bolle, C., Batschauer, A., Oelmüller, R., & Mohr, H. (1994). Coaction of blue light and light absorbed by phytochrome in control of glutamine synthetase gene expression in Scots pine (*Pinus sylvestris* L.) seedlings. *Planta, 192*(2), 189–194.

Elsik, C. G., & Williams, C. G. (2000). Retroelements contribute to the excess low-copy-number DNA in pine. *Molecular and General Genetics MGG, 264*(1), 47–55.

Elvira-Recuenco, M., Iturritxa, E., Majada, J., Alia, R., & Raposo, R. (2014). Adaptive potential of maritime pine (*Pinus pinaster*) populations to the emerging pitch canker pathogen, *Fusarium circinatum*. *PLoS One, 9*(12), e114971.

Emebiri, L. C., Devey, M. E., Matheson, A. C., & Slee, M. U. (1998). Age-related changes in the expression of QTLs for growth in radiata pine seedlings. *Theoretical and Applied Genetics, 97*(7), 1053–1061.

Epperson, B. K., Telewski, F. W., & Willyard, A. (2009). Chloroplast diversity in a putative hybrid swarm of *Ponderosae* (Pinaceae). *American Journal of Botany, 96*(3), 707–712.

Erickson, V. J., & Adams, W. T. (1989). Mating success in a coastal Douglas-fir seed orchard as affected by distance and floral phenology. *Canadian Journal of Forest Research, 19*, 1248–1255.

Eriksson, G., Namkoong, G., & Roberds, J. H. (1993). Dynamic gene conservation for uncertain futures. *Forest Ecology and Management, 62*(1–4), 15–37.

Ernst, S. G., Hanover, J. W., & Keathley, D. E. (1990). Assessment of natural interspecific hybridization of blue and Engelmann spruce in southwestern Colorado. *Canadian Journal of Botany, 68*(7), 1489–1496.

Ersoz, E. S., Wright, M. H., González-Martínez, S. C., Langley, C. H., & Neale, D. B. (2010). Evolution of disease response genes in loblolly pine: Insights from candidate genes. *PLoS One, 5*(12), e14234.

Escapa, I., Cuneo, R., & Axsmith, B. (2008). A new genus of the Cupressaceae (sensu *lato*) from the Jurassic of Patagonia: Implications for conifer megasporangiate cone homologies. *Review of Palaeobotany and Palynology, 151*, 110–122.

Escapa, I. H., & Catalano, S. A. (2013). Phylogenetic analysis of Araucariaceae: Integrating molecules, morphology, and fossils. *International Journal of Plant Sciences, 174*(8), 1153–1170.

Escapa, I. H., Rothwell, G. W., Stockey, R. A., & Cuneo, N. R. (2012). Seed cone anatomy of Cheirolepidiaceae (Conferales): Reinterpreting *Pararaucaria patagonica* Wieland. *American Journal of Botany, 99*(6), 1058–1068.

Eveno, E., Collada, C., Guevara, M. A., Léger, V., Soto, A., Díaz, L., Léger, P., González-Martínez, S. C., Cervera, M. T., Plomion, C., & Garnier-Géré, P. H. (2008). Contrasting patterns of selection at *Pinus pinaster* Ait. drought stress candidate genes as revealed by genetic differentiation analyses. *Molecular Biology and Evolution, 25*(2), 417–437.

Fady, B., & Conkle, M. T. (1993). Allozyme variation and possible phylogenetic implications in *Abies cephalonica* Loudon and some related eastern Mediterranean firs. *Silvae Genetica, 42*, 351–359.

Fady, B., Forest, I., Hochu, I., Ribiollet, A. D., De Beaulieu, J. L., & Pastuszka, P. (1999). Genetic differentiation in *Abies alba* Mill. populations from Southeastern France. *International Journal of Forest Genetics*.

Fageria, M. S., & Rajora, O. P. (2013). Effects of harvesting of increasing intensities on genetic diversity and population structure of white spruce. *Evolutionary Applications, 6*(5), 778–794.

Fageria, M. S., & Rajora, O. P. (2014). Effects of silvicultural practices on genetic diversity and population structure of white spruce in Saskatchewan. *Tree Genetics & Genomes, 10*(2), 287–296.

Falconer, D. S. (1960). *Introduction to quantitative genetics*. Edinburgh/London: Oliver and Boyd.

Fan, F., Cui, B., Zhang, T., Ding, G., & Wen, X. (2014). LTR-retrotransposon activation, IRAP marker development and its potential in genetic diversity assessment of masson pine (*Pinus massoniana*). *Tree Genetics & Genomes, 10*(1), 213–222.

Fan, F., Wen, X., Ding, G., & Cui, B. (2013). Isolation, identification, and characterization of genomic LTR retrotransposon sequences from masson pine (*Pinus massoniana*). *Tree Genetics & Genomes, 9*(5), 1237–1246.

FAO. (2009). Global review of forest pests and diseases. FAO forestry paper No. 156. Rome, Italy.

Farjon, A. (2001). *World checklist and bibliography of conifers* (2nd ed.). Kew: Royal Botanic Gardens.

Farjon, A. (2005). *A monograph of Cupressaceae and Sciadopitys*. Kew: Royal Botanic Gardens.

Farjon, A. (2008). *A natural history of conifers*. Portland: Timber Press.

Farjon, A. (2010). *A handbook of the world's conifers*. Leiden, The Netherlands: Brill Academic Publishers.

Farjon, A., & Filer, D. (2013). *An atlas of the world's conifers: An analysis of their distribution, biogeography, diversity and conservation status*. Leiden, The Netherlands: Brill Academic Publishers.

Farjon, A., & Garcia, S. O. (2003). Cone and ovule development in *Cunninghamia* and *Taiwania* (Cupressaceae sensu lato) and its significance for conifer evolution. *American Journal of Botany, 90*(1), 8–16.

Farjon, A., Hiep, N. T., Harder, D. K., Loc, P. K., & Averyanov, L. (2002). A new genus and species in Cupressaceae (Coniferales) from northern Vietnam, *Xanthocyparis vietnamensis*. *Novon, 12*(2), 179–189.

Farjon, A., & Rushforth, K. D. (1989). A classification of *Abies* Miller (Pinaceae). *Notes Roy Bot Gard Edinburgh, 46*(1), 59–79.

Fassett, N. C. (1944). Juniperus virginiana, *J. horizontalis* and *J. scopulorum* II. Hybrid swarms of *J. virginiana* and *J. scopulorum*. *Bulletin of the Torrey Botanical Club, 71*(5), 475–483.

Fazekas, A. J., & Yeh, F. C. (2006). Postglacial colonization and population genetic relationships in the *Pinus contorta* complex. *Canadian Journal of Botany, 84*(2), 223–234.

Felsenstein, J. (2004). *Inferring phylogenies*. Sunderland: Sinauer Assoc.

Ferguson, M. C. (1904). Contributions to the knowledge of the life history of Pinus with special reference to sporogenesis, the development of the gametophytes and fertilization (Vol. 6). The Academy.

Fernández-Pozo, N., Canales, J., Guerrero-Fernández, D., Villalobos, D. P., Díaz-Moreno, S. M., Bautista, R., et al. (2011). EuroPineDB: A high-coverage web database for maritime pine transcriptome. *BMC Genomics, 12*(1), 366.

Ferreyra, L. L., Latino, A., Calderon, A., & Gardenal, C. N. (1996). Allozyme polymorphism in *Austrocedrus chilensis* (D Don) florin and Boutelie from Patagonia Argentina. *Silvae Genetica, 45*, 61–63.

Fins, L., Byler, J., Ferguson, D., Harvey, A., Mahalovich, M. F., McDonald, G., Miller, D., Schwandt, J., & Zack, A. (2002). Return of the giants: Restoring western white pine to the inland northwest. *Journal of Forestry, 100*(4), 20–26.

Fins, L., & Seeb, L. W. (1986). Genetic variation in allozymes of western larch. *Canadian Journal of Forest Research, 16*(5), 1013–1018.

Florin, R. (1951). Evolution in cordaites and conifers. *Acta Horti Gergiani, 15*, 285–388.

Foster, A. J., Hall, D. E., Mortimer, L., Abercromby, S., Gries, R., Gries, G., et al. (2013). Identification of genes in *Thuja plicata* foliar terpenoid defenses. *Plant Physiology, 161*(4), 1993–2004.

Fowler, D. P., & Park, Y. S. (1983). Population studies of white spruce. I. Effects of self-pollination. *Canadian Journal of Forest Research, 13*, 1133–1138.

Fraga, M. F., Cañal, M., & Rodríguez, R. (2002a). Phase-change related epigenetic and physiological changes in *Pinus radiata* D. Don. *Planta, 215*(4), 672–678.

Fraga, M. F., Rodríguez, R., & Cañal, M. J. (2002b). Genomic DNA methylation-demethylation during aging and reinvigoration of *Pinus radiata*. *Tree Physiology, 22*(11), 813–816.

Frampton, J., & Benson, D. M. (2004). Phytophthora root rot mortality in Fraser fir seedlings. *Hortscience, 39*(5), 1025–1026.

Franceschi, V. R., Krokene, P., Christiansen, E., & Krekling, T. (2005). Anatomical and chemical defenses of conifer bark against bark beetles and other pests. *New Phytologist, 167*(2), 353–376.

Franich, R. A., Carson, M. J., & Carson, S. D. (1986). Synthesis and accumulation of benzoic acid in *Pinus radiata* needles in response to tissue injury by dothistromin, and correlation with resistance of *P. radiata* families to *Dothistroma pini*. *Physiological and Molecular Plant Pathology, 28*(2), 267–286.

Franich, R. A., Gadgil, P. D., & Shain, L. (1983). Fungistaxic effects of *Pinus radiata* needle epicuticular fatty and resin acids on *Dothistroma pini*. *Physiological Plant Pathology, 23*(2), 183–195.

Franich, R. A., Wells, L. G., & Barnett, J. R. (1977). Variation with tree age of needle cuticle topography and stomatal structure in Pinus radiata D Don. *Annals of Botany, 41*, 621–626.

Fraser, S., Martín-García, J., Perry, A., Kabir, M. S., Owen, T., Solla, A., Brown, A. V., Bulman, L. S., Barnes, I., Hale, M. D., & Vasconcelos, M. W. (2016). A review of Pinaceae resistance

mechanisms against needle and shoot pathogens with a focus on the *Dothistroma–Pinus* interaction. *Forest Pathology, 46*(5), 453–471.
Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution, 30*(7), 1687–1699.
Friedmann, M., Ralph, S. G., Aeschliman, D., Zhuang, J., Ritland, K., Ellis, B. E., et al. (2007). Microarray gene expression profiling of developmental transitions in Sitka spruce (*Picea sitchensis*) apical shoots. *Journal of Experimental Botany, 58*(3), 593–614.
Friesen, N., Brandes, A., & Heslop-Harrison, J. S. P. (2001). Diversity, origin, and distribution of retrotransposons (gypsy and copia) in conifers. *Molecular Biology and Evolution, 18*(7), 1176–1188.
Fujimoto, A., Kado, T., Yoshimaru, H., Tsumura, Y., & Tachida, H. (2008). Adaptive and slightly deleterious evolution in a conifer, *Cryptomeria japonica*. *Journal of Molecular Evolution, 67*(2), 201–210.
Fujimoto, Y., Maeta, T., & Tajima, M. (1983). Studies on resistant tree breeding to sugi bark midges (*Resseliella odai* Inouye). *Bulletin of the Forest Tree Breeding Institute, 1*, 109–123.
Fukui, M., Futamura, N., Mukai, Y., Wang, Y., Nagao, A., & Shinohara, K. (2001). Ancestral MADS box genes in sugi, Cryptomeria japonica D. Don (*Taxodiaceae*), homologous to the B function genes in angiosperms. *Plant and Cell Physiology, 42*(6), 566–575.
Furnier, G. R., & Adams, W. T. (1986). Geographic Patterns of Allozyme Variation in Jeffrey Pine. *American Journal of Botany, 73*(7), 1009–1015.
Furnier, G. R., Knowles, P., Clyde, M. A., & Dancik, B. P. (1987). Effects of avian seed dispersal on the genetic structure of whitebark pine populations. *Evolution, 41*(3), 607–612.
Furnier, G. R., Stine, M., Mohn, C. A., & Clyde, M. A. (1991). Geographic patterns of variation in allozymes and height growth in white spruce. *Canadian Journal of Forest Research, 21*(5), 707–712.
Futamura, N., Totoki, Y., Toyoda, A., Igasaki, T., Nanjo, T., Seki, M., et al. (2008). Characterization of expressed sequence tags from a full-length enriched cDNA library of *Cryptomeria japonica* male strobili. *BMC Genomics, 9*(1), 383.
Gadek, P. A., Alpers, D. L., Heslewood, M. M., & Quinn, C. J. (2000). Relationships within Cupressaceae sensu lato: A combined morphological and molecular approach. *American Journal of Botany, 87*(7), 1044–1057.
Gajurel, J. P., Cornejo, C., Werth, S., Shrestha, K. K., & Scheidegger, C. (2013). Development and characterization of microsatellite loci in the endangered species *Taxus wallichiana* (Taxaceae). *Applications in Plant Sciences, 1*(3), 1200281. http://www.bioone.org/loi/apps.
Galliano, H., Cabané, M., Eckerskorn, C., Lottspeich, F., Sandermann, H., & Ernst, D. (1993). Molecular cloning, sequence analysis and elicitor-/ozone-induced accumulation of cinnamyl alcohol dehydrogenase from Norway spruce (*Picea abies* L.). *Plant Molecular Biology, 23*(1), 145–156.
Gallis, A. T., Doulis, A. G., & Papageorgiou, A. C. (2007). Variability of cortex terpene composition in *Cupressus sempervirens* L. provenances grown in Crete, Greece. *Silvae Genetica, 56*(6), 294–299.
Gamache, I., Jaramillo-Correa, J. P., Payette, S., & Bousquet, J. (2003). Diverging patterns of mitochondrial and nuclear DNA diversity in subarctic black spruce: Imprint of a founder effect associated with postglacial colonization. *Molecular Ecology, 12*(4), 891–901.
Gapare, W. J., Aitken, S. N., & Ritland, C. E. (2005). Genetic diversity of core and peripheral Sitka spruce (*Picea sitchensis* (Bong.) Carr) populations: Implications for conservation of widespread species. *Biological Conservation, 123*(1), 113–123.
Gapare, W. J., Ivkovic, M., Dutkowski, G. W., Spencer, D. J., Buxton, P., & Wu, H. X. (2012). Genetic parameters and provenance variation of *Pinus radiata* D. Don. 'Eldridge collection' in Australia 1: Growth and form traits. *Tree Genetics and Genomes, 8*, 391–407.
García-Gil, M. R. (2008). Evolutionary aspects of functional and pseudogene members of the phytochrome gene family in Scots pine. *Journal of Molecular Evolution, 67*(2), 222–232.

García-Gil, M. R., Mikkonen, M., & Savolainen, O. (2003). Nucleotide diversity at two phytochrome loci along a latitudinal cline in *Pinus sylvestris*. *Molecular Ecology, 12*(5), 1195–1206.

Garrett, P. W. (1979). Species hybridization in the genus Pinus. USDA Forest Service Research Paper NE-436.

Gaudeul, M., Gardner, M. F., Thomas, P., Ennos, R. A., & Hollingsworth, P. M. (2014). Evolutionary dynamics of emblematic Araucaria species (Araucariaceae) in New Caledonia: Nuclear and chloroplast markers suggest recent diversification, introgression, and a tight link between genetics and geography within species. *BMC Evolutionary Biology, 14*(1), 1–21.

Gea, L. D., & Low, C. B. (1997). Genetic parameters and growth, form and canker resistance of *Cupressus macrocarpa* in New Zealand. *New Zealand Journal of Forest Science, 27*, 245–254.

Geada López, G., Kamiya, K., & Harada, K. (2002). Phylogenetic relationships of diploxylon pines (subgenus Pinus) based on plastid sequence data. *International Journal of Plant Sciences, 163*, 737–747.

Geisler, K., Jensen, N. B., Yuen, M. M. S., Madilao, L., & Bohlmann, J. (2016). Modularity of conifer diterpene resin acid biosynthesis: P450 enzymes of different CYP720B clades use alternative substrates and converge on the same products. *Plant Physiology, 171*, 152–164.

Gerber, S., Rodolphe, F., Bahrman, N., & Baradat, P. (1993). Seed-protein variation in maritime pine (*Pinus pinaster* Ait.) revealed by two-dimensional electrophoresis: Genetic determinism and construction of a linkage map. *Theoretical and Applied Genetics, 85*(5), 521–528.

Gernandt, D. S., Hernández-León, S., Salgado-Hernández, E., & Pérez de La Rosa, J. A. (2009). Phylogenetic relationships of *Pinus* subsection *Ponderosae* inferred from rapidly evolving cpDNA regions. *Systematic Botany, 34*(3), 481–491.

Gernandt, D. S., Holman, G., Campbell, C., Parks, M., Mathews, S., Raubeson, L. A., Liston, A., Stockey, R. A., & Rothwell, G. W. (2016). Phylogenetics of extant and fossil Pinaceae: Methods for increasing topological stability. *Botany, 94*(9), 863–884.

Gernandt, D. S., Liston, A., & Piñero, D. (2001). Variation in the nrDNA ITS of *Pinus* subsection *Cembroides*: Implications for molecular systematic studies of pine species complexes. *Molecular Phylogenetics and Evolution, 21*(3), 449–467.

Gernandt, D. S., Liston, A., & Piñero, D. (2003). Phylogenetics of *Pinus* subsections *Cembroides* and *Nelsoniae* inferred from cpDNA sequences. *Systematic Botany, 28*(4), 657–673.

Gernandt, D. S., Lopez, G. G., Garcia, S. O., & Liston, A. (2005). Phylogeny and classification of *Pinus*. *Taxon, 54*(1), 29–42.

Gernandt, D. S., Magallon, S., Lopez, G. G., Flores, O. Z., Willyard, A., & Liston, A. (2008). Use of simultaneous analyses to guide fossil-based calibrations of Pinaceae phylogeny. *International Journal of Plant Sciences, 169*(8), 1086–1099.

Gernandt, D. S., Willyard, A., Syring, J. V., & Liston, A. (2011). The conifers (Pinophyta). In C. Plomion, J. Bousquet, & C. Kole (Eds.), *Genetics, genomics and breeding of conifers* (pp. 1–39). Enfield: Science Publishers.

Gesell, A., Blaukopf, M., Madilao, L., Yuen, M. M. S., Withers, S. G., Mattsson, J., Russell, J. H., & Bohlmann, J. (2015). The gymnosperm cytochrome P450 CYP750B1 catalyzes stereospecific monoterpene hydroxylation of (+)-sabinene in thujone biosynthesis in *Thuja plicata*. *Plant Physiology, 168*, 94–106.

Ghimire, B., & Heo, K. (2014). Cladistic analysis of Taxaceae s.l. *Plant Systematics and Evolution, 300*, 217–223.

Giannini, R., Morcante, M., & Vendramin, G. G. (1991). Allozyme variation in Italian populations of *Picea abies* (L) Karst. *Silvae Genetica, 40*, 160–166.

Giertych, M. (1979). Summary of results on Scots pine (*Pinus sylvestris* L.) height growth in IUFRO provenance experiments. *Silvae Genetica, 28*, 136–152.

Gill, G. P., Brown, G. R., & Neale, D. B. (2003). A sequence mutation in the cinnamyl alcohol dehydrogenase gene associated with altered lignification in loblolly pine. *Plant Biotechnology Journal, 1*(4), 253–258.

Gilmore, S., & Hill, K. D. (1997). Relationships of the Wollemi pine (*Wollemi nobilis*) and a molecular phylogeny of the Auraucariaceae. *Telopea, 9*, 275–291.

Gion, J. M., Lalanne, C., Le Provost, G., Ferry-Dumazet, H., Paiva, J., Chaumeil, P., et al. (2005). The proteome of maritime pine wood forming tissue. *Proteomics, 5*(14), 3731–3751.

Glaubitz, J. C., El-Kassaby, Y. A., & Carlson, J. E. (2000). Nuclear restriction fragment length polymorphism analysis of genetic diversity in western redcedar. *Canadian Journal of Forest Research, 30*(3), 379–389.

Gleiker, K. P., & Carroll, A. L. (2011). Rating introgression between lodgepole and jack pine at the individual tree level using morphological traits. *Northern Journal of Applied Forestry, 28*(3), 138–145.

Godbout, J., Beaulieu, J., & Bousquet, J. (2010). Phylogeographic structure of jack pine (*Pinus banksiana*; Pinaceae) supports the existence of a coastal glacial refugium in northeastern North America. *American Journal of Botany, 97*, 1903–1912.

Godbout, J., Fazekas, A., Newton, C. H., Yeh, F. C., & Bousquet, J. (2008). Glacial vicariance in the Pacific Northwest: Evidence from a lodgepole pine mitochondrial DNA minisatellite for multiple genetically distinct and widely separated refugia. *Molecular Ecology, 17*(10), 2463–2475.

Godbout, J., Jaramillo-Correa, J. P., Beaulieu, J., & Bousquet, J. (2005). A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (*Pinus banksiana*), a broad-range North American conifer. *Molecular Ecology, 14*(11), 3497–3512.

Godbout, J., Yeh, F. C., & Bousquet, J. (2012). Large-scale asymmetric introgression of cytoplasmic DNA reveals Holocene range displacement in a North American boreal pine complex. *Ecology and Evolution, 2*(8), 1853–1866.

Goddard, R. E., McDonald, G. I., & Steinhoff, R. J. (1985). Measurement of field resistance, rust hazard, and deployment of blister rust-resistant western white pine. USDA Forest Service Res. Pap. INT-358.

Goldfarb, B., Lanz-Garcia, C., Lian, Z., & Whetten, R. (2003). Aux/IAA gene family is conserved in the gymnosperm, loblolly pine (*Pinus taeda*). *Tree Physiology, 23*(17), 1181–1192.

Gömöry, D., Foffová, E., Kmeť, J., Longauer, R., & Romšáková, I. (2010). Norway spruce (*Picea abies* [L.] Karst.) provenance variation in autumn cold hardiness: Adaptation or acclimation. *Acta Biologica Cracoviensia Series Botanica, 52*(2), 42–49.

Gömöry, D., Longauer, R., Liepelt, S., Ballian, D., Brus, R., Kraigher, H., et al. (2004). Variation patterns of mitochondrial DNA of *Abies alba* Mill. in suture zones of postglacial migration in Europe. *Acta societatis botanicorum Poloniae, 73*(3), 203–206.

Gömöry, D., Paule, L., Krajmerová, D., Romšáková, I., & Longauer, R. (2012). Admixture of genetic lineages of different glacial origin: A case study of *Abies alba* Mill. in the Carpathians. *Plant Systematics and Evolution, 298*(4), 703–712.

Gompert, Z., & Alex Buerkle, C. (2010). INTROGRESS: A software package for mapping components of isolation in hybrids. *Molecular Ecology Resources, 10*(2), 378–384.

González-Ibeas, D., Martinez-García, P. J., Famula, R. A., Delfino-Mix, A., Stevens, K. A., Loopstra, C. A., et al. (2016). Assessing the gene content of the megagenome: Sugar pine (*Pinus lambertiana*). *G3: Genes, Genomes, Genetics, 6*(12), 3787–3802.

González-Martínez, S. C., Alía, R., & Gil, L. (2002). Population genetic structure in a Mediterranean pine (*Pinus pinaster* Ait.): A comparison of allozyme markers and quantitative traits. *Heredity, 89*(3), 199–206.

González-Martínez, S. C., Ersoz, E., Brown, G. R., Wheeler, N. C., & Neale, D. B. (2006a). DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L. *Genetics, 172*(3), 1915–1926.

González-Martínez, S. C., Huber, D., Ersoz, E., Davis, J. M., & Neale, D. B. (2008). Association genetics in *Pinus taeda* L. II. Carbon isotope discrimination. *Heredity, 101*(1), 19.

González-Martínez, S. C., Krutovsky, K. V., & Neale, D. B. (2006b). Forest-tree population genomics and adaptive evolution. *New Phytologist, 170*(2), 227–238.

González-Martínez, S. C., Wheeler, N. C., Ersoz, E., Nelson, C. D., & Neale, D. B. (2007). Association genetics in *Pinus taeda* L I. Wood property traits. *Genetics, 175*(1), 399–409.

Gordon, T. R., Wikler, K. R., Clark, S. L., Okamoto, D., Storer, A. J., & Bonello, P. (1998). Resistance to pitch canker disease, caused by *Fusarium subglutinans* f. sp. *pini*, in Monterey pine (*Pinus radiata*). *Plant Pathology, 47*(6), 706–711.

Goto, Y., Kondo, T., Kuramoto, N., Ide, T., Yamamoto, K., Inaoka, K., & Yasueda, H. (2003). Mapping the gene encoding cry j 1; a major *Cryptomeria japonica* pollen allergen. *Silvae Genetica, 52*, 3–4.

Gould, P. J., Harrington, C. A., & Clair, J. B. S. (2011). Incorporating genetic variation into a model of budburst phenology of coast Douglas-fir (*Pseudotsuga menziesii* var. *menziesii*). *Canadian Journal of Forest Research, 41*, 139–150.

Gould, P. J., Harrington, C. A., & Clair, J. B. S. (2012). Growth phenology of coast Douglas-fir seed sources planted in diverse environments. *Tree Physiology, 32*, 1482–1496.

Govindaraju, D. R., & Cullis, C. A. (1992). Ribosomal DNA variation among populations of a *Pinus rigida* Mill. (pitch pine) ecosystem: I. Distribution of copy numbers. *Heredity, 69*(2), 133.

Grace, L. J., Charity, J. A., Gresham, B., Kay, N., & Walter, C. (2005). Insect-resistant transgenic *Pinus radiata*. *Plant Cell Reports*.

Gramzow, L., Weilandt, L., & Theißen, G. (2014). MADS goes genomic in conifers: Towards determining the ancestral set of MADS-box genes in seed plants. *Annals of Botany, 114*(7), 1407–1429.

Grant, V. (1971). *Plant speciation*. New York and London: Columbia University Press.

Graudal, L., Aravanopoulos, F., Bennadji, Z., Changtragoon, S., Fady, B., Kjær, E. D., et al. (2014). Global to local genetic diversity indicators of evolutionary potential in tree species within and outside forests. *Forest Ecology and Management, 333*, 35–51.

Gray, D. J., & Amerson, H. V. (1983). In vitro resistance of embryos of *Pinus taeda* to *Cronartium quercuum* f. sp. *fusiforme*: Ultrastructure and histology. *Phytopathology, 73*(11), 1492–1499.

Greenwood, D. R., Archibald, S. B., Mathewes, R. W., & Moss, P. T. (2005). Fossil biotas from the Okanagan Highlands, southern British Columbia and northeastern Washington State: Climates and ecosystems across an Eocene landscape. *Canadian Journal of Earth Sciences, 42*, 167–185.

Greenwood, M. S., Hopper, C. A., & Hutchison, K. W. (1989). Maturation in larch I. effect of age on shoot growth, foliar characteristics, and DNA methylation. *Plant Physiology, 90*(2), 406–412.

Griffin, J. R., & Critchfield, W. B. (1976). The distribution of forest trees in California. USDA Forest Service Res Paper PSW-82.

Grivet, D., Climent, J., Zabal-Aguirre, M., Neale, D. B., Vendramin, G. G., & González-Martínez, S. C. (2013). Adaptive evolution of Mediterranean pines. *Molecular Phylogenetics and Evolution, 68*(3), 555–566.

Grivet, D., Sebastiani, F., Alía, R., Bataillon, T., Torre, S., Zabal-Aguirre, M., et al. (2010). Molecular footprints of local adaptation in two Mediterranean conifers. *Molecular Biology and Evolution, 28*(1), 101–116.

Grivet, D., Sebastiani, F., González-Martínez, S. C., & Vendramin, G. G. (2009). Patterns of polymorphism resulting from long-range colonization in the Mediterranean conifer Aleppo pine. *New Phytologist, 184*(4), 1016–1028.

Groover, A., Devey, M., Fiddler, T., Lee, J., Megraw, R., Mitchel-Olds, T., et al. (1994). Identification of quantitative trait loci influencing wood specific gravity in an outbred pedigree of loblolly pine. *Genetics, 138*(4), 1293–1300.

Grossnickle, S. C., Sutton, B. C., Fan, S., & King, J. (1997). Characterization of Sitka × interior spruce hybrids: A biotechnological approach to seedlot determination. *The Forestry Chronicle, 73*(3), 357–362.

Grossnickle, S. C., Sutton, B. C., Folk, R. S., & Gawley, R. J. (1996). Relationship between nuclear DNA markers and physiological parameters in Sitka × interior spruce populations. *Tree Physiology, 16*(6), 547–555.

Guan, C., Zhang, H., Zhang, L., Li, X., Deng, J., & Jiang, T. (2011). Construction of genetic linkage maps of larch (*Larix kaempferi* × *Larix gmelini*) by RAPD markers and mapping of QTLS for larch. *Biotechnology & Biotechnological Equipment, 25*(1), 2197–2202.

Gugerli, F., Sperisen, C., Buchler, U., Brunner, I., Brodbeck, S., Palmer, J. D., & Qiu, Y.-L. (2001). The evolutionary split of Pinaceae from other conifers: Evidence from an intron loss and multigene phylogeny. *Molecular Phylogenetics and Evolution, 21*(2), 167–175.

Guillet-Claude, C., Isabel, N., Pelgas, B., & Bousquet, J. (2004). The evolutionary implications of knox-I gene duplications in conifers: Correlated evidence from phylogeny, gene mapping, and analysis of functional divergence. *Molecular Biology and Evolution, 21*(12), 2232–2245.

Günther, T., & Coop, G. (2013). Robust identification of local adaptation from allele frequencies. *Genetics, 195*(1), 205–220.

Guo, W., Grewe, F., Cobo-Clark, A., Fan, W., Duan, Z., Adams, R. P., et al. (2014). Predominant and substoichiometric isomers of the plastid genome coexist within *Juniperus* plants and have shifted multiple times during cupressophyte evolution. *Genome Biology and Evolution, 6*(3), 580–590.

Guries, R. P., & Ledig, F. T. (1982). Genetic diversity and population structure in pitch pine (*Pinus rigida* Mill.). *Evolution, 36*(2), 387–402.

Guttenberger, H., Mueller, M., & Grill, D. (1996). Cytogenetic studies on Norway spruce (*Picea abies* (L.) Karst.). *Phyton, 36*, 147–154.

Gwaze, D. P., Lott, L. H., & Nelson, D. C. (2003b). The efficacy of breeding for brown spot disease resistance in longleaf pine. In C. R. McKinley (Ed.), *Proceedings 27*th *Southern Forest Tree Improvement Conference* (pp. 63–71), Stillwater, Oklahoma, USA, June 24–27, 2003.

Gwaze, D. P., Zhou, Y., Reyes-Valdes, M. H., Al-Rababah, M. A., & Williams, C. G. (2003a). Haplotypic QTL mapping in an outbred pedigree. *Genetics Research, 81*(1), 43–50.

Hair, J. B. (1968). The chromosomes of the Cupressaceae: 1. *Tetraclineae* and *Actinostrobeae* (*Callitroideae*). *New Zealand Journal of Botany, 6*(3), 277–284.

Hajibabaei, M., Xia, J., & Drouin, G. (2006). Seed plant phylogeny: Gnetophytes are derived conifers and a sister group to Pinaceae. *Molecular Phylogenetics and Evolution., 40*(1), 208.

Hall, D. E., Robert, J. A., Keeling, C. I., Domanski, D., Qesada, A. L., Jancsik, S., Kuzyk, M., Br, H., Borchers, C. H., & Bohlmann, J. (2011). An integrated genomic, proteomic, and biochemical analysis of (+)-3-carene biosynthesis in Sitka spruce (*Picea sitchensis*) genotypes which are resistant or susceptible to white pine weevil. *The Plant Journal, 65*, 936–948.

Hall, D. E., Yuen, M. M., Jancsik, S., Quesada, A. L., Dullat, H. K., Li, M., et al. (2013). Transcriptome resources and functional characterization of monoterpene synthases for two host species of the mountain pine beetle, lodgepole pine (*Pinus contorta*) and jack pine *(Pinus banksiana*). *BMC Plant Biology, 13*(1), 80.

Hall, M. T. (1952). Variation and hybridization in *Juniperus*. *Annals of the Missouri Botanical Garden, 39*(1), 1–64.

Hamberger, B., Hall, D., Yuen, M., Oddy, C., Hamberger, B., Keeling, C. I., et al. (2009). Targeted isolation, sequence assembly and characterization of two white spruce (*Picea glauca*) BAC clones for terpenoid synthase and cytochrome P450 genes involved in conifer defence reveal insights into a conifer genome. *BMC Plant Biology, 9*(1), 106.

Hamberger, B., Ohnishi, T., Hamberger, B., Séguin, A., & Bohlmann, J. (2011). Evolution of diterpene metabolism: Sitka spruce CYP720B4 catalyses multiple oxidations in resin acid biosynthesis of conifer defense against insects. *Plant Physiology, 157*, 1677–1695.

Hamilton, J. A., Lexer, C., & Aitken, S. N. (2013a). Differential introgression reveals candidate genes for selection across a spruce (*Picea sitchensis* × *P. glauca*) hybrid zone. *New Phytologist, 197*(3), 927–938.

Hamilton, J. A., Lexer, C., & Aitken, S. N. (2013b). Genomic and phenotypic architecture of a spruce hybrid zone (*Picea sitchensis* × *P. glauca*). *Molecular Ecology, 22*(3), 827–841.

Hamilton, J. A., Royauté, R., Wright, J. W., Hodgskiss, P., & Ledig, F. T. (2017). Genetic conservation and management of the California endemic, Torrey pine (*Pinus torreyana* Parry): Implications of genetic rescue in a genetically depauperate species. *Ecology and Evolution, 7*(18), 7370–7381.

Hamrick, J. L. (2004). Response of forest trees to global environmental changes. *Forest Ecology and Management, 197*(1–3), 323–335.

Hamrick, J. L., & Godt, M. J. W. (1989). Allozyme diversity in plant species. In A. H. D. Brown, M. T. Clegg, A. L. Kahler, & B. S. Weir (Eds.), *Plant population genetics, breeding and genetic resources* (pp. 43–63). Sunderland, MA: Sinauer Associates Inc.

Hamrick, J. L., Godt, M. J. W., & Sherman-Broyles, S. L. (1992). Factors influencing levels of genetic diversity in woody plant species. In W. T. Adams, S. H. Strauss, D. L. Copes, A. R. Griffen (Eds.), *Population genetics of forest trees* (pp. 95–124). Proceedings of the international symposium population genet forest trees, Corvallis, OR, USA, July 31–August 2, 1990. Springer, Netherlands.

Hamrick, J. L., Linhart, Y. B., & Mitton, J. B. (1979). Relationships between life history characteristics and electrophoretically-detectable genetic variation in plants. *Annual Review of Ecology and Systematics, 10*, 173–200.

Hänninen, H., Beuker, E., Johnsen, Ø., Leinonen, I., Murray, M., Sheppard, L., & Skrøppa, T. (2001). Impacts of climate change on cold hardiness of conifers. In *Conifer cold hardiness* (pp. 305–333). Dordrecht: Springer.

Hansen, E. M., Reeser, P., Sutton, W., & Sniezko, R. A. (2012). Methods for screening Port-Orford-cedar for resistance to *Phytophthora lateralis*. In *Proceedings of the 4th International Workshop on Genetics of Host-Parasite Interactions in Forestry, Eugene, Oregon. KM Palmieri, tech. coord. Gen. Tech. Rep. PSW-GTR-240. US Department of Agriculture, Forest Service, Pacific Southwest Research Station, Albany, CA* (pp. 181–188).

Hansen, O. K., Nielsen, U. B., Edvardsen, O. M., Skulason, B., & Skage, J.-O. (2004). Nordic provenance trials with *Abies lasiocarpa* and *Abies lasiocarpa* var. *arizonica*: Three-year results. *Scandinavian Journal of Forest Research, 19*, 112–126.

Hao, D., Yang, L., & Xiao, P. (2011). The first insight into the *Taxus* genome via fosmid library construction and end sequencing. *Molecular Genetics and Genomics, 285*(3), 197–205.

Hao, D. C., Huang, B., & Yang, L. (2008a). Phylogenetic relations of the genus *Taxus* inferred from chloroplast intergenic spacer and nuclear coding DNA. *Biological & Pharmaceutical Bulletin, 31*(2), 260–265.

Hao, D. C., Xiao, P. G., Huang, B., Ge, G. B., & Yang, L. (2008b). Interspecific relationships and origins of Taxaceae and Cephalotaxaceae revealed by partitioned Bayesian analyses of chloroplast and nuclear DNA sequences. *Plant Systematics and Evolution, 276*, 89–104.

Harding, S., Rouland, H., & Wellendorf, H. (2003). Consistency of resistance by the spruce green aphid in different ontogenetic stages of Sitka spruce. *Agricultural and Forest Entomology., 5*(2), 107.

Hare, R. C., & Switzer, G. L. (1969). *Introgression with shortleaf pine may explain rust resistance in western loblolly pine* (USDA Forest Service, research note SO-88). New Orleans, LA: Southern Forest Experiment Station, Forest Service, U.S. Dept. of Agriculture.

Harfouche, A., Bahrman, N., Baradat, P., Guyon, J. P., Petit, R. J., & Kremer, A. (1999). Provenance hybridization in a diallel mating scheme of maritime pine (*Pinus pinaster*). II. Heterosis. *Canadian Journal of Forest Research, 30*, 10–16.

Harju, A. M., Venälainen, M., Beuker, E., Velling, P., & Viitanen, H. (2001). Genetic variation in the decay resistance of Scots pine wood against brown rot fungus. *Canadian Journal of Forest Research, 31*(7), 1244–1249.

Harkins, D. M., Johnson, G. N., Skaggs, P. A., Mix, A. D., Dupper, G. E., Devey, M. E., Kinloch, B. B., Jr., & Neale, D. B. (1998). Saturation mapping of a major gene for resistance to white pine blister rust in sugar pine. *Theoretical and Applied Genetics, 97*, 1355–1360.

Harper, C. J., Taylor, T. N., Krings, M., & Taylor, E. L. (2015). Arbuscular mycorrhizal fungi in a voltzialean conifer from the Triassic of Antarctica. *Review of Palaeobotany and Palynology, 215*, 76–84.

Harry, D. E., Temesgen, B., & Neale, D. B. (1998). Codominant PCR-based markers for *Pinus taeda* developed from mapped cDNA clones. *Theoretical and Applied Genetics, 97*(3), 327–336.

Hart, J. A. (1987). A cladistics analysis of conifers: Preliminary results. *Journal of the Arnold Arboretum, 68*, 269–307.

Haselhorst, M. S., & Buerkle, C. A. (2013). Population genetic structure of *Picea engelmannii*, *P. glauca* and their previously unrecognized hybrids in the central Rocky Mountains. *Tree Genetics & Genomes, 9*(3), 669–681.

Hastings, J. M., Potter, K. M., Koch, F. H., Megalos, M., & Jetton, R. M. (2017). Prioritizing conservation seed banking locations for imperiled hemlock species using multi-attribute frontier mapping. *New Forests, 48*(2), 301–316.

Hawkins, B. J., & Stoehr, M. (2009). Growth, phenology, and cold hardiness of 32 Douglas-fir full-sib families. *Canadian Journal of Forest Research, 39*(10), 1821–1834.

Hawley, G. J., & DeHayes, D. H. (1985). Hybridization among several North American firs. II. Hybrid verification. *Canadian Journal of Forest Research, 15*(1), 50–55.

Hawley, G. J., Schaberg, P. G., DeHayes, D. H., & Brissette, J. C. (2005). Silviculture alters the genetic structure of an eastern hemlock forest in Maine, USA. *Canadian Journal of Forest Research, 35*(1), 143–150.

Hayashi, E., Kondo, T., Terada, K., Kuramoto, N., & Kawasaki, S. (2004). Identification of AFLP markers linked to a resistance gene against pine needle gall midge in Japanese black pine. *Theoretical and Applied Genetics, 108*(6), 1177–1181.

Hayata, B. (1931). The Sciadopityaceae represented by *Sciadopitys verticillata* Sieb. Et Zucc., an endemic species of Japan. *Botanical Magazine (Tokyo), 45*, 567–569.

Hayden, K. J., Nettel, A., Dodd, R. S., & Garbelotto, M. (2011). Will all the trees fall? Variable resistance to an introduced forest disease in a highly susceptible host. *Forest Ecology and Management, 261*, 1781–1791.

He, C. Y., Zhang, J. G., & Duan, A. G. (2012b). Physiological and protein responses to drought in four pine seedlings. *Silvae Genetica, 61*(3), 93–103.

He, C. Y., Zhang, J. G., Duan, A. G., Sun, H. G., Fu, L. H., & Zheng, S. X. (2007). Proteins responding to drought and high-temperature stress in *Pinus armandii* Franch. *Botany, 85*(10), 994–1001.

He, T., Pausas, J. G., Belcher, C. M., Schwilk, D. W., & Lamont, B. B. (2012a). Fire-adapted traits of *Pinus* arose in the fiery Cretaceous. *New Phytologist, 194*, 751–759.

Hedman, H., Zhu, T., von Arnold, S., & Sohlberg, J. J. (2013). Analysis of the Wuschel-related gene family in the conifer *Picea abies* reveals extensive conservation as well as dynamic patterns. *BMC Plant Biology, 13*(1), 89.

Heller, G., Adomas, A., Li, G., Osborne, J., van Zyl, L., Sederoff, R., et al. (2008). Transcriptional analysis of *Pinus sylvestris* roots challenged with the ectomycorrhizal fungus *Laccaria bicolor*. *BMC Plant Biology, 8*(1), 19.

Hennig, W. (1950). Grundzuge einer Theorie der phylogenetishchen Systemmatik. Berlin: Deutscher Zentralverlag.

Hennig, W. (1966). *Phylogenetic systematics*. Urbana: University of Illinois Press.

Heredia, U. L., López, R., Collada, C., Emerson, B. C., & Gil, L. (2014). Signatures of volcanism and aridity in the evolution of an insular pine (*Pinus canariensis* Chr. Sm. Ex DC in Buch). *Heredity, 113*(3), 240–249.

Hermann, R. K., & Lavender, D. P. (1968). Early growth of Douglas-fir from various altitudes and aspects in southern Oregon. *Silvae Genetica, 17*(4), 143–151.

Hernandez-Castillo, G. R., Rothwell, G. W., & Mapes, G. (2001). Thucydiaceae fam. Nov., with a review and reevaluation of Paleozoic walchian conifers. *International Journal of Plant Sciences, 162*, 1155–1185.

Hernandez-Leon, S., Gernandt, D. S., Perez de la Rosa, J. A., & Jardon-Barbolla, L. (2013). Phylogenetic relationships and species delimitation in *Pinus* section *Trifoliae* inferred from plastid DNA. *PLoS One, 8*(7), e70501. https://doi.org/10.1371/journal.pone.0070501.

Hernandez-Serrano, A., Verdu, M., Santos-del-Blanco, L., Climent, J., Gonzalez-Martinez, S. C., & Pausas, J. G. (2014). Heritability and quantitative genetic divergence of serotiny, a fire-persistence plant trait. *Annals of Botany, 114*, 571–577.

Herrera, R., Krier, C., Lalanne, C., Ba, E. M., Stokes, A., Salin, F., et al. (2010). (Not) Keeping the stem straight: A proteomic analysis of maritime pine seedlings undergoing phototropism and gravitropism. *BMC Plant Biology, 10*(1), 217.

Heslop-Harrison, J. (1964). Forty years of genecology. *Advances in Ecological Research, 2*, 159–247.

Heuertz, M., De Paoli, E., Källman, T., Larsson, H., Jurman, I., Morgante, M., et al. (2006). Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce [*Picea abies* (L.) Karst]. *Genetics, 174*(4), 2095–2105.

Heuertz, M., Teufel, J., González-Martínez, S. C., Soto, A., Fady, B., Alía, R., & Vendramin, G. G. (2010). Geography determines genetic relationships between species of mountain pine (*Pinus mugo* complex) in western Europe. *Journal of Biogeography, 37*(3), 541–556.

Heusser, C. J. (1965). A Pleistocene phytogeographical sketch of the Pacific Northwest and Alaska. In *The quaternary of the United States* (pp. 469–483). Princeton: Princeton University Press.

Hietala, A. M., Eikenes, M., Kvaalen, H., Solheim, H., & Fossdal, C. G. (2003). Multiplex real-time PCR for monitoring *Heterobasidion annosum* colonization in Norway spruce clones that differ in disease resistance. *Applied and Environmental Microbiology, 69*(8), 4413–4420.

Highsmith, M. T., Lott, L. H., Gwaze, D. P. Lominy, D., & Nelson, C. D. (2003). Evaluation of the inheritance of tip moth susceptibility using pine hybrids planted in Southeast Mississippi. In C. R. McKinley (Ed.), *Proceedings 27th Southern Forest Tree Improvement Conference*, Stillwater, Oklahoma, USA, June 24–27, 2003.

Hilfiker, K., Gugerli, F., Schütz, J. P., Rotach, P., & Holderegger, R. (2004). Low RAPD variation and female-biased sex ratio indicate genetic drift in small populations of the dioecious conifer *Taxus baccata* in Switzerland. *Conservation Genetics, 5*(3), 357–365.

Hill, R. S., & Brodribb, T. J. (1999). Southern conifers in time and space. *Australian Journal of Botany, 47*, 639–696.

Hillis, D. M., Moritz, C., Mable, B. K., & Meyer, A. (Eds.). (1996). *Molecular systematics* (2nd ed.). Sunderland, MA: Sinauer Associates, Inc.

Hipkins, V. D., Krutovsky, K. V., & Strauss, S. (1994). Organelle genomes in conifers: structure, evolution, and diversity. *Forest Genetics, 1*, 179–189.

Hipkins, V. D., Marshall, K. A., Neale, D. B., Rottmann, W. H., & Strauss, S. H. (1995). A mutation hotspot in the chloroplast genome of a conifer (Douglas-fir: *Pseudotsuga*) is caused by variability in the number of direct repeats derived from a partially duplicated tRNA gene. *Current Genetics, 27*(6), 572–579.

Hirao, T., Watanabe, A., Kurita, M., Kondo, T., & Takata, K. (2008). Complete nucleotide sequence of the *Cryptomeria japonica* D. Don. chloroplast genome and comparative chloroplast genomics: Diversified genomic structure of coniferous species. *BMC Plant Biology, 8*(1), 70.

Hirayoshi, I., & Nakamura, Y. (1943). Chromosome number of *Sequoia sempervirens*. *Advances in Zoology and Botany, 2*, 73–75.

Hizume, M. (2015a). Fluorescent band pattern of chromosomes in *Pseudolarix amabilis*, *Pinaceae*. *Cytologia, 80*(2), 151–157.

Hizume, M. (2015b). Fluorescent banding pattern in chromosomes of *Tsuga forrestii* and *T. sieboldii*, *Pinaceae*. *Chromosome Botany, 10*(3), 95–100.

Hizume, M., & Kan, M. (2015). Fluorescent banding pattern of chromosomes in *Araucaria araucana*, *Araucariaceae*. *Cytologia, 80*(4), 399–403.

Hizume, M., Ohtaka, K. N., Takeda, K. M., Fujii, S., Yamasaki, Y., & Kondo, K. (2016). Chromosome banding in the genus *Pinus* IV. Fluorescent banding patterns of chromosomes in eight taxa of haploxylone pines. *Chromosome Botany, 11*(3), 61–71.

Hizume, M., Shibata, F., Matsusaki, Y., & Garajova, Z. (2002). Chromosome identification and comparative karyotypic analyses of four *Pinus* species. *Theoretical and Applied Genetics, 105*(4), 491–497.

Hodge, G. R., & Dvorak, W. S. (2000). Differential responses of Central American and Mexican pine species and *Pinus radiata* to infection by the pitch canker fungus. *New Forests, 19*(3), 241–258.

Hodge, G. R., Dvorak, W. S., & Tighe, M. E. (2012). Comparisons between laboratory and field results of frost tolerance of pines from the southern USA and Mesoamerica planted as exotics. *Southern Forests: Journal of Forest Science, 74*(1), 7–17.

Hoff, R. J. (1985). Susceptibility of lodgepole pine to the needle cast fungus, *Lophodermella corzcolor*. USDA Forest Service, Res. Note INT-349.5.

Hoff, R. J. (1986). Inheritance of the bark reaction resistance mechanism in *Pinus monticola* infected by *Cronartium ribicola*. USDA For. Serv. Res. Note INT-361.

Hoff, R. J. (1987). Susceptibility of Inland Douglas-fir to *Rhabdocline* needle cast. USDA Forest Service, Res. Note INT-375.

Hoff, R. J. (1988a). Resistance of ponderosa pine to the gouty pitch midge (*Cecidomyia piniinopis*). Forest Service. Res. Paper INT-387.

Hoff, R. J. (1988b). Susceptibility of Ponderosa pine to the needle cast fungus *Loplzodermium baculiferrim*. USDA Forest Service. Res. Paper INT-386. 6p.

Hoff, R. J. (1989). Differential susceptibility of ponderosa pine to the gouty pitch midge (*Cecidomyia piniillopis*). USDA Forest Service. Res. Paper INT-q99.

Hoff, R. J., & McDonald, G. I. (1971). Resistance of *Pinus monticola* to *Cronartium ribicola*: Short shoot fungicidal reaction. *Canadian Journal of Botany, 49*, 1235–1239.

Hoff, R. J., & McDonald, G. I. (1975). Hypersensitive reaction in *Pinus armandii* caused by *Cronartium ribicola*. *Canadian Journal of Forest Research, 5*(1), 146–148.

Hoff, R. J., & McDonald, G. I. (1978). Genetic variation in susceptibility of western white pine to needle blight. USDA Forest Service Res. Not INT-249.

Hoff, R. J., & McDonald, G. I. (1993). Variation of virulence of white pine blister rust. *European Journal of Forest Pathology, 23*, 103–109.

Hoff, R. J., & McDonald, G. I. (1994). Disease and insect resistance in conifers associated with the cedar/hemlock ecosystem. In D. M. Baumgartner, J. E. Lotan, & J. R. Tonn (Eds.), *Interior cedar-hemlock-white pine forests: Ecology and management* (pp. 151–155). March 24, 1994, Spokane, WA.

Holliday, J. A., Ralph, S. G., White, R., Bohlmann, J., & Aitken, S. N. (2008). Global monitoring of autumn gene expression within and among phenotypically divergent populations of Sitka spruce (*Picea sitchensis*). *New Phytologist, 178*(1), 103–122.

Holliday, J. A., Ritland, K., & Aitken, S. N. (2010). Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). *New Phytologist, 188*(2), 501–514.

Holliday, J. A., Wang, T., & Aitken, S. (2012). Predicting adaptive phenotypes from multilocus genotypes in Sitka spruce (*Picea sitchensis*) using random forest. *G3: Genes, Genomes. Genetics, 2*(9), 1085–1093.

Hong, Y. P., Hipkins, V. D., & Strauss, S. H. (1993). Chloroplast DNA diversity among trees, populations and species in the California closed-cone pines (*Pinus radiata, Pinus muricata* and *Pinus attenuata*). *Genetics, 135*(4), 1187–1196.

Hornoy, B., Pavy, N., Gerardi, S., Beaulieu, J., & Bousquet, J. (2015). Genetic adaptation to climate in white spruce involves small to moderate allele frequency shifts in functionally diverse genes. *Genome Biology and Evolution, 7*(12), 3269–3285.

Hotter, G. S. (1997). Elicitor-induced oxidative burst and phenylpropanoid metabolism in *Pinus radiata* cell suspension cultures. *Functional Plant Biology, 24*(6), 797–804.

Howe, G. T., Aitken, S. N., Neale, D. B., Jermstad, K. D., Wheeler, N. C., & Chen, T. H. H. (2003). From genotype to phenotype: Unraveling the complexities of cold adaptation in forest trees. *Canadian Journal of Botany, 81*, 1247–1266.

Howe, G. T., Jayawickrama, K., Cherry, M., Johnson, G. R., & Wheeler, N. C. (2006). Breeding Douglas-fir. *Plant Breeding Reviews, 27*, 245.

Howe, G. T., Yu, J., Knaus, B., Cronn, R., Kolpak, S., Dolan, P., et al. (2013). A SNP resource for Douglas-fir: de novo transcriptome assembly and SNP detection and validation. *BMC Genomics, 14*(1), 137.

Hsu, C. Y., Wu, C. S., & Chaw, S. M. (2014). Ancient nuclear plastid DNA in the yew family (*Taxaceae*). *Genome Biology and Evolution, 6*(8), 2111–2121.

https://colloque.inra.fr/tree-parasite-interactions2015/Abstracts for 5th International Workshop on the Genetics of Tree-Parasite Interactions. Orleans, France, 23–28 Aug. 2015.

https://www.fs.fed.us/psw/publications/documents/psw_gtr240/ this is proceedings of the 4th international workshop.

Huang, H. H., Xu, L. L., Tong, Z. K., Lin, E. P., Liu, Q. P., Cheng, L. J., & Zhu, M. Y. (2012a). De novo characterization of the Chinese fir (*Cunninghamia lanceolata*) transcriptome and analysis of candidate genes involved in cellulose and lignin biosynthesis. *BMC Genomics, 13*(1), 648.

Huang, L. C., Hsiao, L. J., Pu, S. Y., Kuo, C. I., Huang, B. L., Tseng, T. C., et al. (2012b). DNA methylation and genome rearrangement characteristics of phase change in cultured shoots of *Sequoia sempervirens*. *Physiologia Plantarum, 145*(2), 360–368.

Huber, D., & Amerson, H. (2011). Performance of the loblolly pine fusiform rust disease resistance gene (Fr1) in a slash X loblolly pine hybrid family. *Tree Genetics & Genomes, 7*(3), 535–540.

Hudgins, J. W., Christiansen, E., & Franceschi, V. R. (2004). Induction of anatomically based defense responses in stems of diverse conifers by methyl jasmonate: A phylogenetic perspective. *Tree Physiology, 24*(3), 251–264.

Hunt, R. S. (1981). Pine needle casts and blights in the Pacific Region. Can. For. Serv. Pac. For. Res. Cent., For. Pest Leaf. No. 43.

Hunt, R. S. (2004). *Blister-rust-resistant western white pines for British Columbia, Information report BC-X-397*. Victoria: Natural Resources Canada.

Hunt, R. S., & Jensen, G. D. (2001). Frequency of resistant western white pine seedlings from parents of different phenotypes. *Western Journal of Applied Forestry, 16*(4), 149–152.

Hunt, R. S., Ying, C. C., & Ashbee, D. (1987). Variation in damage among *Pinus contorta* provenances caused by the needle cast fungus *Lophoderrnelia concolor*. *Canadian Journal of Forest Research, 17*, 359.

Hurme, P., Repo, T., Savolainen, O., & Pääkkönen, T. (1997). Climatic adaptation of bud set and frost hardiness in Scots pine (*Pinus sylvestris*). *Canadian Journal of Forest Research, 27*(5), 716–723.

Hurme, P., Sillanpää, M. J., Arjas, E., Repo, T., & Savolainen, O. (2000). Genetic basis of climatic adaptation in Scots pine by Bayesian quantitative trait locus analysis. *Genetics, 156*(3), 1309–1322.

Hutchison, K. W., Harvie, P. D., Singer, P. B., Brunner, A. F., & Greenwood, M. S. (1990). Nucleotide sequence of the small subunit of ribulose-1, 5-bisphosphate carboxylase from the conifer *Larix laricina*. *Plant Molecular Biology, 14*(2), 281–284.

Huxley, J. (1938). Clines: An auxiliary taxonomic principle. *Nature, London, 142*, 219–220.

Hyun, S. K. (1977). Interspecific hybridization in pines with the special reference to *Pinus rigida* x *P. taeda*. *Silvae Genetica, 25*, 188–191.

Illingworth, K. (1978). Study of lodgepole pine genotype-environment interaction. In *Proceedings of IUFRO joint meeting of working parties: Douglas-fir provenances, lodgepole pine provenances, Sitka spruce provenances, and Abies provenances* (pp. 151–158). Vancouver.

Ilstedt, B., & Eriksson, G. (1986). Quality of intra- and inter-provenance families of *Picea abies* (L.) Karst. *Scandinavian Journal of Forest Research, 1*, 153–166.

Ingouff, M., Farbos, I., Lagercrantz, U., & von Arnold, S. (2001). PaHB1 is an evolutionary conserved HD-GL2 homeobox gene expressed in the protoderm during Norway spruce embryo development. *Genesis, 30*(4), 220–230.

Ingouff, M., Farbos, I., Wiweger, M., & von Arnold, S. (2003). The molecular characterization of PaHB2, a homeobox gene of the HD-GL2 family expressed during embryo development in Norway spruce. *Journal of Experimental Botany, 54*(386), 1343–1350.

Ingvarsson, P. K., Hvidsten, T. R., & Street, N. R. (2016). Towards integration of population and comparative genomics in forest trees. *New Phytologist, 212*(2), 338–344.

Isaac-Renton, M. G., Roberts, D. R., Hamann, A., & Spiecker, H. (2014). Douglas-fir plantations in Europe: A retrospective test of assisted migration to address climate change. *Global Change Biology, 20*, 2607–2617. doi: 10:1111/gcb.12604.

Isik, F., Amerson, H. V., Whetten, R. W., Garcia, S. A., & McKeand, S. E. (2012). Interactions of Fr genes and mixed-pathogen inocula in the loblolly pine-fusiform rust pathosystem. *Tree Genetics & Genomes, 8*(1), 15–25.

Isik, F., Isik, K., & Lee, S. J. (1999). Genetic variation in *Pinus brutia* Ten. in Turkey: I. Growth, biomass and stem quality traits. *Forest Genetics, 6*(2), 89–99.

Isik, F., Keskin, S., & McKeand, S. E. (2000). Provenance variation and provenance-site interaction in *Pinus brutia* Ten.: Consequences of defining breeding zones. *Silvae Genetica, 49*(4–5), 213–223.

Isik, F., Li, B., & Frampton, J. (2003). Estimates of addditive, dominance and epistatic genetic variances from a replicated test of loblolly pine. *Forest Science, 49*(1), 77–88.

Islam-Faridi, M. N., & Nelson, C. D. (2011). Cytogenetics. In C. Plomion, J. Bousquet, & C. Kole (Eds.), *Genetics, genomics and breeding of conifers* (pp. 128–140). Enfield: Science Publishers.

Islam-Faridi, M. N., Nelson, C. D., & Kubisiak, T. L. (2007). Reference karyotype and cytomolecular map for loblolly pine (*Pinus taeda* L.). *Genome, 50*(2), 241–251.

Ito, M., Suyama, Y., Ohsawa, T. A., & Watano, Y. (2008). Airborne-pollen pool and mating pattern in a hybrid zone between *Pinus pumila* and *P. parviflora* var. *pentaphylla*. *Molecular Ecology, 17*(23), 5092–5103.

IUCN. (2017). The IUCN red list of threatened species. http://www.iucnredlist.org/search

IUCN Red List of Threatened Species. (2013). Available at http://www.iucnredlist.org/details/32479/0

Jackman, S. D., Warren, R. L., Gibb, E. A., Vandervalk, B. P., Mohamadi, H., Chu, J., et al. (2016). Organellar genomes of white spruce (*Picea glauca*): Assembly and annotation. *Genome Biology and Evolution, 8*(1), 29–41.

Jacobs, J. J., Burnes, T. A., David, A. J., & Blanchette, R. A. (2009). Histopathology of primary needles and mortality associated with white pine blister rust in resistant and susceptible *Pinus strobus*. *Forest Pathology, 39*(6), 361–376.

Jansson, S., & Gustafsson, P. (1990). Type I and type II genes for the chlorophyll a/b-binding protein in the gymnosperm *Pinus sylvestris* (Scots pine): cDNA cloning and sequence analysis. *Plant Molecular Biology, 14*(3), 287–296.

Jansson, S., & Gustafsson, P. (1991). Evolutionary conservation of the chlorophyll a/b-binding proteins cDNAs encoding Type I, II and III LHC I polypeptides from the gymnosperm Scots pine. *Molecular and General Genetics MGG, 229*(1), 67–76.

Jaquish, B.C. (1990, August 20–24). Geographic variation in ten-year height growth of interior Douglas-fir in British Columbia. P. 2.144–2.155. In: Proc. Joint meeting Western Forest Genetics Assoc. and IUFRO Working Parties S2.02-05, 06,12 and 14 – Douglas-fir, contorta pine, Sitka spruce, and Abies breeding and genetic resources, Olympia, WA. Weyerhaeuser Co., Federal Way, WA.

Jaramillo-Correa, J. P., Aguirre-Planter, E., Khasa, D. P., Eguiarte, L. E., PiNEro, D., Furnier, G. R., & Bousquet, J. (2008). Ancestry and divergence of subtropical montane forest isolates: Molecular biogeography of the genus *Abies* (Pinaceae) in southern México and Guatemala. *Molecular Ecology, 17*(10), 2476–2490.

Jaramillo-Correa, J. P., Beaulieu, J., & Bousquet, J. (2001). Contrasting evolutionary forces driving population structure at expressed sequence tag polymorphisms, allozymes and quantitative traits in white spruce. *Molecular Ecology, 10*(11), 2729–2740.

Jaramillo-Correa, J. P., Beaulieu, J., Khasa, D. P., & Bousquet, J. (2009). Inferring the past from the present phylogeographic structure of North American forest trees: Seeing the forest for the genes. *Canadian Journal of Forest Research, 39*(2), 286–307.

Jaramillo-Correa, J. P., Beaulieu, J., Ledig, F. T., & Bousquet, J. (2006). Decoupled mitochondrial and chloroplast DNA population structure reveals Holocene collapse and population isolation in a threatened Mexican-endemic conifer. *Molecular Ecology, 15*(10), 2787–2800.

Jaramillo-Correa, J. P., & Bousquet, J. (2003). New evidence from mitochondrial DNA of a progenitor-derivative species relationship between black spruce and red spruce (Pinaceae). *American Journal of Botany, 90*(12), 1801–1806.

Jaramillo-Correa, J. P., & Bousquet, J. (2005). Mitochondrial genome recombination in the zone of contact between two hybridizing conifers. *Genetics, 171*(4), 1951–1962.

Jardón-Barbolla, L., Delgado-Valerio, P., Geada-López, G., Vázquez-Lobo, A., & Piñero, D. (2011). Phylogeography of *Pinus* subsection Australes in the Caribbean basin. *Annals of Botany, 107*(2), 229–241.

Jasińska, A. K., Wachowiak, W., Muchewicz, E., Boratyńska, K., Montserrat, J. M., & Boratyński, A. (2010). Cryptic hybrids between *Pinus uncinata* and *P. sylvestris*. *Botanical Journal of the Linnean Society, 163*(4), 473–485.

Jermstad, K. D., Bassoni, D. L., Jech, K. S., Ritchie, G. A., Wheeler, N. C., & Neale, D. B. (2003). Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas fir III. Quantitative trait loci-by-environment interactions. *Genetics, 165*(3), 1489–1506.

Jermstad, K. D., Bassoni, D. L., Jech, K. S., Wheeler, N. C., & Neale, D. B. (2001a). Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir I. Timing of vegetative bud flush. *Theoretical and Applied Genetics, 102*(8), 1142–1151.

Jermstad, K. D., Bassoni, D. L., Kinlaw, C. S., & Neale, D. B. (1998). Partial DNA sequencing of Douglas-fir cDNAs used for RFLP mapping. *Theoretical and Applied Genetics, 97*(5–6), 771–776.

Jermstad, K. D., Bassoni, D. L., Wheeler, N. C., Anekonda, T. S., Aitken, S. N., Adams, W. T., & Neale, D. B. (2001b). Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir II. Spring and fall cold-hardiness. *Theoretical and Applied Genetics, 102*(8), 1152–1158.

Jermstad, K. D., Eckert, A. J., Wegrzyn, J. L., Delfino-Mix, A., Davis, D. A., Burton, D. C., & Neale, D. B. (2011). Comparative mapping in Pinus: sugar pine (*Pinus lambertiana* Dougl.) and loblolly pine (*Pinus taeda* L.). *Tree Genetics & Genomes, 7*(3), 457–468.

Jermstad, K. D., Sheppard, L. A., Kinloch, B. B., Delfino-Mix, A., Ersoz, E. S., Krutovsky, K. V., & Neale, D. B. (2006). Isolation of a full-length CC–NBS–LRR resistance gene analog candidate from sugar pine showing low nucleotide diversity. *Tree Genetics & Genomes, 2*(2), 76.

Jetton, R. M., Whittier, W. A., Dvorak, W. S., & Rhea, J. R. (2013). Conserved ex situ genetic resources of eastern and Carolina hemlock: Eastern North American conifers threatened by the hemlock woolly adelgid. *Tree Planter's Notes, 56*(2), 59–71.

Jewell, F. F. (1990). Histopathology of longleaf pine needles infected by *Ploioderma hedgcockii* (Dearn.) Darker. *Forest Pathology, 20*(1), 24–31.

Johansen, A. D., & Latta, R. G. (2003). Mitochondrial haplotype distribution, seed dispersal and patterns of postglacial expansion of ponderosa pine. *Molecular Ecology, 12*(1), 293–298.

Johnsen, Ø., Dæhlen, O. G., Østreng, G., & Skrøppa, T. (2005). Daylength and temperature during seed production interactively affect adaptive performance of *Picea abies* progenies. *New Phytologist, 168*(3), 589–596.

Johnsen, Ø., Skrøppa, T., Junttila, O., & Dæhlen, O. G. (1996). Influence of the female flowering environment on autumn frost-hardiness of *Picea abies* progenies. *Theoretical and Applied Genetics, 92*(7), 797–802.

Johnson, G. R. (2002). Genetic variation in tolerance of Douglas-fir to Swiss needle cast as assessed by symptom expression. *Silvae Genetica, 51*(2–3), 80–88.

Jones, F. A., Hamrick, J. L., Peterson, C. J., & Squiers, E. R. (2006). Inferring colonization history from analyses of spatial genetic structure within populations of *Pinus strobus* and *Quercus rubra*. *Molecular Ecology, 15*(3), 851–861.

Jones, G. E., & Cregg, B. M. (2006). Screening exotic firs for the midwestern United States: Interspecific variation in adaptive traits. *Hortscience, 41*(2), 323–328.

Jones, W. G., Hill, K. D., & Allen, J. M. (1995). *Wollemia nobilis*, a new living Australian genus and species in the Araucariaceae. *Telopea, 6*, 173–176.

Joosen, R. V., Lammers, M., Balk, P. A., Bronnum, P., Konings, M. C. J. M., Perks, M., et al. (2006). Correlating gene expression to physiological parameters and environmental conditions during cold acclimation of *Pinus sylvestris*, identification of molecular markers using cDNA microarrays. *Tree Physiology, 26*(10), 1297.

Jorge, V., Dowkiw, A., Bankhead, S., & F-X. Saintonge. (2015). https://colloque.inra.fr/tree-parasite-interactions2015/Abstracts.

Jørgensen, S., Hamrick, J. L., & Wells, P. V. (2002). Regional patterns of genetic diversity in *Pinus flexilis* (Pinaceae) reveal complex species history. *American Journal of Botany, 89*(5), 792–800.

Jorgensen, S. M., & Hamrick, J. L. (1997). Biogeography and population genetics of whitebark pine, *Pinus albicaulis*. *Canadian Journal of Forest Research, 27*(10), 1574–1585.

Jost, L. O. U. (2008). GST and its relatives do not measure differentiation. *Molecular Ecology, 17*(18), 4015–4026.

Joyce, D. G., & Rehfeldt, G. E. (2013). Climatic niche, ecological genetics and impact of climate change on eastern white pine (*Pinus strobus* L.): Guidelines for land managers. *Forest Ecology and Management, 295*, 173–192.

Joyner, K. L., Wang, X. R., Johnston, J. S., Price, H. J., & Williams, C. G. (2001). DNA content for Asian pines parallels New World relatives. *Canadian Journal of Botany, 79*(2), 192–196.

Juang, Y., Jin, N., Diner, A., Tauer, C., Zhang, Y., & Damicone, J. (2003). Genetic analysis of a disease resistance gene from loblolly pine. In C. R. McKinley (Ed.), *Proceedings 27th Southern Forest Tree Improvement Conference* (p. 214), Stillwater, Oklahoma, USA, June 24–27, 2003.

Judd, W. S., Campbell, C. S., Kellogg, E. A., Stevens, P. F., & Donoghue, M. J. (2008). *Plant systematics: a phylogenetics approach* (3rd ed.). Sunderland: Sinauer Associates, Inc.

Jurgens, J. A., Blanchette, R. A., Zambino, P. J., & David, A. (2003). Histology of white pine blister rust in needles of resistant and susceptible eastern white pine. *Plant Disease, 87*(9), 1026–1030.

Kado, T., Ushio, Y., Yoshimaru, H., Tsumura, Y., & Tachida, H. (2006). Contrasting patterns of DNA variation in natural populations of two related conifers, *Cryptomeria japonica* and *Taxodium distichum* (*Cupressaceae sensu lato*). *Genes & Genetic Systems, 81*(2), 103–113.

Kado, T., Yoshimaru, H., Tsumura, Y., & Tachida, H. (2003). DNA variation in a conifer, *Cryptomeria japonica* (*Cupressaceae sensu lato*). *Genetics, 164*(4), 1547–1559.

Kamm, A., Doudrick, R. L., Heslop-Harrison, J. S., & Schmidt, T. (1996). The genomic and physical organization of Ty1-copia-like sequences as a component of large genomes in *Pinus elliottii* var. *elliottii* and other gymnosperms. *Proceedings of the National Academy of Sciences, 93*(7), 2708–2713.

Kamunya, S. M., Olng'otie, P. S., Chagala, E., Day, R. K., & Kipkore, W. K. (1999). Genetic variation and heritability of resistance to *Cinara cupressi* in Cupressus *lusitanica*. *Journal of Tropical Forest Science, 11*(3), 587–598.

Kandemir, G. E., Kaya, Z., Temel, F., & Önde, S. (2010). Genetic variation in cold hardiness and phenology between and within Turkish red pine (*Pinus brutia* Ten.) populations: Implications for seed transfer. *Silvae Genet, 59*, 49–57.

Kapeller, S., Lexer, M. J., Gegurek, T., Hiebl, J., & Schueler, S. (2012). Intraspecific variation in climate response of Norway spruce in the eastern Alpine range: Selecting appropriate provenances for future climate. *Forest Ecology and Management, 271*, 46–57.

Kapheim, K. M., Pan, H., Li, C., Salzberg, S. L., Puiu, D., Magoc, T., et al. (2015). Genomic signatures of evolutionary transitions from solitary to group living. *Science*. https://doi.org/10.1126/science.aaa4788.

Karkkainen, K., Koski, V., & Savolainen, O. (1996). Geographical variation in the inbreeding depression of scots pine. *Evolution, 50*(1), 111–119.

Karlsson, B., & Swedjemark, G. (2006). Genotypic variation in natural infection of *Heterobasidion* spp. in a *Picea abies* clone trial in southern Sweden. *Scandinavian Journal of Forest Research, 21*, 108–114.

Karlsson, B., Wellendorf, H., Roulund, H., & Werner, M. (2001). Genotype x trial interaction and stability across sites in 11 combined provenance and clone experiments with *Picea abies* in Denmark and Sweden. *Canadian Journal of Forest Research, 31*, 1826–1836.

Karvonen, P., Karjalainen, M., & Savolainen, O. (1993). Ribosomal RNA genes in Scots pine (*Pinus sylvestris* L.): Chromosomal organization and structure. *Genetica, 88*(1), 59–68.

Karvonen, P., & Savolainen, O. (1993). Variation and inheritance of ribosomal DNA in *Pinus sylvestris* L. (Scots pine). *Heredity, 71*(6), 614–622.

Kato, K. (1996). Comparison of resistance in elite hinoki trees (*Chamaecyparis obtusa* (Sieb. et Zucc.) Endl) to *Seiridium unicorne* (Cooke et Ellis) Sutto. *Bulletin of the National Forest Tree Breeding Center, 14*, 145–157.

Kato, K., & Kawamura, K. (1993). Clonal variation in Sugi [*Cryptomeria japonica*] on the resistance to cryptomeria bark borer (*Semanotus Japonicus* Lacordaire) and the population dynam-

ics by setting the bark borer free in a cage. *Bulletin of the National Forest Tree Breeding Center, 11*, 1–15.

Kato, K., & Taniguchi, T. (2003). Ten years examination in the primary screening test in a project for selecting Japanese cedar resistant to *Semanotus japonicus* (Coleoptera: Cerambycidae) conducted in Kanto Breeding Region. *Bulletin of the National Forest Tree Breeding Center, 19*, 13–24.

Kauffmann, M. E. (2012). *Conifer country*. Kneeland: Backcountry Press.

Kavanagh, K. L., Bond, B. J., Aitken, S. N., Gartner, B. L., & Knowe, S. (1999). Shoot and root vulnerability to xylem cavitation in four populations of Douglas-fir seedlings. *Tree Physiology, 19*(1), 31–37.

Kaya, Z., Sewell, M. M., & Neale, D. B. (1999). Identification of quantitative trait loci influencing annual height-and diameter-increment growth in loblolly pine (*Pinus taeda* L.). *Theoretical and Applied Genetics, 98*(3–4), 586–592.

Kayihan, G. C., Huber, D. A., Morse, A. M., White, T. L., & Davis, J. M. (2005). Genetic dissection of fusiform rust and pitch canker disease traits in loblolly pine. *Theoretical and Applied Genetics, 110*(5), 948–958.

Keane, R. E., Tomback, D. F., Aubry, C. A., Bower, A. D., Campbell, E. M., Cripps, C. L., et al. (2012). *A range-wide restoration strategy for whitebark pine (Pinus albicaulis)* (General Technical Report RMRS-GTR-279). Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. 108 p., 279.

Keane, R. E., Tomback, D. F., Murray, M. P., & Smith, C. M. (2011). *The future of high-elevation, five-needle white pines in Western North America: Proceedings of the High Five Symposium* (Proceedings RMRS-P-63). Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. 376 p., 63.

Keeley, J. E. (2012). Ecology and evolution of pine life histories. *Annals of Forest Science, 69*(4), 445–453.

Keeley, J. E., Pausas, J. G., Rundel, P. W., Bond, W. J., & Bradstock, R. A. (2011). Fire as an evolutionary pressure shaping plant traits. *Trends in Plant Science, 16*, 406–411.

Keeley, J. E., & Zedler, P. H. (1998). Evolution of life histories in *Pinus*. In D. M. Richardson (Ed.), *Ecology and biogeography of Pinus* (pp. 219–249). Cambridge, UK: Cambridge University Press.

Keeling, C. I., & Bohlmann, J. (2006a). Diterpene resin acids in conifers. *Phytochemistry, 67*(22), 2415–2423.

Keeling, C. I., & Bohlmann, J. (2006b). Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defence of conifers against insects and pathogens. *The New Phytologist, 170*, 657–675.

Keeling, C. I., Weisshaar, S., Ralph, S. G., Jancsik, S., Hamberger, B., Dullat, H. K., & Bohlmann, J. (2011). Transcriptome mining, functional characterization, and phylogeny of a large terpene synthase gene family in spruce (*Picea* spp.). *BMC Plant Biology, 11*(1), 43.

Kegley, S., Stipe, L., & Hepner, C. (1994). Tip moth control at the Lenore Tree Improvement Area 1993. USDA Forest Service, Northern Region, FPM Report 94-4 (3450), p. 7.

Kelch, D. G. (1998). Phylogeny of Podocarpaceae: comparison of evidence from morphology and 18S rDNA. *American Journal of Botany, 85*, 986.

Kenny, J. R., Dancik, B. P., Florence, L. Z., & Nargang, F. E. (1988). Nucleotide sequence of the carboxy-terminal portion of a lodgepole pine actin gene. *Canadian Journal of Forest Research, 18*(12), 1595–1602.

Ketchum, R. E., Rithner, C. D., Qiu, D., Kim, Y. S., Williams, R. M., & Croteau, R. B. (2003). *Taxus* metabolomics: Methyl jasmonate preferentially induces production of taxoids oxygenated at C-13 in *Taxus* x media cell cultures. *Phytochemistry, 62*(6), 901–909.

Khan, M. A., & Korban, S. S. (2012). Association mapping in forest trees and fruit crops. *Journal of Experimental Botany, 63*(11), 4045–4060.

Khasa, D. P., Jaramillo-Correa, J. P., Jaquish, B., & Bousquet, J. (2006). Contrasting microsatellite variation between subalpine and western larch, two closely related species with different distribution patterns. *Molecular Ecology, 15*(13), 3907–3918.

Khasa, P. D., & Dancik, B. P. (1996). Rapid identification of white-Engelmann spruce species by RAPD markers. *Theoretical and Applied Genetics, 92*(1), 46–52.

Khatab, I. A., Ishiyama, H., Inomata, N., Wang, X. R., & Szmidt, A. E. (2008). Phylogeography of Eurasian *Larix* species inferred from nucleotide variation in two nuclear genes. *Genes & Genetic Systems, 83*(1), 55–66.

Khoshoo, T. N. (1959). Polyploidy in gymnosperms. *Evolution, 13*(1), 24–39.

Khoshoo, T. N. (1961). Chromosome numbers in gymnosperms. *Silvae Genetica, 10*(1), 1–32.

Kim, C.S., Hong, S.H., Ryu, J.B., Choi, C., & Kim, J.S. (1978). Breeding of varieties of pines resistant to pine gall midge (*Thecodiplosis japonensis* Uchida et Inouye). 2. Seasonal variation of needle monoterpene composition in resistant *Pinus thunbergii*. In *Third World Consultation on Forest Tree Breeding, Canberra (Australia), 21 Mar 1977, Session 4 – Constraints on progress* (pp. 948–958).

Kim, I.-S., Kwon, H.-Y., Ryu, K.-O., & Choi, W. Y. (2008). Provenance by site interaction of *Pinus densiflora* in Korea. *Silvae Genetica, 57*(3), 131–138.

Kim, Z. S., Hwang, J. W., Lee, S. W., Yang, C., & Gorovoy, P. G. (2005). Genetic variation of Korean Pine (*Pinus koraiensis* Sieb. et Zucc.) at allozyme and RAPD markers in Korea, China and Russia. *Silvae Genetica, 54*(4–5), 235–245.

Kim, Z. S., Lee, S. W., & Hwang, J. W. (1997). Genetic diversity and structure of natural populations of *Pinus thunbergii* in Korea. *Silvae Genetica, 46*(2), 120–123.

Kimberley, M. O., & Hood, I. A. (2005). Douglas-fir provenance susceptibility to Swiss needle cast in New Zealand. *Australasian Plant Pathology, 34*(1), 57–62.

Kimura, M. (1968). Evolutionary rate at the molecular level. *Nature, 217*, 624–626.

Kimura, M. (1983). *The neutral theory of molecular evolution*. Cambridge: Cambridge University Press.

King, J. L., & Jukes, T. H. (1969). Non-Darwinian evolution. *Science, 164*, 788–797.

King, J. N., & Alfaro, R. I. (2004). Breeding for resistance to a shoot weevil of Sitka spruce in British Columbia, Canada. In C. Walter & M. Carson (Eds.). *Proc. of Conference Plantation forest biotechnology for the 21st century*.

King, J. N., Alfaro, R. I., Ott, P., & van Akker L. (2012). Phenotypic evidence suggests a possible major-gene element to weevil resistance in Sitka spruce. In *Proc. Fourth International Workshop on the Genetics of Host-Parasite Interactions in Forestry: Disease and Insect Resistance in Forest Trees*. United States Department of Agriculture, Forest Service Gen. Tech. Rep. PSW-GTR-240.

King, J. N., Yanchuk, A. D., Kiss, G. K., & Alfaro, R. I. (1997). Genetic and phenotypic relationships between weevil (*Pissodes strobi*) resistance and height growth in spruce populations of British Columbia. *Canadian Journal of Forest Research, 27*, 732–739.

Kinlaw, C. S., Ho, T., Ljungkvist, V., & Baysdorfer, C. (1997). Gene discovery in loblolly pine through cDNA sequencing [abstract]. In D. B. Neale (Ed.), *Forest Tree Genome Workshop*. Placerville, CA: Institute of Forest Genetics, USDA Forest Service (1997). Abstract nr 4.

Kinlaw, C. S., & Neale, D. B. (1997). Complex gene families in pine genomes. *Trends in Plant Science, 2*(9), 356–359.

Kinloch, B. B., Jr. (1982). Resistance to white pine blister rust in sugar pine: Research at the Pacific Southwest Forest and Range Experiment Station. In *Breeding insect and disease resistant forest trees workshop, 19–23 July 1982, Eugene, Ore* (pp. 86–91).

Kinloch, B. B., Jr. (2003). White pine blister rust in North America: Past and prognosis. *Phytopathology, 93*, 1044–1047.

Kinloch, B. B., Jr., & Dupper, G. E. (2002). Genetic specificity in the white pine-blister rust pathosystem. *Phytopathology, 92*(3), 278–280.

Kinloch, B. B., Jr., & Littlefield, J. L. (1977). White pine blister rust: Hypersensitive resistance in sugar pine. *Canadian Journal of Botany, 55*(9), 1148–1155.

Kinloch, B. B., Jr., Parks, G. K., & Fowler, C. W. (1970). White pine blister rust: Simply inherited resistance in sugar pine. *Science*, 193–195.

Kinloch, B. B., Jr., Sniezko, R. A., Barnes, G. D., & Greathouse, T. E. (1999). A major gene for resistance to white pine blister rust in western white pine from the western Cascade Range. *Phytopathology, 89*, 861–867.

Kinloch, B. B., Jr., Sniezko, R. A., & Dupper, G. E. (2003). Origin and distribution of Cr2, a gene for resistance to white pine blister rust in natural populations of western white pine. *Phytopathology, 93*(6), 691–694.

Kinloch, B. B., Jr., Davis, D. A., & Burton, D. C. (2004). Resistance and virulence interactions between two white pine species and blister rust in a 30-year field trial. In B. Li, & S. McKeand (Eds.), *Forest genetics and tree breeding in the age of genomics: Progress and future, Proceedings of 2004 IUFRO Joint Conference of Division 2, 1–5 November 2004*. Charleston, S.C.

Kirst, M., Johnson, A. F., Baucom, C., Ulrich, E., Hubbard, K., Staggs, R., et al. (2003). Apparent homology of expressed genes from wood-forming tissues of loblolly pine (*Pinus taeda* L.) with *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences, 100*(12), 7383–7388.

Kitzmiller, J. H. (1976). Tree improvement master plan for the California region. USDA For. Serv. Region 5. Appendix C: Breeding for white pine blister rust resistance in sugar pine: An operational plan for Region 5. pp. 97–123.

Kitzmiller, J. H., & Sniezko, R. A. (2000). Range-wide genetic variation in Port-Orford-cedar (*Chamaecyparis lawsoniana* (A. Murr.) Parl.). I. Early height growth at coastal and inland nurseries. *Journal of Sustainable Forestry, 10*, 57–67.

Kjellsen, T. D., Shiryaeva, L., Schröder, W. P., & Strimbeck, G. R. (2010). Proteomics of extreme freezing tolerance in Siberian spruce (*Picea obovata*). *Journal of Proteomics, 73*(5), 965–975.

Klein, T., Di Matteo, G., Rotenberg, E., Cohen, S., & Yakir, D. (2013). Differential ecophysiological response of a major Mediterranean pine species across a climate gradient. *Tree Physiology, 33*, 26–36.

Kleinhentz, M., Raffin, A., & Jactel, H. (1998). Genetic parameters and gain expected from direct selection for resistance to *Dioryctria sylvestrella* Ratz. (Lepidoptera: Pyralidae) in *Pinus pinaster* Ait., using a full diallel mating design. *Forest Genetics, 5*(3), 147–154.

Klumpp, R., & Dhar, A. (2011). Genetic variation of *Taxus baccata* L populations in the Eastern Alps and its implications for conservation management. *Scandinavian Journal of Forest Research, 26*(4), 294–304.

Knapp, M., Mudaliar, R., Havell, D., Wagstaff, S. J., & Lockhart, P. J. (2007). The drowning of New Zealand and the problem of *Agathis*. *Systematic Biology, 56*(5), 862–870.

Knott, S. A., Neale, D. B., Sewell, M. M., & Haley, C. S. (1997). Multiple marker mapping of quantitative trait loci in an outbred pedigree of loblolly pine. *Theoretical and Applied Genetics, 94*(6–7), 810–820.

Knowles, P. (1985). Comparison of isozyme variation among natural stands and plantations: Jack pine and black spruce. *Canadian Journal of Forest Research, 15*(5), 902–908.

Kobayashi, T., & Muramoto, M. (1989). Pitch canker of *Pinus luchuensis*, a new disease of Japanese forests. *Forest Pests, 40*, 169–173.

Kohda, H., Chiba, S., & Nagata, Y. (1981). Studies on the breeding of *Larix* species 9. Resistance of selected trees to Needle cast disease and genetic analysis by crossed families of Japanese larch. Technical note, Institute for Forest Tree Improvement, Oji Paper Co., Ltd., No. 196, Reprinted from Trans. Mtg. Hokkaido Br. Jpn. For. Soc. No. 30.

Kohmann, K., & Johnsen, O. (1994). The timing of bud set in seedlings of *Picea abies* from seed crops of a cool versus a warm spring and summer. *Silvae Genetica, 43*(5), 329–332.

Kolb, T. E., Fettig, C. J., Ayres, M. P., Bentz, B. J., Hicke, J. A., Mathiasen, R., Stewart, J. E., & Weed, A. S. (2016). Observed and anticipated impacts of drought on forest insects and diseases in the United States. *Forest Ecology and Management, 380*, 321–334.

Kole, C. (Ed.). (2007). *Forest trees* (Vol. 7). Berlin, Germany: Springer Science & Business Media.

Komulainen, P., Brown, G. R., Mikkonen, M., Karhu, A., Garcia-Gil, M. R., O'malley, D., et al. (2003). Comparing EST-based genetic maps between *Pinus sylvestris* and *Pinus taeda*. *Theoretical and Applied Genetics, 107*(4), 667–678.

Kondo, T., Terada, K., Hayashi, E., Kuramoto, N., Okamura, M., & Kawasaki, H. (2000). RAPD markers linked to a gene for resistance to pine needle gall midge in Japanese black pine (*Pinus thunbergii*). *Theoretical and Applied Genetics, 100*(3–4), 391–395.

Kondo, T., Tsumura, Y., Kawahara, T., & Okamura, M. (1998). Paternal inheritance of chloroplast and mitochondrial DNA in interspecific hybrids of *Chamaecyparis* spp. *Japanese Journal of Breeding, 48*(2), 177–179.

Konnert, M., & Bergmann, F. (1995). The geographical distribution of genetic variation of silver fir (*Abies alba*, *Pinaceae*) in relation to its migration history. *Plant Systematics and Evolution, 196*(1–2), 19–30.

Konnert, M., Maurer, W., Degen, B., & Kätzel, R. (2011). Genetic monitoring in forests-early warning and controlling system for ecosystemic changes. *iForest-Biogeosciences and Forestry, 4*(2), 77.

Kormutak, A., Vooková, B., Čamek, V., Salaj, T., Galgóci, M., Maňka, P., Boleček, P., Kuna, R., Kobliha, J., Lukáčik, I., & Gömöry, D. (2013). Artificial hybridization of some *Abies* species. *Plant Systematics and Evolution, 299*(6), 1175–1184.

Korol, L., Shklar, E., & Schiller, G. (2002). Diversity among circum-Mediterranean populations of Aleppo pine and differentiation from Brutia pine in their isoenzymes: Additional results. *Silvae Genetica, 51*(1), 35–41.

Korshikov, I. I., Tkacheva, Y. A., & Privalikhin, S. N. (2012). Cytogenetic abnormalities in Norway spruce (*Picea abies* (L.) Karst.) seedlings from natural populations and an introduction plantation. *Cytology and Genetics, 46*(5), 280–284.

Koski, V. (1970). A study of pollen dispersal as a mechanism of gene flow in conifers. *Communicationes Instituti Forestalis Fenniae, 70*(4), 1–78.

Koski, V. (1971). Embryonic lethals of *Picea abies* and *Pinus sylvestris*. *Communicationes Instituti Forestalis Fenniae, 75*, 1–30.

Kossack, D. S., Barrios, M., & Kinlaw, C. S. (1990). IFG: An abundant family of retrotransposons in pine. In *UFRO Molecular Genetics Meeting, Lake Tahoe*, September 30–October 4, 1990. (Abstract).

Kossack, D. S., & Kinlaw, C. S. (1999). IFG, a gypsy-like retrotransposon in *Pinus* (Pinaceae), has an extensive history in pines. *Plant Molecular Biology, 39*(3), 417–426.

Kou, Y. X., Shang, H. Y., Mao, K. S., Li, Z. H., Rushforth, K., & Adams, R. P. (2014). Nuclear and cytoplasmic DNA sequence data further illuminate the genetic composition of Leyland cypresses. *Journal of the American Society for Horticultural Science, 139*(5), 558–566.

Kovach, A., Wegrzyn, J. L., Parra, G., Holt, C., Bruening, G. E., Loopstra, C. A., et al. (2010). The *Pinus taeda* genome is characterized by diverse and highly diverged repetitive sequences. *BMC Genomics, 11*(1), 420.

Krakowski, J., Aitken, S. N., & El-Kassaby, Y. A. (2003). Inbreeding and conservation genetics in whitebark pine. *Conservation Genetics, 4*(5), 581–593.

Krakowski, J., & Stoehr, M. U. (2009). Coastal Douglas-fir provenance variation: Patterns and predictions for British Columbia seed transfer. *Annals of Forest Science, 66*, 811, 10 p.

Kraus, J. F. (1986). Breeding shortleaf X loblolly pine hybrids for the development of fusiform rust resistant loblolly pine. *Southern Journal of Applied Forestry, 10*, 195–197.

Krauss, K. W., Chambers, J. L., Allen, J. A., Soileau, D. M., Jr., & DeBosier, A. S. (2000). Growth and nutrition of baldcypress families planted under varying salinity regimes in Louisiana, USA. *Journal of Coastal Research, 16*(1), 153–163.

Kreyling, J., Wiesenberg, G. L., Thiel, D., Wohlfart, C., Huber, G., Walter, J., Jentsch, A., Konnert, M., & Beierkuhnlein, C. (2012). Cold hardiness of *Pinus nigra* Arnold as influenced by geographic origin, warming, and extreme summer drought. *Environmental and Experimental Botany, 78*, 99–108.

Kriebel, H. B. (1985). DNA sequence components of the *Pinus strobus* nuclear genome. *Canadian Journal of Forest Research, 15*(1), 1–4.

Krokene, P. (2015). Conifer defense and resistance to bark beetles. In *Bark beetles: Biology and ecology of native and invasive species* (pp. 177–207). ScienceDirect.

Kruglyak, L., & Lander, E. S. (1995). A nonparametric approach for mapping quantitative trait loci. *Genetics, 139*(3), 1421–1428.

Krupkin, A. B., Liston, A., & Strauss, S. H. (1996). Phylogenetic analysis of the hard pines (Pinus subgenus Pinus, Pinaceae) from chloroplast DNA restriction site analysis. *American Journal of Botany, 83*(4), 489–498.

Krutovskii, K. V., & Bergmann, F. (1995). Introgressive hybridization and phylogenetic relationships between Norway, *Picea abies* (L.) Karst., and Siberian, *P. obovata* Ledeb., spruce species studied by isozyme loci. *Heredity, 74*(5), 464–480.

Krutovsky, K. V., & Neale, D. B. (2005). Nucleotide diversity and linkage disequilibrium in cold-hardiness-and wood quality-related candidate genes in Douglas fir. *Genetics, 171*(4), 2029–2041.

Krutovsky, K. V., Troggio, M., Brown, G. R., Jermstad, K. D., & Neale, D. B. (2004). Comparative mapping in the Pinaceae. *Genetics, 168*(1), 447–461.

Kuang, H., Richardson, T. E., Carson, S. D., & Bongarten, B. C. (1999a). Genetic analysis of inbreeding depression in plus tree 850.55 of *Pinus radiata* D. Don. II. Genetics of viability genes. *Theoretical and Applied Genetics, 99*(1–2), 140–146.

Kubisiak, T. L., Nelson, C. D., Nowak, J., & Friend, A. L. (1999). Genetic linkage mapping of genomic regions conferring tolerance to high aluminum in slash pine. *Journal of Sustainable Forestry, 10*(1–2), 69–78.

Kuhlman, E. G. (1997). Inoculum density and expression of major gene resistance to fusiform rust disease in loblolly pine. *Plant Disease, 81*, 597–600.

Kujala, S. T., & Savolainen, O. (2012). Sequence variation patterns along a latitudinal cline in scots pine (*Pinus sylvestris*): Signs of clinal adaptation? *Tree Genetics & Genomes, 8*(6), 1451–1467.

Kumar, S., Spelman, R. J., Garrick, D. J., Richardson, T. E., Lausberg, M., & Wilcox, P. L. (2000). Multiple-marker mapping of wood density loci in an outbred pedigree of radiata pine. *Theoretical and Applied Genetics, 100*(6), 926–933.

Kunzmann, L. (2007). Araucariaceae (Pinopsida): Aspects in palaeobiogeography and palaeobiodiveristy in the Mesozoic. *Zoologischer Anzeiger, 246*, 257–277.

Kusumi, J., Zidong, L., Kado, T., Tsumura, Y., Middleton, B. A., & Tachida, H. (2010). Multilocus patterns of nucleotide polymorphism and demographic change in *Taxodium distichum* (Cupressaceae) in the lower Mississippi River alluvial valley. *American Journal of Botany, 97*(11), 1848–1857.

Kvarnheden, A., Albert, V. A., & Engström, P. (1998). Molecular evolution of cdc2 pseudogenes in spruce (*Picea*). *Plant Molecular Biology, 36*(5), 767–774.

Kvarnheden, A., & Engström, P. (1992). Genetically stable, individual-specific differences in hypervariable DNA in Norway spruce, detected by hybridization to a phage M13 probe. *Canadian Journal of Forest Research, 22*(1), 117–123.

Kvarnheden, A., Tandre, K., & Engström, P. (1995). A cdc2 homologue and closely related processed retropseudogenes from Norway spruce. *Plant Molecular Biology, 27*(2), 391–403.

L'Homme, Y., Séguin, A., & Tremblay, F. M. (2000). Different classes of retrotransposons in coniferous spruce species. *Genome, 43*(6), 1084–1089.

La Farge, T. (1975). Genetic differences in stem form of ponderosa pine grown in Michigan. *Silvae Genetica, 23*, 211–213.

Lachance, D., Hamel, L. P., Pelletier, F., Valéro, J., Bernier-Cardou, M., Chapman, K., Van Frankenhuyzen, K., & Séguin, A. (2007). Expression of a *Bacillus thuringiensis* cry1Ab gene in transgenic white spruce and its efficacy against the spruce budworm (*Choristoneura fumiferana*). *Tree Genetics & Genomes, 3*(2), 153–167.

Ladrach, W. E. (1998). Provenance research: The concept, application and achievement. In A. K. Mandal & G. I. Gibson (Eds.), *Forest genetics and tree breeding* (pp. 16–37). New Delhi: CBS Publishers and Distributors.

Laflamme, G., Rioux, D., Simard, M., Bussières, G., & Mallett, K. (2006). Resistance of *Pinus contorta* to the European race of *Gremmeniella abietina*. *Forest Pathology, 36*, *83–96*.

Lamant, T. (2012). Vegetative reproduction in gymnosperms. Bulletin de l'Association des Parcs Botaniques de France, n 53.

Lamara, M., Raherison, E., Lenz, P., Beaulieu, J., Bousquet, J., & MacKay, J. (2016). Genetic architecture of wood properties based on association analysis and co-expression networks in white spruce. *New Phytologist, 210*(1), 240–255.

Lambeth, C. (2000). Realized genetic gains for first generation improved loblolly pine in 45 tests in coastal North Carolina. *Southern Journal of Applied Forestry, 24*(3), 140–144.

Lambeth, C., Lee, B. C., O'Malley, D., & Wheeler, N. (2001). Polymix breeding with parental analysis of progeny: An alternative to full-sib breeding and testing. *Theoretical and Applied Genetics, 103*(6–7), 930–943.

Lambeth, C. C., McKeand, S., Rousseau, R., & Schmidtling, R. (2005). Planting nonlocal seed sources of loblolly pine - managing benefits and risks. *Southern Journal of Applied Forestry, 29*(2), 96–104.

Lamont, B., Lemaitre, D., Cowling, R., & Enright, N. (1991). Canopy seed storage in woody-plants. *The Botanical Review, 57*(4), 277–317.

Lamy, J.-B., Delzon, S., Bouche, P. S., Alia, R., Vendramin, G. G., Cochard, H., & Plomion, C. (2014). Limited genetic variability and phenotypic plasticity detected for cavitation resistance in a Mediterranean pine. *New Phytologist, 201*, 874–886.

Lander, E. S., & Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics, 121*(1), 185–199.

Lander, E. S., & Schork, N. J. (1994). Genetic dissection of complex traits. *Science, 265*(5181), 2037–2048.

Langlet, O. (1971). Two hundred years of genecology. *Taxon, 20*, 653–721.

Lanner, R. M. (1966). Needed: A new approach to the study of pollen dispersion. *Silvae Genetica, 15*, 50–52.

Lanner, R. M., & Phillips, A. M., III. (1992). Natural hybridization and introgression of pinyon pines in northwestern Arizona. *International Journal of Plant Sciences, 153*(2), 250–257.

Larochelle, F., Lavallée, R., Plourde, A., Daoust, G., & Bauce, E. (1993). Sélection d'épinettes de Norvège résistantes au charançon du pin blanc (résumé) Page 285 (Vol. 61) dans Association canadienne-française pour l'avancement des sciences (ACFAS), Compte rendu : 61e congrès de l'ACFAS. May 17–21, 1993, Université du Québec à Rimouski, Rimouski (Québec). ACFAS, Montréal (Québec).

Larsen, C. S. (1956). *Genetics in silviculture*. Fairlawn, New Jersey: Essential Books.

Latta, R. G., & Mitton, J. B. (1999). Historical separation and present gene flow through a zone of secondary contact in ponderosa pine. *Evolution, 53*(3), 769–776.

Lawrence, G. H. M. (1951). *Taxonomy of vascular plants*. New York/Toronto: Macmillan Company.

Le Provost, G., Paiva, J., Pot, D., Brach, J., & Plomion, C. (2003). Seasonal variation in transcript accumulation in wood-forming tissues of maritime pine (*Pinus pinaster* Ait.) with emphasis on a cell wall glycine-rich protein. *Planta, 217*(5), 820–830.

Ledig, F. T. (1988). The conservation of diversity in forest trees: Why and how should genes be conserved? *Bioscience, 38*(7), 471–479.

Ledig, F. T. (1992). Human impacts on genetic diversity in forest ecosystems. *Oikos, 63*, 87–108.

Ledig, F. T. (1998). Genetic variation in *Pinus*. In D. M. Richardson (Ed.), *Ecology and biogeography of Pinus* (pp. 251–280). Cambridge, UK: Cambridge University Press.

Ledig, F. T. (1999). Genetic diversity, genetic structure and biogeography of Pinus sabiniana Dougl. *Diversity and Distributions, 5*, 77–99.

Ledig, F. T. (2000). Founder effects and the genetic structure of Coulter pine. *Journal of Heredity, 91*(4), 307–315.

Ledig, F. T. (2012). Climate change and conservation. *Acta Silvatica et Lignaria Hungarica, 8*(1), 57–74.

Ledig, F. T., Bermejo-Velázquez, B., Hodgskiss, P. D., Johnson, D. R., Flores-López, C., & Jacob-Cervantes, V. (2000). The mating system and genic diversity in Martinez spruce, an extremely rare endemic of Mexico's Sierra Madre Oriental: An example of facultative selfing and survival in interglacial refugia. *Canadian Journal of Forest Research, 30*(7), 1156–1164.

Ledig, F. T., Capó-Arteaga, M. A., Hodgskiss, P. D., Sbay, H., Flores-López, C., Conkle, M. T., & Bermejo-Velázquez, B. (2001). Genetic diversity and the mating system of a rare Mexican piñon, *Pinus pinceana*, and a comparison with *Pinus maximartinezii* (Pinaceae). *American Journal of Botany, 88*(11), 1977–1987.

Ledig, F. T., & Conkle, M. T. (1983). Gene diversity and genetic structure in a narrow endemic, Torrey pine (*Pinus torreyana* Parry ex Carr.). *Evolution, 37*, 79–85.

Ledig, F. T., Conkle, M. T., Bermejo-Velazquez, B., Eguiluz-Piedra, T., Hodgskiss, P. D., Johnson, D. R., & Dvorak, W. S. (1999). Evidence for an extreme bottleneck in a rare Mexican pinyon: Genetic diversity, disequilibrium, and the mating system in *Pinus maximartinezii*. *Evolution, 53*, 91–99.

Ledig, F. T., Hodgskiss, P. D., & Jacob-Cervantes, V. (2002). Genetic diversity, mating system, and conservation of a Mexican subalpine relict, *Picea mexicana* Martinez. *Conservation Genetics, 3*(2), 113–122.

Ledig, F. T., Hodgskiss, P. D., & Johnson, D. R. (2005). Genic diversity, genetic structure, and mating system of Brewer spruce (Pinaceae), a relict of the Arcto-Tertiary forest. *American Journal of Botany, 92*(12), 1975–1986.

Ledig, F. T., Hodgskiss, P. D., & Johnson, D. R. (2006a). Genetic diversity and seed production in Santa Lucia fir (*Abies bracteata*), a relict of the Miocene broadleaved evergreen forest. *Conservation Genetics, 7*(3), 383–398.

Ledig, F. T., Hodgskiss, P. D., & Johnson, D. R. (2006b). The structure of genetic diversity in Engelmann spruce and a comparison with blue spruce. *Botany, 84*(12), 1806–1828.

Ledig, F. T., Hodgskiss, P. D., Krutovskii, K. V., Neale, D. B., & Eguiluz-Piedra, T. (2004). Relationships among the spruces (*Picea*, *Pinaceae*) of southwestern North America. *Systematic Botany, 29*(2), 275–295.

Ledig, F. T., Jacob-Cervantes, V., Hodgskiss, P. D., & Eguiluz-Piedra, T. (1997). Recent evolution and divergence among populations of a rare Mexican endemic, Chihuahua spruce, following Holocene climatic warming. *Evolution, 51*, 1815–1827.

Lee, S. W., Choi, W. Y., Kim, W. W., & Kim, Z. S. (2000). Genetic variation of *Taxus cuspidata* Sieb. Et Zucc. In Korea. *Silvae Genetica, 49*(3), 124–129.

Lee, S. W., Ledig, F. T., & Johnson, D. R. (2002). Genetic variation at allozyme and RAPD markers in *Pinus longaeva* (Pinaceae) of the White Mountains, California. *American Journal of Botany, 89*(4), 566–577.

Leibing, C., van Zonneveld, M., Jarvis, A., & Dvorak, W. (2009). Adaptation of tropical and subtropical pine plantation forestry to climate change: Realignment of Pinus patula and *Pinus tecunumanii* genotypes to 2020 planting site climates. *Scandinavian Journal of Forest Research, 24*, 483–493.

Leites, L. P., Rehfeldt, G. E., Robinson, A. P., Crookston, N. L., & Jaquish, B. (2012). Possibilities and limitations of using historic provenance tests to infer forest species growth responses to climate change. *Natural Resource Modeling, 25*(3), 409–433.

Lemieux, M. J., Beaulieu, J., & Bousquet, J. (2011). Chloroplast DNA polymorphisms in eastern hemlock: Range-wide genogeographic analyses and implications for gene conservation. *Canadian Journal of Forest Research, 41*(5), 1047–1059.

Lendvay, B., Höhn, M., Brodbeck, S., Mîndrescu, M., & Gugerli, F. (2014). Genetic structure in Pinus cembra from the Carpathian Mountains inferred from nuclear and chloroplast microsatellites confirms post-glacial range contraction and identifies introduced individuals. *Tree Genetics and Genomes, 10*(5), 1419–1433.

Lepoittevin, C., Harvengt, L., Plomion, C., & Garnier-Géré, P. (2012). Association mapping for growth, straightness and wood chemistry traits in the *Pinus pinaster* Aquitaine breeding population. *Tree Genetics & Genomes, 8*(1), 113–126.

Lerceteau, E., Plomion, C., & Andersson, B. (2000). AFLP mapping and detection of quantitative trait loci (QTLs) for economically important traits in *Pinus sylvestris*: A preliminary study. *Molecular Breeding, 6*(5), 451–458.

Leslie, A. B., Beaulieu, U. J. M., Rai, H. S., Crane, P. R., Donoghue, M. J., & Mathews, S. (2012). Hemisphere-scale differences in conifer evolutionary dynamics. *Proceedings of the National Academy of Sciences of the United States of America, 109*(40), 16217–16221.

Lesney, M. S. (1989). Growth responses and lignin production in cell suspensions of *Pinus elliottii* 'elicited' by chitin, chitosan or mycelium of *Cronartium quercum* f. sp. *fusiforme*. *Plant Cell, Tissue and Organ Culture, 19*(1), 23–31.

Lesser, M. R., Parchman, T. L., & Jackson, S. T. (2013). Development of genetic diversity, differentiation and structure over 500 years in four ponderosa pine populations. *Molecular Ecology, 22*(10), 2640–2652.

Lesser, M. R., & Parker, W. H. (2004). Genetic variation in *Picea glauca* for growth and phenological traits from provenance tests in Ontario. *Silvae Genetica, 53*(4–5), 141–148.

Lewandowski, A., & Burczyk, J. (2000). Mating system and genetic diversity in natural populations of European larch (*Larix decidua*) and stone pine (*Pinus cembra*) located at higher elevations. *Silvae Genetica, 49*(3), 158–160.

Lewandowski, A., Burczyk, J., & Mejnartowicz, L. (1995). Genetic structure of English yew (*Taxus baccata* L) in the Wierzchlas reserve: Implications for genetic conservation. *Forest Ecology and Management, 73*(1), 221–227.

Lewandowski, A., Nikkanen, T., & Burczyk, J. (1994). Production of hybrid seed in a seed orchard of two larch species, *Larix sibirica* and *Larix decidua*. *Scandinavian Journal of Forest Research, 9*(1–4), 214–217.

Lewis, I. M. (1908). The behaviour of the chromosomes in *Pinus* and *Thuja*. *Annals of Botany, 22*(88), 529–556.

Lewontin, R. C., & Birch, L. C. (1966). Hybridization as a source of variation for adaptation to new environments. *Evolution, 20*, 315–336.

Li, H. (1953). A reclassification of *Libocedrus* and Cupressaceae. *Journal of the Arnold Arboretum, 34*, 17–35.

Li, M., & Ritchie, G. A. (1999a). Eight hundred years of clonal forestry in China: I. traditional afforestation with Chinese fir (*Cunninghamia lanceolata* (Lamb.) Hook.). *New Forests, 18*, 131–142.

Li, M., & Ritchie, G. A. (1999b). Eight hundred years of clonal forestry in China: II. Mass production of rooted cuttings of Chinese fir (*Cunninghamia lanceolata* (Lamb.) Hook.). *New Forests, 18*, 143–159.

Li, P., & Adams, W. T. (1989). Range-wide patterns of allozyme variation in Douglas-fir (*Pseudotsuga menziesii*). *Canadian Journal of Forest Research, 19*(2), 149–161.

Li, P., Beaulieu, J., Corriveau, A., & Bousquet, J. (1993). Genetic variation in juvenile growth and phenology in a white spruce provenance-progeny test. *Silvae Genetica, 42*(1), 52–60.

Li, X., Wu, H. X., Dillon, S. K., & Southerton, S. G. (2009a). Generation and analysis of expressed sequence tags from six developing xylem libraries in *Pinus radiata* D. Don. *BMC Genomics, 10*(1), 41.

Li, X., Wu, H. X., & Southerton, S. G. (2010). Seasonal reorganization of the xylem transcriptome at different tree ages reveals novel insights into wood formation in *Pinus radiata*. *New Phytologist, 187*(3), 764–776.

Li, X., Wu, H. X., & Southerton, S. G. (2011a). Transcriptome profiling of *Pinus radiata* juvenile wood with contrasting stiffness identifies putative candidate genes involved in microfibril orientation and cell wall mechanics. *BMC Genomics, 12*(1), 480.

Li, X., Wu, H. X., & Southerton, S. G. (2011b). Transcriptome profiling of wood maturation in *Pinus radiata* identifies differentially expressed genes with implications in juvenile and mature wood variation. *Gene, 487*(1), 62–71.

Li, X., Yang, X., & Wu, H. X. (2013). Transcriptome profiling of radiata pine branches reveals new insights into reaction wood formation with implications in plant gravitropism. *BMC Genomics, 14*(1), 768.

Li, Y., Stocks, M., Hemmilä, S., Källman, T., Zhu, H., Zhou, Y., et al. (2009b). Demographic histories of four spruce (*Picea*) species of the Qinghai-Tibetan Plateau and neighboring areas inferred from multiple nuclear loci. *Molecular Biology and Evolution, 27*(5), 1001–1014.

Li, Y., Wilcox, P., Telfer, E., Graham, N., & Stanbra, L. (2016). Association of single nucleotide polymorphisms with form traits in three New Zealand populations of radiata pine in the presence of genotype by environment interactions. *Tree Genetics & Genomes, 12*(4), 63.

Li, Z., Baniaga, A. E., Sessa, E. B., Scascitelli, M., Graham, S. W., Rieseberg, L. H., & Barker, M. S. (2015). Early genome duplications in conifers and other seed plants. *Science Advances, 1*(10), e1501084–e1501084.

Li, Z., Hallingbäck, H. R., Abrahamsson, S., Fries, A., Gull, B. A., Sillanpää, M. J., & García-Gil, M. R. (2014). Functional multi-locus QTL mapping of temporal trends in scots pine wood traits. *G3: Genes, Genomes, Genetics, 4*(12), 2365–2379.

Li, Z., Vikneswaran, G., Li, X., Davis, J. M., & Casella, G. (2012a). Simultaneous SNP identification in association studies with missing data. *Annals of Applied Statistics, 6*, 432–456.

Li, Z., Zou, J., Mao, K., Lin, K., Li, H., Liu, J., Källman, T., & Lascoux, M. (2012b). Population genetic evidence for complex evolutionary histories of four high altitude juniper species in the Qinghai–Tibetan Plateau. *Evolution, 66*(3), 831–845.

Liang, C., Wang, G., Liu, L., Ji, G., Fang, L., Liu, Y., et al. (2007). ConiferEST: An integrated bioinformatics system for data reprocessing and mining of conifer expressed sequence tags (ESTs). *BMC Genomics, 8*(1), 134.

Liepelt, S., Bialozyt, R., & Ziegenhagen, B. (2002). Wind-dispersed pollen mediates postglacial gene flow among refugia. *Proceedings of the National Academy of Sciences, 99*(22), 14590–14594.

Liepelt, S., Cheddadi, R., de Beaulieu, J. L., Fady, B., Gömöry, D., Hussendörfer, E., et al. (2009). Postglacial range expansion and its genetic imprints in *Abies alba* (Mill.)—A synthesis from palaeobotanic and genetic data. *Review of Palaeobotany and Palynology, 153*(1–2), 139–149.

Lieutier, F. (2007). Host resistance to bark beetles and its variations. In *Bark and wood boring insects in living trees in Europe, a synthesis* (pp. 135–180). Dordrecht: Springer.

Lind, B. M., Friedline, C. J., Wegrzyn, J. L., Maloney, P. E., Vogler, D. R., Neale, D. B., & Eckert, A. J. (2017). Water availability drives signatures of local adaptation in whitebark pine (*Pinus albicaulis* Engelm.) across fine spatial scales of the Lake Tahoe Basin, USA. *Molecular Ecology, 26*(12), 3168–3185.

Lind, M., Källman, T., Chen, J., Ma, X. F., Bousquet, J., Morgante, M., et al. (2014). A *Picea abies* linkage map based on SNP markers identifies QTLs for four aspects of resistance to *Heterobasidion parviporum* infection. *PLoS One, 9*(7), e101049.

Lippert, D., Chowrira, S., Ralph, S. G., Zhuang, J., Aeschliman, D., Ritland, C., et al. (2007). Conifer defense against insects: Proteome analysis of Sitka spruce (*Picea sitchensis*) bark induced by mechanical wounding or feeding by white pine weevils (*Pissodes strobi*). *Proteomics, 7*(2), 248–270.

Lippert, D., Zhuang, J., Ralph, S., Ellis, D. E., Gilbert, M., Olafson, R., et al. (2005). Proteome analysis of early somatic embryogenesis in *Picea glauca*. *Proteomics, 5*(2), 461–473.

Lippert, D. N., Ralph, S. G., Phillips, M., White, R., Smith, D., Hardie, D., et al. (2009). Quantitative iTRAQ proteome and comparative transcriptome analysis of elicitor-induced Norway spruce (*Picea abies*) cells reveals elements of calcium signaling in the early conifer defense response. *Proteomics, 9*(2), 350–367.

Lise, Y., Kaya, Z., Isik, F., Sabuncu, R., Kandemir, I., & Onde, S. (2007). The impact of overexploitation on the genetic structure of Turkish red pine (*Pinus brutia* Ten.) populations determined by RAPD markers. *Silva Fennica, 41*(2), 211.

Liston, A., Parker-Defeniks, M., Syring, J. V., Willyard, A., & Cronn, R. (2007). Interspecific phylogenetic analysis enhances intraspecific phylogeographical inference: A case study in *Pinus lambertiana*. *Molecular Ecology, 16*(18), 3926–3937.

Liston, A., Robinson, W. A., Piñero, D., & Alvarez-Buylla, E. R. (1999). Phylogenetics of *Pinus*(Pinaceae) based on nuclear ribosomal DNA internal transcribed spacer regions sequences. *Molecular Phylogenetics and Evolution, 11*, 95–109.

Little, D. P. (2004). Documentation of hybridization between Californian cypresses: *Cupressus macnabiana* × *C. sargentii*. *Systematic Botany, 29*(4), 825–833.

Little, D. P. (2006). Evolution and circumscription of the true cypresses (Cupressaceae: *Cupressus*). *Systematic Botany, 31*(3), 461–480.

Little, E. L. (1953). A natural hybrid spruce in Alaska. *Journal of Forestry, 51*(10), 745–747.

Little, E. L., Jr. (1971). *Atlas of United States Trees. Volume 1. Conifer s and important hardwoods* (Miscellaneous Publication 1146). Washington, DC: U.S. Department of Agriculture, Forest Service. 9 p., illus. [313 maps, folio].

Little, E. L., Jr., & Critchfield, W. B. (1969). *Subdivisions of the genus Pinus* (p. 1144). Washington, DC: USDA Forest Service Miscellaneous Publication.

Liu, J. J., Chan, D., Sturrock, R., & Sniezko, R. A. (2014). Genetic variation and population differentiation of the endochitinase gene family in *Pinus monticola*. *Plant Systematics and Evolution, 300*(6), 1313–1322.

Liu, J. J., & Ekramoddoullah, A. K. (2011). Genomic organization, induced expression and promoter activity of a resistance gene analog (PmTNL1) in western white pine (*Pinus monticola*). *Planta, 233*(5), 1041–1053.

Liu, J. J., Ekramoddoullah, A. K., & Yu, X. (2003b). Differential expression of multiple PR10 proteins in western white pine following wounding, fungal infection and cold-hardening. *Physiologia Plantarum, 119*(4), 544–553.

Liu, J. J., Ekramoddoullah, A. K., & Zamani, A. (2005). A class IV chitinase is up-regulated by fungal infection and abiotic stresses and associated with slow-canker-growth resistance to *Cronartium ribicola* in western white pine (*Pinus monticola*). *Phytopathology, 95*(3), 284–291.

Liu, J. J., Hammett, C., & Sniezko, R. A. (2013b). *Pinus monticola* pathogenesis-related gene PmPR10-2 alleles as defense candidates for stem quantitative disease resistance against white pine blister rust (*Cronartium ribicola*). *Tree Genetics & Genomes, 9*(2), 397–408.

Liu, J. J., Schoettle, A., Sniezko, R., Sturrock, R., Zamany, A., Williams, H., Ha, A., Chan, D., Danchok, B., Savin, D., & Kegley, A. (2016b). Genetic mapping of *Pinus flexilis* major gene (Cr4) for resistance to white pine blister rust using transcriptome-based SNP genotyping. *BMC Genomics, 17*, 753.

Liu, J. J., Sniezko, R., Murray, M., Wang, N., Chen, H., Zamany, A., et al. (2016a). Genetic diversity and population structure of whitebark pine (*Pinus albicaulis* Engelm.) in western North America. *PloS One, 11*(12), e0167986.

Liu, J. J., Sniezko, R. A., & Ekramoddoullah, A. K. (2011b). Association of a novel *Pinus monticola* chitinase gene (PmCh4B) with quantitative resistance to *Cronartium ribicola*. *Phytopathology, 101*(8), 904–911.

Liu, J. J., Sniezko, R. A., Zamany, A., Williams, H., Wang, N., Kegley, A., Savin, D. P., Chen, H., & Sturrock, R. N. (2017). Saturated genic SNP mapping identified functional candidates and selection tools for the *Pinus monticola Cr2* locus controlling resistance to white pine blister rust. *Plant Biotechnology Journal, 15*, 1149–1162. https://doi.org/10.1111/pbi.12705.

Liu, J. J., Sturrock, R. N., & Benton, R. (2013a). Transcriptome analysis of *Pinus monticola* primary needles by RNA-seq provides novel insight into host resistance to *Cronartium ribicola*. *BMC Genomics, 14*(1), 884.

Liu, J. J., & Xiang, Y. (2011). In silico mining and PCR-based approaches to transcription factor discovery in non-model plants: Gene discovery of the WRKY transcription factors in conifers. In *Plant transcription factors* (pp. 21–43). New York: Humana Press.

Liu, J. J., Zamani, A., & Ekramoddoullah, A. K. (2010a). Expression profiling of a complex thaumatin-like protein family in western white pine. *Planta, 231*(3), 637–651.

Liu, J. J., Zamany, A., & Sniezko, R. A. (2013c). Anti-microbial peptide (AMP): Nucleotide variation, gene expression, and host resistance in the white pine blister rust (WPBR) pathosystem. *Planta, 237*(1), 43–54.

Liu, N., Zhu, Y., Wei, Z., Chen, J., Wang, Q., Jian, S., Zhou, D., et al. (2009). Phylogenetic relationships and divergence times of the family Araucariaceae based on the DNA sequences of eight genes. *Chinese Science Bulletin, 54*, 2648–2655.

Liu, W., Thummasuwan, S., Sehgal, S. K., Chouvarine, P., & Peterson, D. G. (2011a). Characterization of the genome of bald cypress. *BMC Genomics, 12*(1), 553.

Liu, X. Z., Liu, Z., Yang, Y. M., & Zhang, H. Y. (2010b). Production of transgenic *Pinus armandii* plants harbouring btCryIII (A) gene. *Biologia Plantarum, 54*(4), 711–714.

Liu, Z. L., Cheng, C., & Li, J. (2012). High genetic differentiation in natural populations of *Pinus henryi* and *Pinus tabuliformis* as revealed by nuclear microsatellites. *Biochemical Systematics and Ecology, 42*, 1–9.

Liu, Z. L., Zhang, D., Hong, D. Y., & Wang, X. R. (2003a). Chromosomal localization of 5S and 18S–5.8 S–25S ribosomal DNA sites in five Asian pines using fluorescence in situ hybridization. *Theoretical and Applied Genetics, 106*(2), 198–204.

Lockwood, J. D., Aleksic, J. M., Zou, J., Wang, J., Liu, J., & Renner, S. S. (2013). A new phylogeny for the genus *Picea* from plastid, mitochrondrial, and nuclear sequences. *Molecular Phylogenetics and Evolution, 69*, 717–727.

Lopez, R., Lopez de Heredia, U., Collada, C., Cano, F. J., Emerson, B. C., Cochard, H., & Gil, L. (2013). Vulnerability to cavitation, hydraulic efficiency, growth and survival in an insular pine (*Pinus canariensis*). *Annals of Botany, 111*, 1167–1179.

Lopez-Upton, J., White, T. L., & Huber, D. A. (2000). Species differences in early growth and rust incidence of loblolly and slash pine. *Forest Ecology and Management, 132*(2/3), 211–222.

Lorenz, W. W., Alba, R., Yu, Y. S., Bordeaux, J. M., Simões, M., & Dean, J. F. (2011). Microarray analysis and scale-free gene networks identify candidate regulators in drought-stressed roots of loblolly pine (*P. taeda* L.). *BMC Genomics, 12*(1), 264.

Lorenz, W. W., Ayyampalayam, S., Bordeaux, J. M., Howe, G. T., Jermstad, K. D., Neale, D. B., et al. (2012). Conifer DBMagic: A database housing multiple de novo transcriptome assemblies for 12 diverse conifer species. *Tree Genetics & Genomes, 8*(6), 1477–1485.

Lorenz, W. W., & Dean, J. F. (2002). SAGE profiling and demonstration of differential gene expression along the axial developmental gradient of lignifying xylem in loblolly pine (*Pinus taeda*). *Tree Physiology, 22*(5), 301–310.

Lorenz, W. W., Sun, F., Liang, C., Kolychev, D., Wang, H., Zhao, X., et al. (2006). Water stress-responsive genes in loblolly pine (*Pinus taeda*) roots identified by analyses of expressed sequence tag libraries. *Tree Physiology, 26*(1), 1–16.

Loveless, M. D., & Hamrick, J. L. (1984). Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics, 15*, 65–95.

Lu, M., Krutovsky, K. V., Nelson, C. D., Koralewski, T. E., Byram, T. D., & Loopstra, C. A. (2016). Exome genotyping, linkage disequilibrium and population structure in loblolly pine (*Pinus taeda* L.). *BMC Genomics, 17*(1), 730.

Lu, M., Krutovsky, K. V., Nelson, C. D., West, J. B., Reilly, N. A., & Loopstra, C. A. (2017). Association genetics of growth and adaptive traits in loblolly pine (*Pinus taeda* L.) using whole-exome-discovered polymorphisms. *Tree Genetics & Genomes, 13*(3), 57.

Lu, P., Colombo, S. J., & Sinclair, R. W. (2007). Cold hardiness of interspecific hybrids between *Pinus strobus* and *P. wallichiana* measured by post-freezing needle electrolyte leakage. *Tree Physiology, 27*(2), 243–250.

Lu, P., & Derbowka, D. (2009). Breeding eastern white pine for blister rust resistance: A review of progress in Ontario. *The Forestry Chronicle, 85*(5), 745–755.

Lu, P., Sinclair, R. W., Boult, T. J., & Blake, S. G. (2005). Seedling survival of *Pinus strobus* and its interspecific hybrids after artificial inoculation of *Cronartium ribicola*. *Forest Ecology and Management., 214*, 344–357.

Lu, X., Xu, H., Li, Z., Shang, H., Adams, R. P., & Mao, K. (2014a). Genetic diversity and conservation implications of four Cupressus species in China as revealed by microsatellite markers. *Biochemical Genetics, 52*(3–4), 181–202.

Lu, Y., Ran, J.-H., Guo, D.-M., Yang, Z.-Y., & Wang, X.-Q. (2014b). Phylogeny and edivergence times of gymnopserms inferred from single-copy nuclear genes. *PLoS One, 9*(9), e107679. https://doi.org/10.1371/journal.pone.0107679.

Lubaretz, O., Fuchs, J., Ahne, R., Meister, A., & Schubert, I. (1996). Karyotyping of three Pinaceae species via fluorescent in situ hybridization and computer-aided chromosome analysis. *Theoretical and Applied Genetics, 92*(3–4), 411–416.

Luo, S., Zou, H., & Liang, S. (2000). Study on the introgressive hybridization between *Pinus hwangshanensis* and *P. massoniana*. *Scientia Silvae Sinicae, 37*(6), 118–122.

MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography* (Vol. 1). Princeton: Princeton University Press.

MacDonald, G. M., & Cwynar, L. C. (1985). A fossil pollen based reconstruction of the late Quaternary history of lodgepole pine (*Pinus contorta* ssp. *latifolia*) in the western interior of Canada. *Canadian Journal of Forest Research, 15*(6), 1039–1044.

Macdonald, S. E., Thomas, B. R., Cherniawsky, D. M., & Purdy, B. G. (2001). Managing genetic resources of lodgepole pine in west-central Alberta: Patterns of isozyme variation in natural populations and effects of forest management. *Forest Ecology and Management, 152*(1–3), 45–58.

MacKay, J. J., Liu, W., Whetten, R., Sederoff, R. R., & O'Malley, D. M. (1995). Genetic analysis of cinnamyl alcohol dehydrogenase in loblolly pine: Single gene inheritance, molecular characterization and evolution. *Molecular and General Genetics MGG, 247*(5), 537–545.

MacKay, J. J., O'Malley, D. M., Presnell, T., Booker, F. L., Campbell, M. M., Whetten, R. W., & Sederoff, R. R. (1997). Inheritance, gene expression, and lignin characterization in a mutant pine deficient in cinnamyl alcohol dehydrogenase. *Proceedings of the National Academy of Sciences, 94*(15), 8255–8260.

Magallón, S., & Sanderson, M. J. (2002). Relationships among seed plants inferred from highly conserved genes: Sorting conflicting phylogenetic signals among ancient lineages. *American Journal of Botany, 89*(12), 1991–2006.

Mageroy, M. H., Parent, G., Germanos, G., Giguère, I., Delvas, N., Maaroufi, H., et al. (2015). Expression of the β-glucosidase gene Pgβglu-1 underpins natural resistance of white spruce against spruce budworm. *Plant Journal, 81*, 68–80.

Mahalovich, M. F. (1995). The role of genetics in improving forest health. In Proc. National Silviculture Workshop, May 8–11, 1995, Mescalero, NM, USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO, General Technical Report 267, pp. 200–207. Available at: www.treesearch.fs.fed.us/pubs/23502

Mahalovich, M. F. (2003). The ponderosa pine breeding, seed orchard and restoration program for the Northern Rockies. USDA Forest Service, Northern Region and Inland Empire Tree Improvement Cooperative.

Mahalovich, M. F. (2004). Early survival, height and blister rust performance of *Pinus albicaulis* (whitebark pine) seed sources from the Inland Empire. Proc. Can. Tree Improv. Assoc., IUFRO 7.01.04 meeting. Kelowna, B.C. July 26-29.

Malloch, D. W., Pirozynski, K. A., & Raven, P. H. (1980). Ecological and evolutionary significance of mycorrhizal symbioses in vascular plants (a review). *Proceedings of the National Academy of Sciences of the United States of America, 77*(4), 2113–2118.

Manel, S., Schwartz, M. K., Luikart, G., & Taberlet, P. (2003). Landscape genetics: Combining landscape ecology and population genetics. *Trends in Ecology & Evolution, 18*(4), 189–197.

Manley, S. A. M. (1972). The occurrence of hybrid swarms of red and black spruces in Central New Brunswick. *Canadian Journal of Forest Research, 2*(4), 381–391.

Mann, I. K., Wegrzyn, J. L., & Rajora, O. P. (2013). Generation, functional annotation and comparative analysis of black spruce (*Picea mariana*) ESTs: An important conifer genomic resource. *BMC Genomics, 14*(1), 702.

Manninen, A.-M., Tarhanen, S., Vuorinen, M., & Kainulainen, P. (2002). Comparing the variation of needle and wood terpenoids in scots pine provenances. *Journal of Chemical Ecology, 28*(1), 211–228.

Mao, K., Milne, R. I., Zhang, L., Peng, Y., Liu, J., Thomas, P., Mill, R. R., & Renner, S. S. (2012). Distribution of living Cupressaceae reflects the breakup of Pangea. *Proceedings of the National Academy of Sciences of the United States of America, 109*(20), 7793–7798.

Mao, K., Hao, G., Liu, J., Adams, R. P., & Milne, R. I. (2010). Diversification and biogeography of *Juniperus* (Cupressaceae): Variable diversification rates and multiple intercontinental dispersals. *New Phytologist, 188*(1), 254–272.

Marguerit, E., Bouffier, L., Chancerel, E., Costa, P., Lagane, F., Guehl, J. M., et al. (2014). The genetics of water-use efficiency and its relation to growth in maritime pine. *Journal of Experimental Botany, 65*(17), 4757–4768.

Mariette, S., Chagné, D., Lézier, C., Pastuszka, P., Raffin, A., Plomion, C., & Kremer, A. (2001). Genetic diversity within and among *Pinus pinaster* populations: Comparison between AFLP and microsatellite markers. *Heredity, 86*(4), 469–479.

Markussen, T., Tusch, A., Stephan, B. R., & Fladung, M. (2004). Identification of molecular markers for selected wood properties of Norway spruce *Picea abies* L.(Karst.) I. Wood density. *Silvae Genetica, 53*(1–6), 45–50.

Markussen, T., Tusch, A., Stephan, B. R., & Fladung, M. (2005). Identification of molecular markers for selected wood properties of Norway spruce *Picea abies* L.(Karst.) II. Extractives content. *Silvae Genetica, 54*(1–6), 145–152.

Marquardt, P. E., Echt, C. S., Epperson, B. K., & Pubanz, D. M. (2007). Genetic structure, diversity, and inbreeding of eastern white pine under different management conditions. *Canadian Journal of Forest Research, 37*(12), 2652–2662.

Marquardt, P. E., & Epperson, B. K. (2004). Spatial and population genetic structure of microsatellites in white pine. *Molecular Ecology, 13*(11), 3305–3315.

Marris, E. (2009). Forestry: planting the forest of the future. *Nature News, 459*(7249), 906–908.

Marshall, H. D., Newton, C., & Ritland, K. (2001). Sequence-repeat polymorphisms exhibit the signature of recombination in lodgepole pine chloroplast DNA. *Molecular Biology and Evolution, 18*(11), 2136–2138.

Marshall, H. D., Newton, C., & Ritland, K. (2002). Chloroplast phylogeography and evolution of highly polymorphic microsatellites in lodgepole pine (*Pinus contorta*). *Theoretical and Applied Genetics, 104*(2–3), 367–378.

Marshall, K. A., & Neale, D. B. (1992). The inheritance of mitochondrial DNA in Douglas-fir (*Pseudotsuga menziesii*). *Canadian Journal of Forest Research, 22*(1), 73–75.

Martin, D. M., Fäldt, J., & Bohlmann, J. (2004). Functional characterization of nine Norway spruce TPS genes and evolution of gymnosperm terpene synthases of the TPS-d subfamily. *Plant Physiology, 135*(4), 1908–1927.

Mast, S., Peng, L., Jordan, T. W., Flint, H., Phillips, L., Donaldson, L., et al. (2010). Proteomic analysis of membrane preparations from developing *Pinus radiata* compression wood. *Tree Physiology, 30*(11), 1456–1468.

Matheson, A. C., Bell, J. C., & Barnes, R. D. (1989). Breeding systems and genetic structure in some central American pine populations. *Silvae Genetica, 38*(3), 107–113.

Matheson, A. C., Devey, M. E., Gordon, T. R., Balocchi, C., Carson, M. J., & Werner, W. (2006). The genetics of response to inoculation by pine pitch canker (*Fusarium circinatum* Nirenberg and O'Donnell) infection by seedlings of radiata pine (*Pinus radiata* D.Don). *Australian Forestry, 69*, 101–106.

Matos, J. A., & Schaal, B. A. (2000). Chloroplast evolution in the *Pinus montezumae* complex: A coalescent approach to hybridization. *Evolution, 54*(4), 1218–1233.

Matras, J. (2001). Provenance studies of the Forest research institute on the European larch in 1948-2000. Prace Instytutu Badawczego Lenictwa, Seria A. (908/912): 41–63.

Matyas, C. (1994). Modeling climate change effects with provenance test data. *Tree Physiology, 14*, 797–804.

Mayden, R. L. (1992). *Systematics, historical ecology, and North American freshwater fishes*. Stanford: Stanford University Press.

Mayne, M. B., Subramanian, M., Blake, T. J., Coleman, J. R., & Blumwald, E. (1994). Changes in protein synthesis during drought conditioning in roots of jack pine seedlings (*Pinus banksiana* Lamb.). *Tree Physiology, 14*(5), 509–519.

McDonald, G. I. (1985). Heritability of resistance in Douglas-fir to western spruce budworm. In *The role of the host in the population dynamics of forest insects* (pp. 96–107). Victoria, BC, Canadian Forestry Service, Pacific Forest Res. Centre.

McDonald, G. I., & Hoff, R. J. (1971). Resistance to *Cronartium ribicola* in *Pinus monticola*: Genetic control of needle-spots-only resistance factors. *Canadian Journal of Forest Research, 1*(4), 197–202.

McKenzie, E. A., Elkinton, J. S., Casagrande, R. A., & Mayer, M. (2014). Terpene chemistry of eastern hemlocks resistant to hemlock woolly adelgid. *Journal of Chemical Ecology, 40*(9), 1003–1012.

McLane, S. C., & Aitken, S. N. (2012). Whitebark pine (*Pinus albicaulis*) assisted migration potential: Testing establishment north of the species range. *Ecological Applications, 22*(1), 142–153.

McLane, S. C., Daniels, L. D., & Aitken, S. N. (2011). Climate impacts on lodgepole pine (*Pinus contorta*) radial growth in a provenance experiment. *Forest Ecology and Management, 262*, 115–123.

Medri, C., Ruas, P. M., Higa, A. R., Murakami, M., & de Fatima Ruas, C. (2003). Effects of forest management on the genetic diversity in a population of *Araucaria angustifolia* (bert.) O Kuntze. *Silvae Genetica, 52*(5–6), 202–205.

Mehes-Smith, M., Nkongolo, K. K., & Kim, N. S. (2011). A comparative cytogenetic analysis of five pine species from North America, *Pinus banksiana*, *P. contorta*, *P. monticola, P. resinosa*, and *P. strobus*. *Plant Systematics and Evolution, 292*(3–4), 153–164.

Meirmans, P. G., Gros-Louis, M. C., Lamothe, M., Perron, M., Bousquet, J., & Isabel, N. (2014). Rates of spontaneous hybridization and hybrid recruitment in co-existing exotic and native mature larch populations. *Tree Genetics & Genomes, 10*(4), 965–975.

Men, L., Yan, S., & Liu, G. (2013). De novo characterization of *Larix gmelinii* (Rupr.) Rupr. transcriptome and analysis of its gene expression induced by jasmonates. *BMC Genomics, 14*, 548.

Meyer-Berthaud, B., Scheckler, S. E., & Wendt, J. (1999). Archaeopteris is the earliest known modern tree. *Nature, 398*, 700–701.

Michelmore, R. W., Paran, I., & Kesseli, R. V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences, 88*(21), 9828–9832.

Michelozzi, M., White, T. L., Squillace, A. E., & Lowe, W. J. (1995). Monoterpene composition and fusiform rust resistance in slash and loblolly pines. *Canadian Journal of Forest Research, 25*(2), 193–197.

Miguel, C., Simões, M., Oliveira, M. M., & Rocheta, M. (2008). Envelope-like retrotransposons in the plant kingdom: Evidence of their presence in gymnosperms (*Pinus pinaster*). *Journal of Molecular Evolution, 67*(5), 517–525.

Miksche, J. P. (1967). Variation in DNA content of several gymnosperms. *Canadian Journal of Genetics and Cytology, 9*(4), 717–722.

Miksche, J. P. (1968). Quantitative study of intraspecific variation of DNA per cell in *Picea glauca* and *Pinus banksiana*. *Canadian Journal of Genetics and Cytology, 10*(3), 590–600.

Millar, C. I., & Critchfield, W. B. (1988). Crossability and relationships of *Pinus muricata* (Pinaceae). *Madrono, 35*(1), 39–53.

Millar, C. I., & Marshall, K. A. (1991). Allozyme variation of Port-Orford-Cedar (*Chamaecyparis lawsoniana*): Implications for genetic conservation. *Forest Science, 37*(4), 1060–1077.

Millar, C. I., Strauss, S. H., Conkle, M. T., & Westfall, R. D. (1988). Allozyme differentiation and biosystematics of the Californian closed-cone pines (*Pinus* subsect. Oocarpae). *Systematic Botany, 13*, 351–370.

Miller, J. T., & Ecroyd, C. E. (1987). Introduced forest trees in New Zealand: Recognition, role and see source. 2. *Pinus contorta*. FRI Bulletin 124/2.

Miller, C. N., Jr. (1977). Mesozoic conifers. *The Botanical Review, 43*(2), 217.

Miller, C. N., Jr. (1988). The Origin of modern conifer families. In C. B. Beck (Ed.), *Origin and evolution of gymnosperms*. New York: Columbia University Press.

Mimura, M., & Aitken, S. N. (2007). Adaptive gradients and isolation-by-distance with postglacial migration in *Picea sitchensis*. *Heredity, 99*(2), 224.

Mirov, N. T. (1956). Composition of turpentine of lodgepole × jack pine hybrids. *Canadian Journal of Botany, 34*(4), 443–457.

Mirov, N. T. (1967). *The genus Pinus*. New York: The Ronald Press Co.

Mishima, K., Fujiwara, T., Iki, T., Kuroda, K., Yamashita, K., Tamura, M., et al. (2014). Transcriptome sequencing and profiling of expressed genes in cambial zone and differentiating xylem of Japanese cedar (*Cryptomeria japonica*). *BMC Genomics, 15*(1), 219.

Mitton, J. B. (1992). The dynamic mating systems of conifers. *New Forests, 6*(1–4), 197–216.

Mitton, J. B., Kreiser, B. R., & Latta, R. G. (2000). Glacial refugia of limber pine (*Pinus flexilis* James) inferred from the population structure of mitochondrial DNA. *Molecular Ecology, 9*(1), 91–97.

Molina, R., & Trappe, J. M. (1984). Mycorrhiza management in bareroot nurseries. In M. L. Duryea & T. D. Landis (Eds.), *Forestry nursery manual: Production of bareroot seedlings* (pp. 211–223). The Netherlands: Springer, Dordrecht.

Moraga-Suazo, P., Orellana, L., Quiroga, P., Balocchi, C., Sanfuentes, E., Whetten, R. W., et al. (2014). Development of a genetic linkage map for Pinus radiata and detection of pitch canker disease resistance associated QTLs. *Trees, 28*(6), 1823–1835.

Moreira, X., Mooney, K. A., Rasmann, S., Petry, W. K., Carrillo-Gavilán, A., Zas, R., & Sampedro, L. (2014). Trade-offs between constitutive and induced defenses drive geographical and climatic clines in pine chemical defenses. *Ecology Letters, 17*(5), 537–546.

Morel, A., Teyssier, C., Trontin, J. F., Eliášová, K., Pešek, B., Beaufour, M., et al. (2014a). Early molecular events involved in *Pinus pinaster* Ait. somatic embryo development under reduced water availability: Transcriptomic and proteomic analyses. *Physiologia Plantarum, 152*(1), 184–201.

Morel, A., Trontin, J. F., Corbineau, F., Lomenech, A. M., Beaufour, M., Reymond, I., et al. (2014b). Cotyledonary somatic embryos of *Pinus pinaster* Ait. most closely resemble fresh, maturing cotyledonary zygotic embryos: Biological, carbohydrate and proteomic analyses. *Planta, 240*(5), 1075–1095.

Moreno-Letelier, A., & Piñero, D. (2009). Phylogeographic structure of *Pinus strobiformis* Engelm. across the Chihuahuan Desert filter-barrier. *Journal of Biogeography, 36*(1), 121–131.

Morgenstern, M. (2011). *Geographic variation in forest trees: Genetic basis and application of knowledge in silviculture*. Vancouver: UBC Press.

Moriguchi, Y., Kita, K., Uchiyama, K., Kuromaru, M., & Tsumura, Y. (2008). Enhanced hybridization rates in a *Larix gmelinii* var. *japonica* × *L. kaempferi* interspecific seed orchard with a single maternal clone revealed by cytoplasmic DNA markers. *Tree Genetics & Genomes, 4*(4), 637–645.

Moriguchi, Y., Uchiyama, K., Ueno, S., Ujino-Ihara, T., Matsumoto, A., Iwai, J., et al. (2016). A high-density linkage map with 2560 markers and its application for the localization of the male-sterile genes *ms3* and *ms4* in *Cryptomeria japonica* D. Don. *Tree Genetics & Genomes, 12*(3), 57.

Moriguchi, Y., Ujino-Ihara, T., Uchiyama, K., Futamura, N., Saito, M., Ueno, S., et al. (2012). The construction of a high-density linkage map for identifying SNP markers that are tightly linked to a nuclear-recessive major gene for male sterility in *Cryptomeria japonica* D. Don. *BMC Genomics, 13*(1), 95.

Morin, R. D., Aksay, G., Dolgosheina, E., Ebhardt, H. A., Magrini, V., Mardis, E. R., et al. (2008). Comparative analysis of the small RNA transcriptomes of *Pinus contorta* and *Oryza sativa*. *Genome Research, 18*(4), 571–584.

Morse, A. M., Nelson, C. D., Covert, S. F., Holliday, A. G., Smith, K. E., & Davis, J. M. (2004). Pine genes regulated by the necrotrophic pathogen *Fusarium circinatum*. *Theoretical and Applied Genetics, 109*(5), 922–932.

Morse, A. M., Whetten, R. W., Dubos, C., & Campbell, M. M. (2009). Post-translational modification of an R2R3-MYB transcription factor by a MAP Kinase during xylem development. *New Phytologist, 183*(4), 1001–1013.

Mosca, E., Eckert, A. J., Di Pierro, E. A., Rocchini, D., La Porta, N., Belletti, P., & Neale, D. B. (2012b). The geographical and environmental determinants of genetic diversity for four alpine conifers of the European Alps. *Molecular Ecology, 21*(22), 5530–5545.

Mosca, E., Eckert, A. J., Liechty, J. D., Wegrzyn, J. L., La Porta, N., Vendramin, G. G., & Neale, D. B. (2012a). Contrasting patterns of nucleotide diversity for four conifers of Alpine European forests. *Evolutionary Applications, 5*(7), 762–775.

Mosca, E., González-Martínez, S. C., & Neale, D. B. (2014). Environmental versus geographical determinants of genetic structure in two subalpine conifers. *New Phytologist, 201*(1), 180–192.

Mosca, E., Gugerli, F., Eckert, A. J., & Neale, D. B. (2016). Signatures of natural selection on *Pinus cembra* and *P. mugo* along elevational gradients in the Alps. *Tree Genetics & Genomes, 12*(1), 9.

Moss, E. H. (1949). Natural pine hybrids in Alberta. *Canadian Journal of Research, 27*(5), 218–229.

Moss, S. J., & Wilson, M. E. J. (1998). Biogeographic implications from the Tertiary palaeogeographic evolution of Sulawesi and Borneo. In R. Hall & J. D. Holloway (Eds.), *Biogeography and geological evolution of SE Asia* (pp. 133–163). Leiden: Backhuys.

Moura, V. P. G., Dvorak, W. S., & Hodge, G. R. (1998). Provenance and family variation of *Pinus oocarpa* grown in the Brazilian cerrado. *Forest Ecology and Management, 109*, 315–322.

Muir, J. A., & Cobb, F. W., Jr. (2005). Infection of radiata and bishop pine by *Mycosphaerella pini* in California. *Canadian Journal of Forest Research, 35*(11), 2529–2538.

Muir, P. S., & Lotan, J. E. (1985). Disturbance history and serotiny in *Pinus contorta* in Western Montana. *Ecology, 66*, 1658–1668.

Mukai, Y., Tazaki, K., Fujii, T., & Yamamoto, N. (1992). Light-independent expression of three photosynthetic genes, cab, rbcS and rbcL, in coniferous plants. *Plant and Cell Physiology, 33*(7), 859–866.

Müller, T., Ensminger, I., & Schmid, K. J. (2012). A catalogue of putative unique transcripts from Douglas-fir (*Pseudotsuga menziesii*) based on 454 transcriptome sequencing of genetically diverse, drought stressed seedlings. *BMC Genomics, 13*(1), 673.

Munger, T. T., & Morris, W. G. (1936). *Growth of Douglas-fir trees of known seed source* (Vol. 537, 40 pp). Washington, DC: US Department of Agriculture Technical Bulletin.

Murray, B. G. (1998). Nuclear DNA amounts in gymnosperms. *Annals of Botany, 82*(1), 3–15.

Mutke, S., Gordo, J., Chambel, M. R., Prada, M. A., Alvarez, D., Iglesias, S., & Gil, L. (2010). Phenotypic plasticity is stronger than adaptative differentiation among Mediterranean stone pine provenances. *Forest Systems, 19*(3), 354–366.

Myburg, H., & Harris, S. A. (1997). Genetic variation across the natural distribution of the south east Asian pine, *Pinus kesiya* Royle ex Gordon (Pinaceae). *Silvae Genetica, 46*(5), 295–301.

Myers, E. R., Chung, M. Y., & Chung, M. G. (2007). Genetic diversity and spatial genetic structure of *Pinus strobus* (Pinaceae) across an island landscape inferred from allozyme and cpDNA markers. *Plant Systematics and Evolution, 264*(1–2), 15–30.

Mylecraine, K. A., Kuser, J. E., Smouse, P. E., & Zimmermann, G. L. (2004). Geographic allozyme variation in Atlantic white-cedar, *Chamaecyparis thyoides* (Cupressaceae). *Canadian Journal of Forest Research, 34*(12), 2443–2454.

Mylecraine, K. A., Kuser, J. E., Zimmerman, G. L., & Smouse, P. E. (2005). Rangewide provenance variation in Atlantic white-cedar (*Chamaecyparis thyoides*): Early survival and growth in New Jersey and North Carolina plantations. *Forest Ecology and Management, 216*, 91–104.

Nadeau, S., Meirmans, P. G., Aitken, S. N., Ritland, K., & Isabel, N. (2016). The challenge of separating signatures of local adaptation from those of isolation by distance and colonization history: The case of two white pines. *Ecology and Evolution, 6*(24), 8649–8664.

Nagata, Y., Tomaki, K., & Chiba, S. (1989). Variation and selection of vole resistance in *Larix gmelini* var. *japonica* and its hybrid. Technical note, Institute for Forest Tree Improvement, Oji Paper Co., Ltd., No.261, Reprinted from Trans. Mtg. Hokkaido Br. Jpn. For. Soc. 37th 1989.

Nairn, C. J., & Haselkorn, T. (2005). Three loblolly pine CesA genes expressed in developing xylem are orthologous to secondary cell wall CesA genes of angiosperms. *New Phytologist, 166*(3), 907–915.

Nakamura, R. R., & Wheeler, N. C. (1992a). Pollen competition and paternal success in Douglas-fir. *Evolution, 46*(3), 846–851.

Nakamura, R. R., & Wheeler, N. C. (1992b). Self-fertility variation and paternal success through outcrossing in Douglas-fir. *Theoretical and Applied Genetics, 83*, 851–854.

Namkoong, G. (1969). Non-optimality of local races. In *Proceedings of the 10th southern forest tree improvement conference* (pp. 149–153). Houston: Texas Forest Service, Texas A & M University.

Namroud, M. C., Beaulieu, J., Juge, N., Laroche, J., & Bousquet, J. (2008). Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. *Molecular Ecology, 17*(16), 3599–3613.

Namroud, M. C., Bousquet, J., Doerksen, T., & Beaulieu, J. (2012). Scanning SNPs from a large set of expressed genes to assess the impact of artificial selection on the undomesticated genetic diversity of white spruce. *Evolutionary Applications, 5*(6), 641–656.

Namroud, M. C., Guillet-Claude, C., Mackay, J., Isabel, N., & Bousquet, J. (2010). Molecular evolution of regulatory genes in spruces from different species and continents: Heterogeneous patterns of linkage disequilibrium and selection but correlated recent demographic changes. *Journal of Molecular Evolution, 70*(4), 371–386.

Nardmann, J., Reisewitz, P., & Werr, W. (2009). Discrete shoot and root stem cell-promoting WUS/WOX5 functions are an evolutionary innovation of angiosperms. *Molecular Biology and Evolution, 26*(8), 1745–1755.

Nasri, N., Bojovic, S., Vendramin, G. G., & Fady, B. (2008). Population genetic structure of the relict Serbian spruce, *Picea omorika*, inferred from plastid DNA. *Plant Systematics and Evolution, 271*(1–2), 1–7.

Nathorst, A. G. (1908). Palaobotanische Mitteilungen, 7: Uber *Palissya, Stachyotaxus* und *Palaeotaxus*. *Kungliga Svenska Veterskapsakademiens Handlingar, 43*: 3–37. (from Rothwell et al. 2012).

Natural Resources Canada. (2017). http://www.nrcan.gc.ca/forests/fire-insects-disturbances/pest-management/13361. Last checked 31 Jan 2018.

Naydenov, K. D., Alexandrov, A., Matevski, V., Vasilevski, K., Naydenov, M. K., Gyuleva, V., Carcaillet, C., Wahid, N., & Kamary, S. (2014). Range-wide genetic structure of maritime pine predates the last glacial maximum: Evidence from nuclear DNA. *Hereditas, 151*(1), 1–13.

Naydenov, K. D., Tremblay, F. M., Alexandrov, A., & Fenton, N. J. (2005). Structure of *Pinus sylvestris* L. populations in Bulgaria revealed by chloroplast microsatellites and terpenes analysis: Provenance tests. *Biochemical Systematics and Ecology, 33*, 1226–1245.

Neale, D. B. (1985). Genetic implications of shelterwood regeneration of Douglas-fir in Southwest Oregon. *Forest Science, 31*(4), 995–1005.

Neale, D. B. (2007). Genomics to tree breeding and forest health. *Current Opinion in Genetics & Development, 17*(6), 539–544.

Neale, D. B., & Adams, W. T. (1985). The mating system in natural and shelterwood stands of Douglas-fir. *Theoretical and Applied Genetics, 71*(2), 201–207.

Neale, D. B., Devey, M. E., Jermstad, K. D., Ahuja, M. R., Alosi, M. C., & Marshall, K. A. (1992). Use of DNA markers in forest tree improvement research. *New Forests, 6*(1–4), 391–407.

Neale, D. B., & Ingvarsson, P. K. (2008). Population, quantitative and comparative genomics of adaptation in forest trees. *Current Opinion in Plant Biology, 11*(2), 149–155.

Neale, D. B., Kinlaw, C. S., & Sewell, M. M. (1994). Genetic mapping and DNA sequencing of the loblolly pine genome. *International Journal of Forest Genetics*.

Neale, D. B., & Kremer, A. (2011). Forest tree genomics: Growing resources and applications. *Nature Reviews Genetics, 12*(2), 111.

Neale, D. B., Marshall, K. A., & Harry, D. E. (1991). Inheritance of chloroplast and mitochondrial DNA in incense-cedar (*Calocedrus decurrens*). *Canadian Journal of Forest Research, 21*(5), 717–720.

Neale, D. B., Marshall, K. A., & Sederoff, R. R. (1989). Chloroplast and mitochondrial DNA are paternally inherited in *Sequoia sempervirens* D. Don Endl. *Proceedings of the National Academy of Sciences, 86*(23), 9347–9349.

Neale, D. B., Martínez-García, P. J., De La Torre, A. R., Montanari, S., & Wei, X. X. (2017b). Tree genome sequencing: Novel insights into plant biology. *Annual Review of Plant Biology, 68*(1), 457–483.

Neale, D. B., McGuire, P. E., Wheeler, N. C., Stevens, K. A., Crepeau, M. W., Cardeno, C., et al. (2017a). The Douglas-fir genome sequence reveals specialization of the photosynthetic apparatus in Pinaceae. *G3: Genes, Genomes, Genetics, 7*(9), 3157–3167.

Neale, D. B., & Savolainen, O. (2004). Association genetics of complex traits in conifers. *Trends in Plant Science, 9*(7), 325–330.

Neale, D. B., & Sederoff, R. R. (1988). Inheritance and evolution of conifer organelle genomes. In *Genetic manipulation of woody plants* (pp. 251–264). Springer US: Boston, Massachusetts.

Neale, D. B., & Sederoff, R. R. (1989). Paternal inheritance of chloroplast DNA and maternal inheritance of mitochondrial DNA in loblolly pine. *Theoretical and Applied Genetics, 77*(2), 212–216.

Neale, D. B., Sewell, M. M., & Brown, G. R. (2002). Molecular dissection of the quantitative inheritance of wood property traits in loblolly pine. *Annals of Forest Science, 59*(5–6), 595–605.

Neale, D. B., Wegrzyn, J. L., Stevens, K. A., Zimin, A. V., Puiu, D., Crepeau, M. W., et al. (2014). Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biology, 15*(3), R59. http://genomebiology.com/2014/15/3/R59.

Neale, D. B., Wheeler, N. C., & Allard, R. W. (1986). Paternal inheritance of chloroplast DNA in Douglas-fir. *Canadian Journal of Forest Research, 16*(5), 1152–1154.

Neale, D. B., & Williams, C. G. (1991). Restriction fragment length polymorphism mapping in conifers and applications to forest genetics and tree improvement. *Canadian Journal of Forest Research, 21*(5), 545–554.

Nei, M. (1972). Genetic distance between populations. *American Naturalist, 106*, 283–292.

Nei, M. (1987). *Molecular evolutionary genetics*. New York: Columbia University Press.

Nei, M. (2005). Selectionism and neutralism in molecular evolution. *Molecular Biology and Evolution, 22*, 2318–2342.

Nei, M. (2013). *Mutation-driven evolution*. Oxford: Oxford University Press.

Nei, M., Suzuki, Y., & Nozawa, M. (2010). The neutral theory of molecular evolution in the genomic era. *Annual Review of Genomics and Human Genetics, 11*, 265–289.

Nelson, C. D., Nance, W. L., & Wagner, D. B. (1994). Chloroplast DNA variation among and within taxonomic varieties of *Pinus caribaea* and *Pinus elliottii*. *Canadian Journal of Forest Research, 24*(2), 424–426.

Neves, L. G., Davis, J. M., Barbazuk, W. B., & Kirst, M. (2013). Whole-exome targeted sequencing of the uncharacterized pine genome. *The Plant Journal, 75*(1), 146–156.

Nevo, E., Beiles, A., & Ben-Shlomo, R. (1984). The evolutionary significance of genetic diversity: Ecological, demographic and life history correlates. In *Evolutionary dynamics of genetic diversity* (pp. 13–213). Berlin/Heidelberg: Springer.

Newton, A. C., Allnutt, T. R., Dvorak, W. S., Del Castillo, R. F., & Ennos, R. A. (2002). Patterns of genetic variation in *Pinus chiapensis*, a threatened Mexican pine, detected by RAPD and mitochondrial DNA RFLP markers. *Heredity, 89*(3), 191.

Nielsen, R. (2005). Molecular signatures of natural selection. *Annual Review of Genetics, 39*, 197–218.

Nielsen, U. B., Kirkeby-Thomsen, A., & Roulund, H. (2002). Genetic variation in resistance to *Dreyfusia nordmannianae* Eckst. infestation in *Abies nordmanniana* (Stev.). *Forest Ecology and Management, 165*, 271–283.

Nienstaedt, H. (1985). Inheritance and correlations of frost injury, growth, flowering, and cone characteristics in white spruce, Picea glauca (Moench) Voss. *Canadian Journal of Forest Research, 15*(3), 498–504.

Niklas, K. J. (1997). *The evolutionary biology of plants*. University of Chicago Press.

Nikles, D. G. (1992, October). Hybrids of forest trees: The bases of hybrid superiority and a discussion of breeding methods. In: Proc. of the International Union of Forest Research Organizations (IUFRO), Conference Resolving Tropical Forest Research Concerns Through

Tree Improvement, Gene Conservation, and Domestication of New Species. Cali, Colombia (Vol. 2, pp. 333–347).

Nikles, D. G. (2000). Experience with some *Pinus* hybrids in Queensland, Australia. In *Hybrid breeding and genetics of forest trees. Proc. Queensland Forest Research Institute/Cooperative Research Center – Sustainable Production Forestry Symposium* (pp. 27–43). Noosa, Queensland, Australia.

Nishiyama, K., Watanabe, A., & Kubota, M. (2002). Discrimination of bark borer-resistant clones of *Cryptomeria japonica* using RAPD markers. *Journal of the Japanese Forestry Society, 84*(4), 262–266.

Niu, S. H., Li, Z. X., Yuan, H. W., Chen, X. Y., Li, Y., & Li, W. (2013). Transcriptome characterization of *Pinus tabuliformis* and evolution of genes in the *Pinus* phylogeny. *BMC Genomics, 14*(1), 263.

Nkongolo, K. K. (1996). Chromosome analysis and DNA homology in three *Picea* species, *P. mariana*, *P. rubens*, and *P. glauca* (Pinaceae). *Plant Systematics and Evolution, 203*(1), 27–40.

Nkongolo, K. K., & Klimaszewska, K. (1995). Cytological and molecular relationships between *Larix decidua*, *L. leptolepis* and *Larix x eurolepis*: identification of species-specific chromosoms and synchronization of mitotic cells. *Theoretical and Applied Genetics, 90*(6), 827–834.

Nowicka, A., Ukalska, J., Simińska, J., & Szyp-Borowska, I. (2013). Characterization and mapping of QTL used in breeding of Scots pine (*Pinus sylvestris* L.). *Folia Forestalia Polonica, Seria A-Forestry, 55*(4), 168–173.

Nystedt, B., Street, N. R., Wetterbom, A., Zuccolo, A., Lin, Y. C., Scofield, D. G., et al. (2013). The Norway spruce genome sequence and conifer genome evolution. *Nature, 497*(7451), 579–584.

O'Connell, L. M., Ritland, K., & Thompson, S. L. (2008). Patterns of post-glacial colonization by western redcedar (*Thuja plicata*, Cupressaceae) as revealed by microsatellite markers. *Botany, 86*(2), 194–203.

O'Malley, D. M., Porter, S., & Sederoff, R. R. (1992). Purification, characterization, and cloning of cinnamyl alcohol dehydrogenase in loblolly pine (*Pinus taeda* L.). *Plant Physiology, 98*(4), 1364–1371.

O'Neill, G. A., Hamann, A., & Wang, T. (2008). Accounting for population variation improves estimates of the impact of climate change on species' growth and distribution. *Journal of Applied Ecology, 45*(4), 1040–1049.

O'Neill, G. A., Stoehr, M., & Jaquish, B. (2014). Quantifying safe seed transfer distance and impacts of tree breeding on adaptation. *Forest Ecology and Management, 328*, 122–130.

Ogilvie, R. T., & Rudloff, E. V. (1968). Chemosystematic studies in the genus *Picea* (Pinaceae). IV. The introgression of white and Engelmann spruce as found along the Bow River. *Canadian Journal of Botany, 46*(7), 901–908.

Oh, E., Hansen, E., & Sniezko, R. A. (2006). Port-Orford-cedar resistant to *Phytophthora lateralis*. *Forest Pathology, 36*(6), 385–394.

Oh, E., & Hansen, E. M. (2007). Histopathology of infection and colonization of susceptible and resistant Port-Orford-Cedar by *Phytophthora lateralis*. *Phytopathology, 97*(6), 684–693.

Ohba, K., Iwakawa, M., Okada, Y., & Murai, M. (1971). Paternal transmission of a plastid anomaly in some reciprocal crosses of Sugi, *Cryptomeria japonica* D. Don. *Silvae Genetica., 20*, 101–107.

Ohri, D., & Khoshoo, T. N. (1986). Genome size in gymnosperms. *Plant Systematics and Evolution, 153*(1), 119–132.

Ohsawa, T., Nishida, M., & Nishida, H. (1991). Structure and affinities of the petrified plants from the Cretaceous of northern Japan and Saghalien IX. A petrified cone of *Sciadopitys* from the Upper Cretaceous of Hokkaido. *Journal of Phytogeography and Taxonomy., 39*, 97–105.

Ohta, T. (2002). Near-neutrality in evolution of genes and gene regulation. *Proceedings of the National Academy of Sciences of the United States of America, 99*(25), 16134–16137. https://doi.org/10.1073/pnas.252626899. PMC 138577. PMID 12461171.

Oku, T., Sugahara, K., & Tomita, G. (1974). Functional development of photosystems I and II in dark-grown pine seedlings. *Plant and Cell Physiology, 15*(1), 175–178.

Old, K. M., Libby, W. J., Russell, J. H., & Eldridge, K. G. (1986). Genetic variability in susceptibility of *Pinus radiata* to western gall rust. *Silvae Genet., 35*, 145–149.

Oleksyn, J., Reich, P. B., Chalupka, W., & Tjoelker, M. G. (1999). Differential above- and belowground biomass accumulation of European *Pinus sylvestris* populations in a 12-year-old provenance experiment. *Scandinavian Journal of Forest Research, 14*(1), 7–17.

Oleksyn, J., Zytkowiak, R., Karolewski, P., Reich, P. B., & Tjoelker, M. G. (2000). Genetic and environmental control of seasonal carbohydrate dynamics in trees of diverse *Pinus sylvestris* populations. *Tree Physiology, 20*, 837–847.

Oleszek, W., Stochmal, A., Karolewski, P., Simonet, A. M., Macias, F. A., & Tavea, A. (2002). Flavonoids from *Pinus sylvestris* needles and their variation in trees of different origin grown for nearly a century at the same area. *Biochemical Systematics and Ecology, 30*, 1011–1022.

Oline, D. K. (2008). Geographic variation in chloroplast haplotypes in the California red fir-noble fir species complex and the status of Shasta red fir. *Canadian Journal of Forest Research, 38*(10), 2705–2710.

Oline, D. K., Mitton, J. B., & Grant, M. C. (2000). Population and subspecific genetic differentiation in the foxtail pine (*Pinus balfouriana*). *Evolution, 54*(5), 1813–1819.

Oten, K. L., Merkle, S. A., Jetton, R. M., Smith, B. C., Talley, M. E., & Hain, F. P. (2014). Understanding and developing resistance in hemlocks to the hemlock woolly adelgid. *Southeastern Naturalist, 13*(6), 147–167.

Owens, J. N. (1967). Chromosome aberrations in Douglas fir. *Canadian Journal of Botany, 45*(10), 1910–1913.

Owens, J. N. & Blake, M. D. (1985). Forest tree seed production. A review of the literature and recommendations for future research. Information Report – Petawawa National Forestry Institute, Canadian Forestry Service No. PI-X-53 pp.vi + 161 pp.

Paiva, J. A. P., Garnier-Géré, P. H., Rodrigues, J. C., Alves, A., Santos, S., Graça, J., et al. (2008). Plasticity of maritime pine (*Pinus pinaster*) wood-forming tissues during a growing season. *New Phytologist, 179*(4), 1180–1194.

Palle, S. R., Seeve, C. M., Eckert, A. J., Cumbie, W. P., Goldfarb, B., & Loopstra, C. A. (2011). Natural variation in expression of genes involved in xylem development in loblolly pine (*Pinus taeda* L.). *Tree Genetics & Genomes, 7*(1), 193–206.

Palle, S. R., Seeve, C. M., Eckert, A. J., Wegrzyn, J. L., Neale, D. B., & Loopstra, C. A. (2013). Association of loblolly pine xylem development gene expression with single-nucleotide polymorphisms. *Tree Physiology, 33*(7), 763–774.

Pandey, M., & Rajora, O. P. (2012). Higher fine-scale genetic structure in peripheral than in core populations of a long-lived and mixed-mating conifer-eastern white cedar (Thuja occidentalis L). *BMC Evolutionary Biology, 12*(1), 1–14.

Panetsos, K. P., Aravanopoulos, F. A., & Scaltsoyiannes, A. (1998). Genetic variation of *Pinus brutia* from islands of the northeastern Aegean Sea. *Silvae Genetica, 47*(2), 115–119.

Pâques, L. E., Sylvestre-Guinot, G., & Delatour, C. (1999). Variabilité clonale de la race polonica du mélèze d'Europe pour la résistance à *Lachnellula willkommii*. *Annals of Forest Science, 56*(2), 155–166.

Parchman, T. L., Geist, K. S., Grahnen, J. A., Benkman, C. W., & Buerkle, C. A. (2010). Transcriptome sequencing in an ecologically important tree species: Assembly, annotation, and marker discovery. *BMC Genomics, 11*(1), 180.

Parchman, T. L., Gompert, Z., Mudge, J., Schilkey, F. D., Benkman, C. W., & Buerkle, C. A. (2012). Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology, 21*(12), 2991–3005.

Parducci, L., Szmidt, A. E., Villani, F., Wang, X. R., & Cherubini, M. (1996). Genetic variation of *Abies alba* in Italy. *Hereditas, 125*(1), 11–18.

Parent, G. J., Giguère, I., Germanos, G., Lamara, M., Bauce, É., & MacKay, J. J. (2017). Insect herbivory (*Choristoneura fumiferana*, Tortricidea) underlies tree population structure (Picea glauca, Pinaceae). *Scientific Reports, 7*, 42273.

Parent, G. J., Raherison, E., Sena, J., & MacKay, J. J. (2015). Chapter two-forest tree genomics: Review of Progress. *Advances in Botanical Research, 74*, 39–92.

Parker, K. C., Hamrick, J. L., Parker, A. J., & Stacy, E. A. (1997). Allozyme diversity in *Pinus virginiana* (Pinaceae): Intraspecific and interspecific comparisons. *American Journal of Botany, 84*(10), 1372–1382.

Parks, C. R., Wendel, J. F., Sewell, M. M., & Qiu, Y. L. (1994). The significance of allozyme variation and introgression in the *Liriodendron tulipifera* complex (Magnoliaceae). *American Journal of Botany, 81*, 878–889.

Parks, M., Cronn, R., & Liston, A. (2009). Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biology, 7*, 84. https://doi.org/10.1186/1741-7007-7-84. http://www.biomedcentral.com/1741-7007/7/84.

Parks, M., Cronn, R., & Liston, A. (2012). Separating the wheat from the chaff: Mitigating the effects of noise in a plastome phylogenomic data set from *Pinus* L. (Pinaceae). *BMC Evolutionary Biology, 2012*(12), 100. http://www.biomedcentral.com/1471-2148/12/100.

Parraguirre, C., Vargas, J., Ramírez, P., Azpíroz, H. S., & Jasso, J. (2002). Estructura de la diversidad Genética en poblaciones naturales de pinus greggü Engelm. *Revista Fitotecnia Mexicana, 25*(3), 279–287.

Pastorino, M. J., & Gallo, L. A. (2009). Preliminary operational genetic management units of a highly fragmented forest tree species of southern South America. *Forest Ecology and Management, 257*(12), 2350–2358.

Patton, R. F. (1972). *A brief conspectus of pathology and genetics of Cronartium ribicola as related to resistance* (p. 681). Biology of Rust Resistance in Forest Trees. Miscellaneous Publication no. 1221. Washington, DC: U.S. Department of Agriculture, Forest Service.

Patzlaff, A., McInnis, S., Courtenay, A., Surman, C., Newman, L. J., Smith, C., et al. (2003a). Characterization of a pine MYB that regulates lignification. *The Plant Journal, 36*(6), 743–754.

Patzlaff, A., Newman, L. J., Dubos, C., Whetten, R. W., Smith, C., McInnis, S., et al. (2003b). Characterisation of PtMYB1, an R2R3-MYB from pine xylem. *Plant Molecular Biology, 53*(4), 597–608.

Pausas, J. G., & Keeley, J. E. (2009). A burning story, the role of fire in the history of life. *Bioscience, 59*, 593–601.

Pavy, N., Boyle, B., Nelson, C., Paule, C., Giguère, I., Caron, S., et al. (2008). Identification of conserved core xylem gene sets: Conifer cDNA microarray development, transcript profiling and computational analyses. *New Phytologist, 180*(4), 766–786.

Pavy, N., Deschênes, A., Blais, S., Lavigne, P., Beaulieu, J., Isabel, N., et al. (2013). The landscape of nucleotide polymorphism among 13,500 genes of the conifer *Picea glauca*, relationships with functions, and comparison with *Medicago truncatula*. *Genome Biology and Evolution, 5*(10), 1910–1925.

Pavy, N., Namroud, M. C., Gagnon, F., Isabel, N., & Bousquet, J. (2012a). The heterogeneous levels of linkage disequilibrium in white spruce genes and comparative analysis with other conifers. *Heredity, 108*(3), 273–284.

Pavy, N., Paule, C., Parsons, L., Crow, J. A., Morency, M. J., Cooke, J., et al. (2005). Generation, annotation, analysis and database integration of 16,500 white spruce EST clusters. *BMC Genomics, 6*(1), 144.

Pavy, N., Pelgas, B., Laroche, J., Rigault, P., Isabel, N., & Bousquet, J. (2012b). A spruce gene map infers ancient plant genome reshuffling and subsequent slow evolution in the gymnosperm lineage leading to extant conifers. *BMC Biology, 10*(1), 84.

Peakall, R., Ebert, D., Scott, L. J., Meagher, P. F., & Offord, C. A. (2003). Comparative genetic study confirms exceptionally low genetic variation in the ancient and endangered relictual conifer, *Wollemia nobilis* (Araucariaceae). *Molecular Ecology, 12*(9), 2331–2343.

Peattie, D. C. (1953). *A natural history of western trees*. Bonanza Books/Crown Publ, New York.

Pedlar, J. H., McKenney, D. W., Aubin, I., Beardmore, T., Beaulieu, J., Iverson, L., O'Neill, G. A., Winder, R. S., & Ste-Marie, C. (2012). Placing forestry in the assisted migration debate. *Bioscience, 62*(9), 835–842.

Pedrick, L. A. (1967). The structure and identification of the chromosomes of *Pinus radiata* D. Don. *Silvae Genetica, 16*, 69–77.

Pedrick, L. A. (1968). Chromosome inversions in *Pinus radiata*. *Silvae Genetica, 17*, 22–26.

Pedrick, L. A. (1970). Chromosome relationships between *Pinus* species. *Silvae Genetica, 19*, 171–180.

Pelgas, B., Beauseigle, S., Acheré, V., Jeandroz, S., Bousquet, J., & Isabel, N. (2006). Comparative genome mapping among *Picea glauca*, *P. mariana*× *P. rubens* and *P. abies*, and correspondence with other Pinaceae. *Theoretical and Applied Genetics, 113*(8), 1371.

Pelgas, B., Bousquet, J., Meirmans, P. G., Ritland, K., & Isabel, N. (2011). QTL mapping in white spruce: Gene maps and genomic regions underlying adaptive traits across pedigrees, years and environments. *BMC Genomics, 12*(1), 145.

Perron, M., & Bousquet, J. (1997). Natural hybridization between black spruce and red spruce. *Molecular Ecology, 6*(8), 725–734.

Persson, T., Andersson, B., & Ericsson, T. (2010). Relationship between autumn cold hardiness and field performance in northern *Pinus sylvestris*. *Silva Fennica, 44*(2), 255–266.

Peter, G., & Neale, D. (2004). Molecular basis for the evolution of xylem lignification. *Current Opinion in Plant Biology, 7*(6), 737–742.

Petit, R. J., Duminil, J., Fineschi, S., Hampe, A., Salvini, D., & Vendramin, G. G. (2005). Invited review: Comparative organization of chloroplast, mitochondrial and nuclear diversity in plant populations. *Molecular Ecology, 14*(3), 689–701.

Pfeiffer, A., Olivieri, A. M., & Morgante, M. (1997). Identification and characterization of microsatellites in Norway spruce (*Picea abies* K.). *Genome, 40*(4), 411–419.

Pharis, R. P., & Morf, W. (1967). Experiments on the precocious flowering of western red cedar and four species of *Cupressus* with gibberellins A3 and A4/A7 mixture. *Canadian Journal of Botany, 45*(9), 1519–1524.

Piesch, R. F., & Stettler, R. F. (1971). The detection of good selfers for haploid induction in Douglas-fir. *Silvae Genetica, 20*(4), 144–148.

Pilger, R. (1926). Klasse Coniferae, pp. 121–403. In A. Engler & R. Prantl (Eds.), *Die natÜrelichen Pflanzenfamilien* (ed. 2, vol. e). Leipsiz: Verlag von Wilhelm Englemann. (in German).

Piovani, P., Leonardi, S., Piotti, A., & Menozzi, P. (2010). Conservation genetics of small relic populations of silver fir (*Abies alba* Mill.) in the northern Apennines. *Plant Biosystems, 144*(3), 683–691.

Pittermann, J., Stuart, S. A., Dawson, T. E., & Moreau, A. (2012). Cenozoic climate change shaped the evolutionary ecophysiology of the Cupressaceae conifers. *Proceedings of the National Academy of Sciences, 109*(24), 9647–9652.

Plessas, M. E., & Strauss, S. H. (1986). Allozyme differentiation among populations, stands, and cohorts in Monterey pine. *Canadian Journal of Forest Research, 16*(6), 1155–1164.

Plomion, C., Costa, P., Bahrman, N., & Frigerio, J. M. (1997). Genetic analysis of needle proteins in maritime pine. 1. Mapping dominant and codominant protein markers assayed on diploid tissue, in a haploid-based genetic map. *Silvae Genetica, 46*, 161–165.

Plomion, C., Durel, C. E., & O'Malley, D. M. (1996). Genetic dissection of height in maritime pine seedlings raised under accelerated growth conditions. *Theoretical and Applied Genetics, 93*(5–6), 849–858.

Plomion, C., O'Malley, D. M., & Durel, C. E. (1995). Genomic analysis in maritime pine (*Pinus pinaster*). Comparison of two RAPD maps using selfed and open-pollinated seeds of the same individual. *Theoretical and Applied Genetics, 90*(7–8), 1028–1034.

Plomion, C., Pionneau, C., Brach, J., Costa, P., & Bailleres, H. (2000). Compression wood-responsive proteins in developing xylem of maritime pine (*Pinus pinaster* Ait.). *Plant Physiology, 123*(3), 959–970.

Pojar, J., & MacKinnon, A. (Eds.). (2004). *The plants of the Pacific Northwest Coast (Revised)*. Vancouver: Lone Pine Publishing.

Ponoy, B., Hong, Y. P., Woods, J., Jaquish, B., & Carlson, J. E. (1994). Chloroplast DNA diversity of Douglas-fir in British Columbia. *Canadian Journal of Forest Research, 24*(9), 1824–1834.

Pooler, M. R., Riedel, L. G., Bentz, S. E., & Townsend, A. M. (2002). Molecular markers used to verify interspecific hybridization between hemlock (Tsuga) species. *Journal of the American Society for Horticultural Science, 127*(4), 623–627.

Porth, I., Hamberger, B., White, R., & Ritland, K. (2011). Defense mechanisms against herbivory in Picea: Sequence evolution and expression regulation of gene family members in the phenylpropanoid pathway. *BMC Genomics, 12*(1), 608.

Pot, D., McMillan, L., Echt, C., Le Provost, G., Garnier-Géré, P., Cato, S., & Plomion, C. (2005). Nucleotide variation in genes involved in wood formation in two pine species. *New Phytologist, 167*(1), 101–112.

Pot, D., Rodrigues, J. C., Rozenberg, P., Chantre, G., Tibbits, J., Cahalan, C., et al. (2006). QTLs and candidate genes for wood properties in maritime pine (*Pinus pinaster* Ait.). *Tree Genetics & Genomes, 2*(1), 10–24.

Potter, K. M. (2006). *Evolutionary history and genetic conservation of Fraser fir (Abies fraseri [Pursh] Poir.)*. Ph.D. dissertation, North Carolina State University, Raleigh, NC.

Potter, K. M., Dvorak, W. S., Crane, B. S., Hipkins, V. D., Jetton, R. M., Whittier, W. A., & Rhea, R. (2008). Allozyme variation and recent evolutionary history of eastern hemlock (*Tsuga canadensis*) in the southeastern United States. *New Forests, 35*(2), 131–145.

Potter, K. M., & Frampton, J. (2003). Genetic variation in Fraser Fir mortality due to *Phytophthora* root rot. In C. R. McKinley (Ed.), *Proceedings 27th Southern Forest Tree Improvement Conference*, Stillwater, Oklahoma, USA, June 24–27, 2003, pp. 72–74.

Potter, K. M., Hipkins, V. D., Mahalovich, M. F., & Means, R. E. (2015). Nuclear genetic variation across the range of ponderosa pine (*Pinus ponderosa*): Phylogeographic, taxonomic and conservation implications. *Tree Genetics & Genomes, 11*(3), 1–23.

Potter, K. M., Jetton, R. M., Bower, A., Jacobs, D. F., Man, G., Hipkins, V. D., & Westwood, M. (2017). Banking on the future: Progress, challenges and opportunities for the genetic conservation of forest trees. *New Forests, 48*, 153–180.

Potter, K. M., Jetton, R. M., Dvorak, W. S., Hipkins, V. D., Rhea, R., & Whittier, W. A. (2012). Widespread inbreeding and unexpected geographic patterns of genetic variation in eastern hemlock (Tsuga canadensis), an imperiled North American conifer. *Conservation Genetics, 13*(2), 475–498.

Prada, E., Alía, R., Climent, J., & Díaz, R. (2014). Seasonal cold hardiness in maritime pine assessed by different methods. *Tree Genetics & Genomes, 10*(3), 689–701.

Price, R. A., Liston, A., & Strauss, S. H. (1998). Phylogeny and systematics of *Pinus*. P. In D. M. Richardson (Ed.), *Ecology and biogeography of Pinus* (pp. 49–68). Cambridge: Cambridge Univ. Press.

Prior, N., Little, S. A., Pirone, C., Gill, J. E., Smith, D., Han, J., et al. (2013). Application of proteomics to the study of pollination drops. *Applications in Plant Sciences, 1*(4), 1300008.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(2), 945–959.

Prunier, J., Gerardi, S., Laroche, J., Beaulieu, J., & Bousquet, J. (2012). Parallel and lineage-specific molecular adaptation to climate in boreal black spruce. *Molecular Ecology, 21*(17), 4270–4286.

Prunier, J., Laroche, J., Beaulieu, J., & Bousquet, J. (2011). Scanning the genome for gene SNPs related to climate adaptation and estimating selection at the molecular level in boreal black spruce. *Molecular Ecology, 20*(8), 1702–1716.

Prunier, J., Pelgas, B., Gagnon, F., Desponts, M., Isabel, N., Beaulieu, J., & Bousquet, J. (2013). The genomic architecture and association genetics of adaptive characters using a candidate SNP approach in boreal black spruce. *BMC Genomics, 14*(1), 368.

Prunier, J., Verta, J. P., & MacKay, J. J. (2016). Conifer genomics and adaptation: At the crossroads of genetic diversity and genome function. *New Phytologist, 209*(1), 44–62.

Prus-Glowacki, W., & Bernard, E. (1994). Allozyme variation in populations of *Pinus sylvestris* L from a 1912 provenance trial in Pulawy (Poland). *Silvae Genetica, 43*(2), 132–137.

Prus-Glowacki, W., & Stephan, B. R. (1994). Genetic variation of *Pinus sylvestris* from Spain in relation to other European populations. *Silvae Genetica, 43*(1), 7–13.

Pye, M. G., Henwood, M. J., & Gadek, P. A. (2009). Differential levels of genetic diversity and divergence among populations of an ancient Australian rainforest conifer, *Araucaria cunninghamii*. *Plant Systematics and Evolution, 277*(3–4), 173–185.

Pyhäjärvi, T., Garcia-Gil, M. R., Knürr, T., Mikkonen, M., Wachowiak, W., & Savolainen, O. (2007). Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. *Genetics, 177*(3), 1713–1724.

Pyhäjärvi, T., Kujala, S. T., & Savolainen, O. (2011). Revisiting protein heterozygosity in plants – nucleotide diversity in allozyme coding genes of conifer *Pinus sylvestris. Tree Genetics and Genomes 7*, 385–397.

Qiu, Z., Wan, L., Chen, T., Wan, Y., He, X., Lu, S., et al. (2013). The regulation of cambial activity in Chinese fir (*Cunninghamia lanceolata*) involves extensive transcriptome remodeling. *New Phytologist, 199*(3), 708–719.

Quencez, C., & Bastien, C. (2001). Genetic variation within and between populations of *Pinus sylvestris* L. (Scots pine) for susceptibility to *Melampsora pinitorqua* Rostr. (pine twist rust). *Heredity, 86*, 36–44.

Quesada, T., Gopal, V., Cumbie, W. P., Eckert, A. J., Wegrzyn, J. L., Neale, D. B., et al. (2010). Association mapping of quantitative disease resistance in a natural population of loblolly pine (*Pinus taeda* L.). *Genetics, 186*(2), 677–686.

Quesada, T., Resende, M. F., Jr., Muñoz, P., Wegrzyn, J. L., Neale, D. B., Kirst, M., et al. (2014). Mapping fusiform rust resistance genes within a complex mating design of loblolly pine. *Forests, 5*(2), 347–362.

Quinn, C. J. (1982). Taxonomy of *Cacrydium*. *Australian Journal of Botany, 30*, 311–320.

Quiñones-Pérez, C. Z., del Socorro González-Elizondo, M., & Wehenkel, C. (2017). Ruling out genetic erosion in *Picea chihuahuana* Martínez. *New Forests, 48*(2), 201–215.

Quiñones-Pérez, C. Z., Sáenz-Romero, C., & Wehenkel, C. (2014a). Genetic diversity and conservation of *Picea chihuahuana* Martínez: A review. *African Journal of Biotechnology, 13*(28), 2786–2795.

Quiñones-Pérez, C. Z., Sáenz-Romero, C., & Wehenkel, C. (2014b). Influence of Neighbouring tree species on AFLP variants of endangered *Picea chihuahuana* Martínez populations on the Sierra Madre occidental, North-Western Mexico. *Polish Journal of Ecology, 62*(1), 55–65.

Radu, R. G., Curtu, L. A., & Spârchez, G. (2014). Genetic diversity of Norway spruce [*Picea abies* (L) Karst.] in Romanian Carpathians. *Annals of Forest Research, 57*, 1):19–1):29.

Rafil, Z. A., Dodd, R. S., & Zavarin, E. (1996). Genetic diversity in foliar terpenoids among natural populations of European black pine. *Biochemical Systematics and Ecology, 24*(4), 326–339.

Raherison, E. S., Giguère, I., Caron, S., Lamara, M., & MacKay, J. J. (2015). Modular organization of the white spruce (*Picea glauca*) transcriptome reveals functional organization and evolutionary signatures. *New Phytologist, 207*(1), 172–187.

Rahi, A. A., Bowling, C., & Simpson, D. (2010). A red pine provenance test in northwest Ontario: 48 year results. *The Forestry Chronicle, 86*(3), 348–353.

Rajora, O. P. (1999). Genetic biodiversity impacts of silvicultural practices and phenotypic selection in white spruce. *Theoretical and Applied Genetics, 99*(6), 954–961.

Rajora, O. P., & Dancik, B. P. (2000). Population genetic variation, structure, and evolution in Engelmann spruce, white spruce, and their natural hybrid complex in Alberta. *Canadian Journal of Botany, 78*(6), 768–780.

Rajora, O. P., DeVerno, L., Mosseler, A., & Innes, D. J. (1998). Genetic diversity and population structure of disjunct Newfoundland and Central Ontario populations of eastern white pine (*Pinus strobus*). *Canadian Journal of Botany, 76*(3), 500–508.

Rajora, O. P., Eckert, A. J., & Zinck, J. W. (2016). Single-locus versus multilocus patterns of local adaptation to climate in eastern white pine (*Pinus strobus*, *Pinaceae*). *PLoS One, 11*(7), e0158691.

Rajora, O. P., & Mosseler, A. (2001). Challenges and opportunities for conservation of forest genetic resources. *Euphytica, 118*(2), 197–212.

Rajora, O. P., Mosseler, A., & Major, J. E. (2000a). Indicators of population viability in red spruce, *Picea rubens*. II genetic diversity, population structure, and mating behavior. *Canadian Journal of Botany, 78*(7), 941–956.

Rajora, O. P., & Pluhar, S. A. (2003). Genetic diversity impacts of forest fires, forest harvesting, and alternative reforestation practices in black spruce (*Picea mariana*). *Theoretical and Applied Genetics, 106*(7), 1203–1212.

Rajora, O. P., Rahman, M. H., Buchert, G. P., & Dancik, B. P. (2000b). Microsatellite DNA analysis of genetic effects of harvesting in old-growth eastern white pine (*Pinus strobus*) in Ontario, Canada. *Molecular Ecology, 9*(3), 339–348.

Rake, A. V., Miksche, J. P., Hall, R. B., & Hansen, K. M. (1980). DNA reassociation kinetics of four conifers. *Canadian Journal of Genetics and Cytology, 22*(1), 69–79.

Ralph, J., MacKay, J. J., Hatfield, R. D., O'Malley, D. M., Whetten, R. W., & Sederoff, R. R. (1997). Abnormal lignin in a loblolly pine mutant. *Science, 277*(5323), 235–239.

Ralph, S. G., Chun, H. J. E., Kolosova, N., Cooper, D., Oddy, C., Ritland, C. E., et al. (2008). A conifer genomics resource of 200,000 spruce (*Picea* spp.) ESTs and 6,464 high-quality, sequence-finished full-length cDNAs for Sitka spruce (*Picea sitchensis*). *BMC Genomics, 9*(1), 484.

Ralph, S. G., Yueh, H., Friedmann, M., Aeschliman, D., Zeznik, J. A., Nelson, C. C., et al. (2006). Conifer defence against insects: Microarray gene expression profiling of Sitka spruce (*Picea sitchensis*) induced by mechanical wounding or feeding by spruce budworms (*Choristoneura occidentalis*) or white pine weevils (*Pissodes strobi*) reveals large-scale changes of the host transcriptome. *Plant, Cell & Environment, 29*(8), 1545–1570.

Ranade, S., Abrahamsson, S., Niemi, J., & García-Gil, M. (2013). *Pinus taeda* cDNA microarray as a tool for candidate gene identification for local red/far-red light adaptive response in *Pinus sylvestris*. *American Journal of Plant Sciences, 4*, 479–493.

Ratnam, W., Rajora, O. P., Finkeldey, R., Aravanopoulos, F., Bouvet, J. M., Vaillancourt, R. E., et al. (2014). Genetic effects of forest management practices: Global synthesis and perspectives. *Forest Ecology and Management, 333*, 52–65.

Raubeson, L. A., & Jansen, R. K. (1992). A rare chloroplast-DNA structural mutation is shared by all conifers. *Biochemical Systematics and Ecology, 20*, 17–24.

Read, R. A. (1980). Genetic variation in seedling progeny of ponderosa pine provenances. *Forest Science, 26*(suppl_2), a0001–z0001.

Reglinski, T., Stavely, F. J. L., & Taylor, J. T. (1998). Induction of phenylalanine ammonia lyase activity and control of *Sphaeropsis sapinea* infection in *Pinus radiata* by 5-chlorosalicylic acid. *Forest Pathology, 28*(3), 153–158.

Rehfeldt, G. E. (1977). Growth and cold hardiness of intervarietal hybrids of Douglas-fir. *Theoretical and Applied Genetics, 50*, 3–15.

Rehfeldt, G. E. (1978). Genetic differentiation of Douglas-fir populations from the Northern Rocky Mountains. *Ecology, 59*(6), 1264–1270.

Rehfeldt, G. E. (1979). Ecological adaptations in Douglas-fir (*Pseudotsuga menziesii* var. *glauca*) populations. I. Idaho and northeast Washington. *Heredity, 43*, 383–397.

Rehfeldt, G. E. (1983). Adaptation of *Pinus contorta* populations to heterogeneous environments in northern Idaho. *Canadian Journal of Forest Research, 13*, 405–411.

Rehfeldt, G. E. (1987). Components of adaptive variation in *Pinus contorta* from the Inland Northwest. USDA For. Serv. Inter-mountain Research Station Research Paper INT-375, 11 pp.

Rehfeldt, G. E. (1988). Ecological genetics of *Pinus contorta* from the Rocky Mountains (USA): A synthesis. *Silvae Genetica, 37*(3–4), 131–135.

Rehfeldt, G. E. (1989). Ecological adaptations in Douglas-fir (*Pseudotsuga menziesii* var. *glauca*): A synthesis. *Forest Ecology and Management, 28*, 203–215.

Rehfeldt, G. E. (1990). Genetic differentiation among populations of *Pinus ponderosa* from the upper Colorado river basin. *Botanical Gazette, 151*(1), 125–137.

Rehfeldt, G. E. (1991). A model of genetic variation for *Pinus ponderosa* in the Inland Northwest (USA): Applications in gene resource management. *Canadian Journal of Forest Research, 21*, 1491–1500.

Rehfeldt, G. E. (1992). Breeding strategies for *Larix occidentalis*: Adaptations to the biotic and abiotic environment in relation to improving growth. *Canadian Journal of Forest Research, 22*, 5–13.

Rehfeldt, G. E. (1994a). Evolutionary genetics, the biological species and the ecology of the Interior cedar-hemlock forests. In *Proceedings of the interior cedar-hemlock-white pine forests: Ecology and management* (pp. 91–100). Pullman: Washington State University.

Rehfeldt, G. E. (1994b). Genetic structure of western red cedar populations in the Interior West. *Canadian Journal of Forest Research, 24*, 670–680.

Rehfeldt, G. E. (1997). Quantitative analyses of the genetic structure of closely related conifers with disparate distributions and demographics: The *Cupressus arizonica* (Cupressaceae) complex. *American Journal of Botany, 84*(2), 190–200.

Rehfeldt, G. E. (1999). Systematics and genetic structure of Washoe pine: Applications in conservation genetics. *Silvae Genetica, 48*, 167–172.

Rehfeldt, G. E., & Gallo, L. A. (2001). Introduction of ponderosa pine and Douglas-fir to Argentina. *New Forests, 21*, 35–44.

Rehfeldt, G. E., Hoff, R. J., & Steinhoff, R. J. (1984). Geographic patterns of genetic variation in *Pinus monticola*. *Botanical Gazette, 145*(2), 229–239.

Rehfeldt, G. E., & Jaquish, B. (2010). Ecological impacts and management strategies for western larch in the face of climate-change. *Mitigation and Adaptation Strategies for Global Change, 15*, 283–306.

Rehfeldt, G. E., Jaquish, B. C., Lopez-Upton, J., Saenz-Romero, C., St Clair, J. B., Leites, L. P., & Joyce, D. G. (2014a). Comparative genetic responses to climate for the varieties of *Pinus ponderosa* and *Pseudotuga menziesii*: Realized climate niches. *Forest Ecology and Management, 324*, 126–137.

Rehfeldt, G. E., Jaquish, B. C., Lopez-Upton, J., Saenz-Romero, C., St Clair, J. B., Leites, L. P., & Joyce, D. G. (2014b). Comparative genetic responses to climate for the varieties of *Pinus ponderosa* and *Pseudotuga menziesii*: Clines in growth potential. *Forest Ecology and Management, 324*, 138–146.

Rehfeldt, G. E., Jaquish, B. C., Lopez-Upton, J., Saenz-Romero, C., St Clair, J. B., Leites, L. P., & Joyce, D. G. (2014c). Comparative genetic responses to climate for the varieties of *Pinus ponderosa* and *Pseudotuga menziesii*: Reforestation. *Forest Ecology and Management, 324*, 147–157.

Rehfeldt, G. E., Tchebakova, N. M., & Parfenova, Y. I. (2004). Genetic responses to climate and climate -change in conifers of the temperate and boreal forests. *Recent Research and Development in Genetics Breeding, 1*, 113–130.

Rehfeldt, G. E., Tchebakova, N. M., Parfenova, Y. I., Wykoff, W. R., Kuzmina, N. A., & Milyutin, L. I. (2002). Intraspecific responses to climate in *Pinus sylvestris*. *Global Change Biology, 8*, 912–929.

Rehfeldt, G. E., Ying, C. C., Spittlehouse, D. L., & Hamilton, D. A., Jr. (1999). Genetic responses to climate in *Pinus contorta*: Niche breadth, climate change and reforestation. *Ecological Monographs, 69*(3), 375–407.

Reid, K. E., Holliday, J. A., Yuen, M., Nguyen, A., Aitken, S. N., & Bohlmann, J. (2013). Sequencing of Sitka spruce (*Picea sitchensis*) cDNA libraries constructed from autumn buds and foliage reveals autumn-specific spruce transcripts. *Tree Genetics & Genomes, 9*(3), 683–691.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology, 24*(17), 4348–4370.

Remington, D. L., & O'Malley, D. M. (2000). Evaluation of major genetic loci contributing to inbreeding depression for survival and early growth in a selfed family of *Pinus taeda*. *Evolution, 54*(5), 1580–1589.

Ren, G. P., Abbott, R. J., Zhou, Y. F., Zhang, L. R., Peng, Y. L., & Liu, J. Q. (2012). Genetic divergence, range expansion and possible homoploid hybrid speciation among pine species in Northeast China. *Heredity, 108*(5), 552–562.

Richardson, B. A., Brunsfeld, S. J., & Klopfenstein, N. B. (2002a). DNA from bird-dispersed seed and wind-disseminated pollen provides insights into postglacial colonization and population genetic structure of whitebark pine (*Pinus albicaulis*). *Molecular Ecology, 11*(2), 215–227.

Richardson, B. A., Klopfenstein, N. B., & Brunsfeld, S. J. (2002b). Assessing Clark's nutcracker seed-caching flights using maternally inherited mitochondrial DNA of whitebark pine. *Canadian Journal of Forest Research, 32*(6), 1103–1107.

Richardson, D. M., & Rundel, P. W. (1998). Ecology and biogeography of *Pinus*: An introduction. In D. M. Richardson (Ed.), *Ecology and biogeography of Pinus*. Cambridge, UK: Cambridge University Press.

Rigault, P., Boyle, B., Lepage, P., Cooke, J. E., Bousquet, J., & MacKay, J. J. (2011). A white spruce gene catalog for conifer genome analyses. *Plant Physiology, 157*(1), 14–28.

Ritchie, G. A. (1991). The commercial use of conifer rooted cuttings in forestry: A world overview. *New Forests, 5*, 247–275.

Ritland, C., Pape, T., & Ritland, K. (2001). Genetic structure of yellow cedar (*Chamaecyparis nootkatensis*). *Canadian Journal of Botany, 79*(7), 822–828.

Ritland, K., Krutovsky, K. V., Tsumura, Y., Pelgas, B., Isabel, N., & Bousquet, J. (2011). Genetic mapping in conifers. In *Genetics, genomics and breeding of conifers* (pp. 196–238).

Ro, D.-K., Arimura, G.-I., Lau, S. Y. W., Piers, E., & Bohlmann, J. (2005). Loblolly pine abietadienol/abietadienal oxidase PtAO is a multifunctional, multi-substrate cytochrome P450 monooxygenase. *Proceedings of the National Academy of Sciences USA, 102*, 8060–8065.

Roach, C. R., Hall, D. E., Zerbe, P., & Bohlmann, J. (2014). Plasticity and evolution of (+)-3-carene synthase and (−)-sabinene synthase functions of a Sitka spruce monoterpene synthase gene family associated with weevil resistance. *Journal of Biological Chemistry, 289*(34), 23859–23869.

Robert, J. A., Madilao, L. L., White, R., Yanchuk, A., King, J., & Bohlmann, J. (2010). Terpenoid metabolite profiling in Sitka spruce identifies association of dehydroabietic acid,(+)-3-carene, and terpinolene with resistance against white pine weevil. *Botany, 88*(9), 810–820.

Robinson, A. R., Dauwe, R., Ukrainetz, N. K., Cullis, I. F., White, R., & Mansfield, S. D. (2009). Predicting the regenerative capacity of conifer somatic embryogenic cultures by metabolomics. *Plant Biotechnology Journal, 7*(9), 952–963.

Robinson, A. R., Ukrainetz, N. K., Kang, K. Y., & Mansfield, S. D. (2007). Metabolite profiling of Douglas-fir (*Pseudotsuga menziesii*) field trials reveals strong environmental and weak genetic variation. *New Phytologist, 174*(4), 762–773.

Robinson, R. M., & Morrison, D. J. (1994). Resistance mechanisms of *Larix occidentalis* (western larch) to *Armillaria ostoyae* in the southern interior of British Columbia. Pages 26–33 in Proceedings of the Eighth International Conference on Root and Butt Rots, August 9-16, 1993, Wik, Sweden and Haikko, Finland. Swedish University of Agricultural Sciences, Uppsala, Sweden.

Robledo-Arnuncio, J. J., Alia, R., & Gil, L. (2004). Increased selfing and correlated paternity in a small population of a predominantly outcrossing conifer, *Pinus sylvestris*. *Molecular Ecology, 13*(9), 2567–2577.

Robledo-Arnuncio, J. J., Navascués, M., González-Martínez, S. C., & Gil, L. (2009). Estimating gametic introgression rates in a risk assessment context: A case study with Scots pine relicts. *Heredity, 103*(5), 385–393.

Roche, L. (1969). A genecological study of the genus *Picea* in British Columbia. *New Phytologist, 68*, 505–554.

Rocheta, M., Cordeiro, J., Oliveira, M., & Miguel, C. (2007). PpRT1: The first complete gypsy-like retrotransposon isolated in *Pinus pinaster*. *Planta, 225*(3), 551–562.

Rockwood, D. L., Huber, D. A., & White, T. L. (2001). Provenance and family variability in slash pine (*Pinus elliottii* var. *elliottii* Engelm.) grown in southern Brazil and northeastern Argentina. *New Forests, 21*, 115–125.

Rogers, D. L., Millar, C. I., & Westfall, R. D. (1999). Fine-scale genetic structure of whitebark pine (*Pinus albicaulis*): Associations with watershed and growth form. *Evolution, 53*(1), 74–90.

Rogstad, S. H., Patton, J. C., & Schaal, B. A. (1988). M13 repeat probe detects DNA minisatellite-like sequences in gymnosperms and angiosperms. *Proceedings of the National Academy of Sciences, 85*(23), 9176–9178.

Roschanski, A. M., Csilléry, K., Liepelt, S., Oddou-Muratorio, S., Ziegenhagen, B., Huard, F., et al. (2016). Evidence of divergent selection for drought and cold tolerance at landscape and local scales in *Abies alba* Mill. in the French Mediterranean Alps. *Molecular Ecology, 25*(3), 776–794.

Rossi, S., & Bousquet, J. (2014). The bud break process and its variation among local populations of boreal black spruce. *Frontiers in Plant Science, 5*, 1–9.

Roth, L. F. (1974). Resistance of ponderosa pine to dwarf mistletoe. *Silvae Genetics, 23*(4), 116–120.

Rothwell, G. W. (1988). Cordaitales. In C. B. Beck (Ed.), *Origin and evolution of gymnopserms*. New York: Columbia University Press.

Rothwell, G. W., Mapes, G., & Hernandez-Castillo, G. R. (2005). *Hanskerpia* gen. nov. and phylogenetic relationships among the most ancient conifers (Voltziales). *Taxon, 54*(3), 733–750.

Rothwell, G. W., Mapes, G., & Mapes, R. H. (1997). Late Paleozoic conifers of North America: Structure, diversity and occurrences. *Review of Palaeobotany and Palynology, 95*, 95–113.

Rothwell, G. W., & Scheckler, S. E. (1988). Biology of ancestral gymnosperms. In C. B. Beck (Ed.), *Origin and evolution of gymnosperms*. New York: Columbia University Press.

Rothwell, G. W. G. M., Stockey, R. A., & Hilton, J. (2012). The seed cone *Eathiestrobus* Gen. Nov.: Fossil evidence for a Jurassic origin of Pinaceae. *American Journal of Botany, 99*(4), 708–720.

Rudolph, T. D., & Yeatman, C. W. (1982). *Genetics of jack pine. (WO-38)*. Washington, DC: U.S. Dept. of Agriculture, Forest Service.

Ruhsam, M., Rai, H. S., Mathews, S., Ross, T. G., Graham, S. W., Raubeson, L. A., Mei, W., et al. (2015). Does complete plastid genome sequencing improve species discrimination and phylogenetic resolution in *Araucaria*. *Molecular Ecology Resources, 15*(5), 1067–1078.

Russell, J. H. (2004). Breeding western redcedar for herbivore resistance. Proc. Can. Tree Improv. Assoc., IUFRO 7.01.04 meeting. Kelowna, BC, July 26–29.

Russell, J. H., Kope, H. H., Ades, P., & Collison, H. (2007). Variation in cedar leaf blight (*Didymascella thujina*) resistance of western redcedar. *Canadian Journal of Forest Research, 37*, 1978–1986.

Russell, J. H., & Krakowski, J. (2012). Geographic variation and adaptation to current and future climates of *Callitropsis nootkatensis* populations. *Canadian Journal of Forest Research, 42*(12), 2118–2129.

Rweyongeza, D. M. (2011). Pattern of genotype-environment interaction in *Picea glauca* (Moench) Voss in Alberta, Canada. *Annals of Forest Science, 68*, 245–253.

Saenz-Romero, C., Beaulieu, J., & Rehfeldt, G. E. (2011). Altitudinal genetic variation among *Pinus patula* populations from Oaxaca, Mexico, in growth chambers simulating global warming temperatures. *Agorciencia, 45*(3), 399–411.

Saenz-Romero, C., Guries, R. P., & Monk, A. I. (2001). Landscape genetic structure of *Pinus banksiana*: Allozyme variation. *Canadian Journal of Botany, 79*(8), 871–878.

Saenz-Romero, C., Rehfeldt, G. E., Soto-Correa, J. C., Agular-Auilar, S., Zamarripa-Morales, V., & Lopez-Upton, J. (2012). Altitudinal genetic variation among *Pinus pseudostrobus* populations from Michoacan, Mexico. Two location shadehouse test results. *Revista Fitotecnia Mexicana, 35*(2), 111–121.

Sáenz-Romero, C., & Tapia-Olivares, B. L. (2003). *Pinus oocarpa* isoenzymatic variation along an altitudinal gradient in Michoacán, México. *Silvae Genetica, 52*(5–6), 237–240.

Sagnard, F., Barberot, C., & Fady, B. (2002). Structure of genetic diversity in *Abies alba* Mill. from southwestern Alps: Multivariate analysis of adaptive and non-adaptive traits for conservation in France. *Forest Ecology and Management, 157*(1–3), 175–189.

Saladin, B., Leslie, A. B., Wüest, R. O., Litsios, G., Conti, E., Salamin, N., & Zimmermann, N. E. (2017). Fossils matter: Improved estimates of divergence times in Pinus reveal older diversification. *BMC Evolutionary Biology, 17*(1), 95.

Sampedro, J., Carey, R. E., & Cosgrove, D. J. (2006). Genome histories clarify evolution of the expansin superfamily: New insights from the poplar genome and pine ESTs. *Journal of Plant Research, 119*(1), 11–21.

Sanchez, M., Ingrouille, M. J., Cowan, R. S., Hamilton, M. A., & Fay, M. F. (2014). Spatial structure and genetic diversity of natural populations of the Caribbean pine, *Pinus caribaea* var. *bahamensis* (Pinaceae), in the Bahaman archipelago. *Botanical Journal of the Linnean Society, 174*(3), 359–383.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, 74*(12), 5463–5467.

Santini, A., Camussi, A., & Raddi, P. (1997). Genetic variability of canker resistance trait in *Cupressus sempervirens* L. progenies. *Journal of Applied Genetics, 38*(4), 453–461.

Sartor, C. F., & Neel, W. W. (1971). Variable susceptibility to *Dioryctria amatella* (Hulst) (Lepidoptera: Phycitidae) among pines in clonal seed orchards. In *Proceeding, 11th conference on Southern Forest Tree Improvement* (pp. 91–94).

Sasaki, M., Okanura, M., Fujisawa, Y., & Nishimura, K. (2003). Progress of the breeding project on resistance to sugi bark midges. *Bulletin of the Forest Tree Breeding Institute, 19*, 1–12.

Sasse, J., Elms, S., & Kube, P. (2009). Genetic resistance in *Pinus radiata* to defoliation by the pine aphid *Essigella californica*. *Australian Forestry, 72*, 25–31.

Sato, Y., Wuli, B., Sederoff, R., & Whetten, R. (2001). Molecular cloning and expression of eight laccase cDNAs in loblolly pine (*Pinus taeda*). *Journal of Plant Research, 114*(2), 147–155.

Savolainen, O. (2013). Genetic variation and fitness: Conservation lessons from pines. *Conservation Genetics, 68*, 27.

Savolainen, O., & Kärkkäinen, K. (1992). Effect of forest management on gene pools. *New Forests, 6*(1–4), 329–345.

Savolainen, O., & Pyhäjärvi, T. (2007). Genomic diversity in forest trees. *Current Opinion in Plant Biology, 10*(2), 162–167.

Savva, Y., Schweingruber, F., Milyutin, L., & Vaganov, E. (2002). Genetic and environmental signals in tree rings from different provenances of *Pinus sylvestris* L. planted in the southern taiga, central Siberia. *Trees, 16*, 313–324.

Sawyer, J. O. (2006). *Northwest California*. Berkeley: University of California Press.

Sax, K. (1923). The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics, 8*(6), 552–560.

Sax, K., & Sax, H. J. (1933). Chromosome number and morphology in the conifers. *Journal of the Arnold Arboretum, 14*(4), 356–375.

Saylor, L. C. (1961). A karyotypic analysis of selected species of *Pinus*. *Silvae Genetica, 10*, 77–83.

Saylor, L. C. (1964). Karyotype Analysis of *Pinus*, Group *Lariciones*. *Silvae Genetica, 13*, 165–170.

Saylor, L. C. (1972). Karyotype analysis of the genus *Pinus*-subgenus *Pinus*. *Silvae Genetica, 21*(5), 155–163.

Scalfi, M., Mosca, E., Di Pierro, E. A., Troggio, M., Vendramin, G. G., Sperisen, C., et al. (2014). Micro- and macro-geographic scale effect on the molecular imprint of selection and adaptation in Norway spruce. *PLoS One, 9*(12), e115499.

Schaberg, P. G., Wilkinson, R. C., Shane, J. B., Donnelly, J. R., & Cali, P. F. (1995). Winter photosynthesis of red spruce from three Vermont seed sources. *Tree Physiology, 15*, 345–350.

Scharpf, R. F., & Roth, L. F. (1992). Resistance of ponderosa pine to western dwarf mistletoe in Central Oregon. USDA Forest Service Res. Paper PSW-RP-208.

Scheckler, S. E. (2003). 2003. Consequences of rapid expansion of late Devonian forests. Session No. 146. *Geological Society of America Abstracts with Programs, 35*(6), 385.

Schemske, D. W., & Lande, R. (1985). The evolution of self-fertilization and inbreeding depression in plants. II empirical observations. *Evolution, 39*, 41–52.

Schlarbaum, S. E., Johnson, L. C., & Tsuchiya, T. (1983). Chromosome studies of *Metasequoia glyptostroboides* and *Taxodium distichum*. *Botanical Gazette, 144*(4), 559–565.

Schlarbaum, S. E., & Tsuchiya, T. (1975a). Chromosome study of giant sequoia, *Sequoiadendron giganteum*. *Silvae Genetica, 24*, 23–26.

Schlarbaum, S. E., & Tsuchiya, T. (1975b). Chromosomes of incense cedar. *Journal of Heredity, 66*(1), 41–42.

Schlarbaum, S. E., & Tsuchiya, T. (1984a). The Chromosomes of *Cunninghamia konishii*, *C. lanceolata*, and *Taiwania cryptomerioides* (Taxodiaceae). *Plant Systematics and Evolution, 145*(3), 169–181.

Schlarbaum, S. E., & Tsuchiya, T. (1984b). Cytotaxonomy and phylogeny in certain species of Taxodiaceae. *Plant Systematics and Evolution, 147*, 29–54.
Schlarbaum, S. T., & Tsuchiya, T. (1976). Chromosome study of Japanese umbrella pine. *Journal of Heredity, 67*(1), 65–67.
Schmidtling, R. C. (1994). Use of provenance tests to predict response to climate change: Loblolly pine and Norway spruce. *Tree Physiology, 14*, 805–817.
Schmidtling, R. C. (2003). *The southern pines during the Pleistocene*. In IV International Conifer Conference, ISHS Acta Horticulturae 615, Wye College, United Kingdom (pp. 203–209).
Schmidtling, R. C., Carroll, E., & LaFarge, T. (1999). Allozyme diversity of selected and natural loblolly pine populations. *Silvae Genetics, 48*(1), 35–45.
Schmidtling, R. C., & Hipkins, V. (1998). Genetic diversity in longleaf pine (*Pinus palustris*): Influence of historical and prehistorical events. *Canadian Journal of Forest Research, 28*(8), 1135–1145.
Schneck, V. (1992). Scots pine breeding at Waldsieversdorf and its impact on management in the Northeastern German lowland. *Silvae Genet., 41*(3), 174–180.
Schoenike, R. E. (1976). Geographical variations in jack pine (*Pinus banksiana*). Univ. Minnesota Agric. Exp. Stn., Tech. Bull, 304, p.47.
Schuster, W. S., & Mitton, J. B. (2000). Paternity and gene dispersal in limber pine (*Pinus flexilis* James). *Heredity, 84*(3), 348–361.
Scott, A. D., Stenz, N. W., Ingvarsson, P. K., & Baum, D. A. (2016). Whole genome duplication in coast redwood (*Sequoia sempervirens*) and its implications for explaining the rarity of polyploidy in conifers. *New Phytologist, 211*(1), 186–193.
Scotti, I., Vendramin, G. G., Matteotti, L. S., Scarponi, C., Sari-Gorla, M., & Binelli, G. (2000). Postglacial recolonization routes for *Picea abies* K in Italy as suggested by the analysis of sequence-characterized amplified region (SCAR) markers. *Molecular Ecology, 9*(6), 699–708.
Sedelnikova, T. S., & Muratova, E. N. (2002). Specific karyological features of Siberian stone pine (*Pinus sibirica* Du Tour) in Western Siberian bogs. *Russian Journal of Ecology, 33*(5), 303–308.
Semerikov, V. L., & Lascoux, M. (1999). Genetic relationship among Eurasian and American Larix species based on allozymes. *Heredity, 83*(1), 62–70.
Semerikov, V. L., Semerikova, S. A., & Polezhaeva, M. A. (2013). Nucleotide diversity and linkage disequilibrium of adaptive significant genes in *Larix* (Pinaceae). *Russian Journal of Genetics, 49*(9), 915–923.
Semerikova, S. A., & Semerikov, V. L. (2014). Molecular phylogenetic analysis of the genus Abies (Pinaceae) based on the nucleotide sequence of chloroplast DNA. *Russian journal of genetics, 50*(1), 7–19.
Semerikova, S. A., Semerikov, V. L., & Lascoux, M. (2011). Post-glacial history and introgression in Abies (Pinaceae) species of the Russian Far East inferred from both nuclear and cytoplasmic markers. *Journal of Biogeography, 38*(2), 326–340.
Senjo, M., Kimura, K., Watano, Y., Ueda, K., & Shimizu, T. (1999). Extensive mitochondrial introgression from *Pinus pumila* to *P. parviflora* var. *pentaphylla* (Pinaceae). *Journal of Plant Research, 112*(1), 97–105.
Setoguchi, H., Osawa, T. A., Pintaud, J.-C., Jaffre, T., & Veillon, J.-M. (1998). Phylogenetic relationships within Araucariaceae based on *rbcL* gene sequences. *American Journal of Botany, 85*(11), 1507–1516.
Sewell, M. M., Bassoni, D. L., Megraw, R. A., Wheeler, N. C., & Neale, D. B. (2000). Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.). I. Physical wood properties. *Theoretical and Applied Genetics, 101*(8), 1273–1281.
Sewell, M. M., Davis, M. F., Tuskan, G. A., Wheeler, N. C., Elam, C. C., Bassoni, D. L., & Neale, D. B. (2002). Identification of QTLs influencing wood property traits in loblolly pine (*Pinus taeda* L.). II. Chemical wood properties. *Theoretical and Applied Genetics, 104*(2–3), 214–222.
Sewell, M. M., & Neale, D. B. (2000). Mapping quantitative traits in forest trees. In *Molecular biology of woody plants* (pp. 407–423). Dordrecht, Netherlands: Springer.
Shapcott, A. (1997). Population genetics of the long-lived Huon pine *Lagarostrobos franklinii*: An endemic Tasmanian temperate rainforest tree. *Biological Conservation, 80*(2), 169–179.

Shea, K. L., & Furnier, G. R. (2002). Genetic variation and population structure in central and isolated populations of balsam fir, *Abies balsamea* (Pinaceae). *American Journal of Botany, 89*(5), 783–791.

Shepherd, M., Cross, M., Dieters, M. J., & Henry, R. (2002). Branch architecture QTL for *Pinus elliottii* var. *elliottii* x *Pinus caribaea* var. *hondurensis* hybrids. *Annals of Forest Science, 59*(5–6), 617–625.

Shepherd, M., Huang, S., Eggler, P., Cross, M., Dale, G., Dieters, M., & Henry, R. (2006). Congruence in QTL for adventitious rooting in *Pinus elliottii* × *Pinus caribaea* hybrids resolves between and within-species effects. *Molecular Breeding, 18*(1), 11–28.

Shi, J., Zhen, Y., & Zheng, R. H. (2010). Proteome profiling of early seed development in *Cunninghamia lanceolata* (Lamb.) Hook. *Journal of Experimental Botany, 61*(9), 2367–2381.

Shibata, F., & Hizume, M. (2008). Comparative FISH karyotype analysis of 11 *Picea* species. *Cytologia (Tokyo), 73*(2), 203.

Shibata, F., Matsusaki, Y., & Hizume, M. (2005). AT-rich sequences containing *Arabidopsis*-type telomere sequence and their chromosomal distribution in *Pinus densiflora*. *Theoretical and Applied Genetics, 110*(7), 1253–1258.

Shifley, S. R., & Moser, W. K. (Eds.). (2016). *Future forests of the northern United States. Gen. Tech. Rep. NRS-151. Newtown Square, PA: U.S* (388 p). Forest Service, Northern Research Station: Department of Agriculture.

Shin, D. I., Podila, G. K., Huang, Y., & Karnosky, D. F. (1994). Transgenic larch expressing genes for herbicides and insect resistance. *Canadian Journal of Forest Research, 24*, 2059–2067.

Shutyaev, A. M., & Giertych, M. (2000). Genetic subdivisions of the range of Scots pine (*Pinus sylvestris* L.) based on a transcontinental provenance experiment. *Silvae Genetica, 49*(3), 137–151.

Sierra-DeGrado, R., Pando, V., Martinez-Surimendi, P., Penalvo, A., Bascones, E., & Moulia, B. (2008). Biomechanical differences in the stem straightening process among *Pinus pinaster* provenances. A new approach for early selection of stem straightness. *Tree Physiology, 28*, 835–846.

Sierra-Lucero, V., McKeand, S. E., Huber, D. A., Rockwood, D. L., & White, T. L. (2002). Performance differences and genetic parameters for four coastal provenances of loblolly pine in the Southeastern United States. *Forest Science, 48*(4), 732–742.

Silim, S. N., & Lavender, D. P. (1994). Seasonal patterns and environmental regulation of frost hardiness in shoots of seedlings of *Thuja plicata*, *Chamaecyparis nootkatensis*, and *Picea glauca*. *Canadian Journal of Botany, 72*(3), 309–316.

Simental-Rodríguez, S. L., Quiñones-Pérez, C. Z., Moya, D., Hernandez-Tecles, E., Lopez-Sanchez, C. A., & Wehenkel, C. (2014). The relationship between species diversity and genetic structure in the rare *Picea chihuahuana* tree species community, Mexico. *PLoS One, 9*(11), e111623.

Simon, J.-P., Bergeron, Y., & Gagnon, D. (1986). Isozyme uniformity in populations of red pine (*Pinus resinosa*) in the Abitibi region, Quebec. *Canadian Journal of Forest Research, 16*, 1133–1135.

Simpson, J. A., & Ades, P. K. (1990). Screening *Pinus radiata* families and clones for diseases and pest insect resistance (review). *Australian Forestry, 53*(3), 194–199.

Sinclair, W. T., Mill, R. R., Gardner, M. F., Woltz, P., Jaffré, T., Preston, J., et al. (2002). Evolutionary relationships of the New Caledonian heterotrophic conifer, Parasitaxus usta (Podocarpaceae), inferred from chloroplast trnL-F intron/spacer and nuclear rDNA ITS2 sequences. *Plant Systematics and Evolution, 233*(1–2), 79–104.

Sinclair, W. T., Morman, J. D., & Ennos, R. A. (1998). Multiple origins for scots pine (*Pinus sylvestris* L) in Scotland: Evidence from mitochondrial DNA variation. *Heredity, 80*(2), 233–240.

Singh, G. (2004). *Plant systematics: An integrated approach*. Enfield: Science Publishers.

Skov, E., & Wellendorf, H. (2000). RAPD markers linked to major genes behind field resistance against the green spruce aphid *Elatobium abietinum* (Walker) in Picea sitchensis (Bong. (Carr.)). *Forest Genetics, 7*(3), 233–246.

Slavov, G. T., Howe, G. T., & Adams, W. T. (2005). Pollen contamination and mating patterns in a Douglas-fir seed orchard as measured by simple sequence repeat markers. *Canadian Journal of Forest Research, 35*(7), 1592–1603.

Smith, D. N., & Devey, M. E. (1994). Occurrence and inheritance of microsatellites in *Pinus radiata*. *Genome, 37*(6), 977–983.

Smith, J. A., Blanchette, R. A., Burnes, T. A., Gillman, J. H., & David, A. J. (2006a). Epicuticular wax and white pine blister rust resistance in resistant and susceptible selections of eastern white pine (*Pinus strobus*). *Phytopathology, 96*(2), 171–177.

Smith, J. A., Blanchette, R. A., Burnes, T. A., Jacobs, J. J., Higgins, L., Witthuhn, B. A., et al. (2006b). Proteomic comparison of needles from blister rust-resistant and susceptible *Pinus strobus* seedlings reveals upregulation of putative disease resistance proteins. *Molecular Plant-Microbe Interactions, 19*(2), 150–160.

Smith, J. A., Blanchette, R. A., Burnes, T. A., Jacobs, J. J., Higgins, L., Witthuhn, B. A., David, A. J., & Gillman, J. H. (2005). Proteomic comparison of needles from blister rust-resistant and susceptible *Pinus strobus* seedlings reveals up-regulation of putative disease resistance proteins. *Molecular Plant Microbe Interactions, 19*(2), 150–160.

Smith, K. E., Morse, A. M., Kayihan, G., Huber, D. A., White, T. L., Nelson, C. D., & Davis, J. M. (2003). Molecular pathology of pitch canker disease. In C. R. McKinley, (Ed.), *Proceedings 27th Southern Forest Tree Improvement Conference* (p. 214), Stillwater, Oklahoma, USA, June 24–27, 2003.

Smith, W. K., & Brewer, C. A. (1994). The adaptive importance of shoot and crown architecture in conifer trees. *The American Naturalist, 143*(3), 528–532.

Sniezko, R. A. (2003). Some considerations for using major gene resistance to *Cronartium ribicola* in *Pinus monticola* in Oregon and Washington. In *Proceedings of the 50th Western International Forest Disease Work Conference* (pp. 54–55), 7–11 Oct 2002, Powell River, BC, Canada.

Sniezko, R. A. (2004). Genetic resistance in Port-Orford-cedar to the non-native root rot pathogen *Phytophthora lateralis*—2003 update. In B. W. Geils (Ed.), *Proceedings of the 51st Western International Forest Disease Work Conference* (pp. 127–131), 18–22 Aug 2003, Grants Pass, OR. Flagstaff, AZ: USDA Forest Service, Rocky Mountain Research Station.

Sniezko, R. A., Bower, A., & Danielson, J. (2000a). A comparison of early field results of white pine blister rust resistance of sugar pine and western white pine. *Hort Technology, 10*, 519–522.

Sniezko, R. A., Bower, A., & Kegley, A. (2004a). Variation in *Cronartium ribicola* field resistance among 13 *Pinus monticola* and 12 *P. lambertiana* families: early results from Happy Camp. In R. A. Sniezko, S. Samman, S. E. Schlarbaum, & H. B. Kriebel (Eds.), Breeding and genetic resources of five-needle pines: genetics, breeding, and adaptability, Proceedings of the IUFRO 2.02.15 Working Party Conference. July 23–27, 2001, Medford, OR. Proceedings RMRS-P-32. USDA Forest Service Rocky Mountain Research Station, Ft. Collins, CO, pp. 203–208.

Sniezko, R. A., Elliott, L., Goheen, D. G., Casavan, K., Hansen, E., Frank, C., & Angwin, P. (2003a). Development of *Phytophthora lateralis* resistant Port-Orford-cedar for restoration in the Pacific Northwest. Poster Presentation, In 2002 North American Forest Biology Workshop, Pullman, Washington.

Sniezko, R. A., Elliott, L., Linn, J., Goheen, D. G., Hansen, E., Casavan, K., Kitzmiller, J., & Rose, D. L. (2000b). *Operational breeding program for resistance to an exotic root disease in Port-Orford-Cedar. Poster Presentation*. Washington, DC: Society of American Foresters National Convention.

Sniezko, R. A., Hamlin, J., & Hansen E. M. (2012b). Operational program to develop *Phytophthora lateralis*-resistant populations of Port-Orford-cedar (*Chamaecyparis lawsoniana*). In Proc. Fourth International Workshop on the Genetics of Host-Parasite Interactions in Forestry: Disease and Insect Resistance in Forest Trees. United States Department of Agriculture, Forest Service Gen. Tech. Rep. PSW-GTR-240, pp. 65–79.

Sniezko, R. A., Hamlin, J., Hansen, E. M., & Lucas, S. (2012a). Nine-year survival of 16 *Phytophthora lateralis* resistant and susceptible Port-Orford-Cedar families in a southern

Oregon field trial. In Proc. Fourth International Workshop on the Genetics of Host-Parasite Interactions in Forestry: Disease and Insect Resistance in Forest Trees. United States Department of Agriculture, Forest Service Gen. Tech. Rep. PSW-GTR-240, pp. 348–355.

Sniezko, R. A., & Hansen, E. (2003). Breeding Port-Orford-cedar for resistance to *Phytophthora lateralis*: Current status & considerations for developing durable resistance. In *2nd International IUFRO Working Party 7.02.09 Meeting. Sept. 30-Oct. 5, 2001, Albany, W. Australia* (pp. 197–201). Murdoch: Murdoch University Print.

Sniezko, R. A & Hansen, E. M. (2001). Breeding Port-Orford-Cedar for resistance to *Phytophthora lateralis*. Proc. IUFRO 7.02.09 Albany, Australia, 30 Sept-5 Oct.

Sniezko, R. A., & Kegley, A. (2003a). Blister rust resistance experiences in Oregon and Washington: Evolving perspectives. In *Proceedings of the 50th Western International Forest Disease Work Conference* (pp. 111–117), October 7–11, 2002, Powell River, BC, Canada.

Sniezko, R. A., & Kegley, A. (2003b). Blister rust resistance of five-needle pines in Oregon and Washington. In *Proceedings of the Second IUFRO Rusts of Forest Trees Working Party Conference* (pp. 101–112), 19–23 Aug 2002, Yangling, China. Forest Research. 16 (Suppl.): 1–228.

Sniezko, R. A., & Kegley, A. (2005). White pine blister rust resistance in *Pinus monticola* and *P. lambertiana* seedling families following artificial inoculation with two sources of blister rust. Poster Presentation. In Forest Genetics and Tree Breeding in the Age of Genomics: Progress and Future, IUFRO Joint Conference of Division 2, Conference Proceedings. November 1–5, 2005, Charleston, SC. North Carolina State University, p. 420.

Sniezko, R. A., Kegley, A., Danchok, B., & Conklin, D. (2006a). Variation in white pine blister rust resistance among nine seedling families of southwestern white pine – early results. Poster Presentation. Third Rusts of Forest Trees Conference, IUFRO Unit 7.02.05.

Sniezko, R. A., Kegley, A., Danchok, B., & Long, S. (2007). Variation in resistance to white pine blister rust among whitebark pine families from Oregon and Washington – early results and implications for conservation. In Proceedings of Whitebark pine: a Pacific Coast Perspective Conference. August 27–31, 2006, Ashland, OR.

Sniezko, R. A., Kegley, A. J., & Danchok, R. (2008). White pine blister rust resistance in north American, Asian and European species-results from artificial inoculartion trials in Oregon. *Annals of Forest Research, 51*(1), 51–53.

Sniezko, R. A., Kitzmiller, J., Elliott, L., & Hamlin, J. E. (2003b). Breeding for resistance to *Phytophthora lateralis*. In A range-wide assessment of Port-Orford-Cedar (*Chamaecyparis lawsoniana*) on Federal lands. U.S. Department of Agriculture, Forest Service, and U.S. Department of the Interior, Bureau of Land Management, pp. 77–89.

Sniezko, R. A., Kolpak, S. E., Hansen, E. B., Goheen, D. J., Elliott, L. J., & Angwind, P. A. (2006b). Field survival of *Phytophthora lateralis* resistant and susceptible Port-Orford-cedar families after 53 months. In C. M. Brasier & T. Jung (Eds.), *Advances in Research on Phytophthora Diseases of Forest Trees, Proceedings of the Third IUFRO Workshop on Phytophthoras in Forests and Natural Ecosystems, 11–17 September 2004, Freising, Germany* (pp. 104–108). London: Forestry Commission Bulletin, Her Majesty's Stationary Office.

Sniezko, R. A., Mahalovich, M. F., Schoettle, A. W., & Vogler, D. R. (2011). Past and current investigations of the genetic resistance to C*ronartium ribicola* in high-elevation five-needle pines. In R. E. Keane, D. F. Tomback, M. P. Murray, C. M. Smith (Eds.), The future of high-elevation, five-needle white pines in Western North America: Proceedings of the High Five Symposium. 28–30 June 2010; Missoula, MT. Proceedings RMRS-P-63. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. pp. 246–264.

Sniezko, R. A., Mylecraine, K., Elliott, L., Kolpak, S., Reeser, P., & Hansen, E. (2005). Genetic resistance in Port-Orford-cedar to *Phytophthora lateralis*: Survival of seedlings from first orchard seed in greenhouse testing. Poster Presentation, 53rd WIFDWC 2005.

Sniezko, R. A., Samman, S., Schlarbaum, S. E., & Kriebel, H. B., (2004b). Breeding and genetic resources of five-needle pines: growth, adaptability, and pest resistance. Proceedings RMRS-P-32. Fort Collins, CO: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. 259 pages.

Sniezko, R. A., Smith, J., Liu, J. J., & Hamelin, R. C. (2014). Genetic resistance to fusiform rust in southern pines and white pine blister rust in white pines—A contrasting tale of two rust pathosystems—Current status and future prospects. *Forests, 5*(9), 2050–2083.

Sniezko, R. A., Yanchuk, A. D., Kliejunas, J. T., Palmieri, K. M., Alexander, J. M., & Frankel, S. J., tech. coords. (2012c). Proceedings of the fourth international workshop on the genetics of host-parasite interactions in forestry: Disease and insect resistance in forest trees. Gen. Tech. Rep. PSW-GTR-240. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture. p. 372.

Sniezko, R. A., & Zambino, P. (2005). White pine blister rust resistance in North America white pines: current status and research needs. In Forest Genetics and Tree Breeding in the Age of Genomics: Progress and Future, IUFRO Joint Conference of Division 2, Conference Proceedings. November 1–5, 2005, Charleston, SC. North Carolina State University. pp. 285–286.

Sniezko, R. A., Danchok, R., Savin, D. P., Liu, J.-J., & Kegley, A. (2016). Genetic resistance to white pine blister rust in limber pine (*Pinus flexilis*): Major gene resistance in a northern population. *Canadian Journal of Forest Research., 46*(9), 1173–1178. https://doi.org/10.1139/cjfr-2016-0128.

Soltis, D. E., Buggs, R. J., Doyle, J. J., & Soltis, P. S. (2010). What we still don't know about polyploidy. *Taxon, 59*(5), 1387–1403.

Son, D. S., Eom, T. J., Choi, C. O., & Zhang, R. M. (1999). Effects of controlling the pine needle gall midge by salicylic acid content in needles of some *Pinus* spp. *Journal of Korean Forestry Society, 88*, 31–37. (english summary).

Sonesson, J., & Eriksson, G. (2003). Genetic variation in drought tolerance in *Picea abies* seedlings and its relationship to growth in controlled and field environments. *Scandinavian Journal of Forest Research, 18*(1), 7–18.

Sorensen, F. (1969). Embryonic genetic load in coastal Douglas-fir, *Pseudotsuga menziesii* var. *menziesii*. *The American Naturalist, 103*, 389–398.

Sorensen, F., & Miles, R. S. (1974). Self-pollination effects on Douglas-fir and ponderosa pine seeds and seedlings. *Silvae Genetica, 23*, 135–138.

Sorensen, F. C. (1971). Estimate of self-fertility in coastal Douglas-fir from inbreeding studies. *Silvae Genetica, 20*, 115–120.

Sorensen, F. C. (1979). Provenance variation *in Pseudotsuga menziesii* seedlings from the var. *menziesii*-var. *glauca* transition zone in Oregon. *Silvae Genetica, 28*(2–3), 96–103.

Sorensen, F. C. (1983). Geographic variation in seedling Douglas-fir (*Pseudotsuga menziesii)* from the western Siskiyou Mountains of Oregon. *Ecology, 64*, 696–702.

Sorensen, F. C. (1992). *Genetic variation and seed transfer guidelines for lodgepole pine in central Oregon* (Res. Pap. PNW-RP-453). Portland: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, 30p.

Sorensen, F. C. (1994). *Genetic variation and seed transfer guidelines for ponderosa pine in Central Oregon* (Res. Pap. PNW-RP-472). Portland: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, 24p.

Sork, V. L., Aitken, S. N., Dyer, R. J., Eckert, A. J., Legendre, P., & Neale, D. B. (2013). Putting the landscape into the genomics of trees: Approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes, 9*(4), 901–911.

Souza, M. I. F. D., Salgueiro, F., Carnavale-Bottino, M., Félix, D. B., Alves-Ferreira, M., Bittencourt, J. V. M., & Margis, R. (2009). Patterns of genetic diversity in southern and southeastern *Araucaria angustifolia* (Bert.) O Kuntze relict populations. *Genetics and Molecular Biology, 32*(3), 546–556.

St. Clair, J. B. (2006). Genetic variation in fall cold hardiness in coastal Douglas-fir in western Oregon and Washington. *Canadian Journal of Forest Research, 84*, 1110–1121.

St. Clair, J. B., & Howe, G. T. (2007). Genetic maladaptation of coastal Douglas-fir seedlings to future climates. *Global Change Biology, 13*, 1441–1454.

St. Clair, J. B., Mandel, N. L., & Vance-Borland, K. W. (2005). Genecology of Douglas-fir in western Oregon and Washington. *Annals of Botany, 96*(7), 1199–1214.

Staszkiewicz, J. (1993). Variability of *Pinus mugo* x *P. Sylvestris* [Pinaceae] hybrid swarm in the Tisovnica Nature Reserve [Slovakia]. *Polish Botanical Studies*, (05), 33–41.

Stebbins, G. L. (1948). The chromosomes and relationships of *Metasequoia* and *Sequoia*. *Science, 108*(2796), 95–98.

Stebbins, G. L. (1959). The role of hybridization in evolution. *Proceedings of the American Philosophical Society, 103*(2), 231–251.

Steele, C. L., Crock, J., Bohlmann, J., & Croteau, R. (1998). Sesquiterpene synthases from grand fir (*Abies grandis*) comparison of constitutive and wound-induced activities, and cDNA isolation, characterization, and bacterial expression of δ-selinene synthase and γ-humulene synthase. *Journal of Biological Chemistry, 273*(4), 2078–2089.

Stefanovic, S., Jager, M., Deutsch, J., Broutin, J., & Masselot, M. (1998). Phylogenetic relationships of conifers inferred from partial 28S rRNA gene sequences. *American Journal of Botany, 85*, 688–697.

Stefenon, V. M., Steiner, N., Guerra, M. P., & Nodari, R. O. (2009). Integrating approaches towards the conservation of forest genetic resources: A case study of *Araucaria angustifolia*. *Biodiversity and Conservation, 18*(9), 2433–2448.

Steinhoff, R. J. (1980). *Distribution, ecology, silvicultural characteristics, and genetics of the Abies grandis – Abies concolor complex*. Washington, DC: FAO.

Steinhoff, R. J., Joyce, D. G., & Fins, L. (1983). Isozyme variation in *Pinus monticola*. *Canadian Journal of Forest Research, 13*(6), 1122–1132.

Steinitz, O., Robledo-Arnuncio, J. J., & Nathan, R. (2012). Effects of forest plantations on the genetic composition of conspecific native Aleppo pine populations. *Molecular Ecology, 21*(2), 300–313.

Stephan, B. R. (1973). Uber Anfalligkeit und Resistenz von Douglasien-Herkuften gegenuber *Rhabdocline Pseudotsugae*. *Silvae Genetica, 33*, 149–153.

Stephenson, P. G., Harris, N., Cottrell, J. E., Ralph, S. G., Bohlmann, J., & Taylor, G. (2011). A transcriptomic approach to identify genes associated with wood density in *Picea sitchensis*. *Scandinavian Journal of Forest Research, 26*(S11), 82–96.

Stern, K. & Roche, L. (1974). Genetics of forest ecosystems. 330 pp.

Stevens, K. A., Wegrzyn, J. L., Zimin, A., Puiu, D., Crepeau, M., Cardeno, C., et al. (2016). Sequence of the Sugar Pine Megagenome. *Genetics, 204*(4), 1613–1626.

Stevens, P. F. (1994). *The development of biological systematics: Antoine-Laurent de Jussieu, nature, and the natural system*. New York: Columbia University Press.

Stevenson, D. (1991). Flora of the Guianas, Series A: Phanerogams, Fascicle 9, Sections 208 Cycadaceae, 208.1 Zamiaceae, and 211 Podocarpaceae. Koeltz Scientific Books, USA/ Germany.

Stewart, G. H. (2002). *Structure and canopy tree species regeneration requirements in indigenous forests, Westland, New Zealand* (Doc Science Internal Series 66). Wellington: Department of Conservation.

Stewart, J. F., Liu, Y., Tauer, C. G., & Nelson, C. D. (2010). Microsatellite versus AFLP analyses of pre-management introgression levels in loblolly pine (*Pinus taeda* L.) and shortleaf pine (*P. echinata* mill.). *Tree Genetics & Genomes, 6*(6), 853–862.

Stewart, J. F., Tauer, C. G., & Nelson, C. D. (2012). Bidirectional introgression between loblolly pine (*Pinus taeda* L.) and shortleaf pine (*P. echinata* mill.) has increased since the 1950s. *Tree Genetics & Genomes, 8*(4), 725–735.

Stine, M., & Keathley, D. E. (1990). Paternal inheritance of plastids in Engelmann spruce x blue spruce hybrids. *Journal of Heredity, 81*(6), 443–446.

Stine, M., Sears, B. B., & Keathley, D. E. (1989). Inheritance of plastids in interspecific hybrids of blue spruce and white spruce. *Theoretical and Applied Genetics, 78*(6), 768–774.

Stockey, R. A. (1982). The Araucariaceae: An evolutionary perspective. *Review of Palaebotany and Palynology, 37*, 133–154.

Stoehr, M. U., & El-Kassaby, Y. A. (1997). Levels of genetic diversity at different stages of the domestication cycle of interior spruce in British Columbia. *Theoretical and Applied Genetics, 94*(1), 83–90.

Stoehr, M. U., Orvar, B. L., Vo, T. M., Gawley, J. R., Webber, J. E., & Newton, C. H. (1998). Application of a chloroplast DNA marker in seed orchard management evaluations of Douglas-fir. *Canadian Journal of Forest Research, 28*(2), 187–195.

Stonecypher, R. W., Piesch, R. F., Helland, G. G., Chapman, J. G., & Reno, H. J. (1996). Results from genetic tests of selected parents of Douglas-fir (*Pseudotsuga menziesii* [Mirb.] Franco) in an applied tree improvement program. *Forest Science, 42*., Monograph(32), 1–35.

Strasburg, J. L., Sherman, N. A., Wright, K. M., Moyle, L. C., Willis, J. H., & Rieseberg, L. H. (2012). What can patterns of differentiation across plant genomes tell us about adaptation and speciation? *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 367*(1587), 364–373.

Straub, S. C. K., Parks, M., Weitemier, K., Fishbein, M., Cronn, R. C., & Liston, A. (2012). Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *American Journal of Botany, 99*(2), 349–364.

Strauss, S. H., Bousquet, J., Hipkins, V. D., & Hong, Y. P. (1992). Biochemical and molecular genetic markers in biosystematic studies of forest trees. In *Population Genetics of Forest Trees* (pp. 125–158). Dordrecht: Springer.

Strauss, S. H., & Doerksen, A. H. (1990). Restriction fragment analysis of pine phylogeny. *Evolution, 44*, 1081–1096.

Strauss, S. H., Doerksen, A. H., & Byrne, J. R. (1990). Evolutionary relationships of Douglas-fir and its relatives from DNA restriction fragment analysis. *Canadian Journal of Botany, 68*, 1502–1510.

Strauss, S. H., Hong, Y. P., & Hipkins, V. D. (1993). High levels of population differentiation for mitochondrial DNA haplotypes in *Pinus radiata*, *muricata*, and *attenuata*. *Theoretical and Applied Genetics, 86*(5), 605–611.

Strauss, S. H., Palmer, J. D., Howe, G. T., & Doerksen, A. H. (1988). Chloroplast genomes of two conifers lack a large inverted repeat and are extensively rearranged. *Proceedings of the National Academy of Sciences, 85*(11), 3898–3902.

Strauss, S. H., & Tsai, C. H. (1988). Ribosomal gene number variability in Douglas-fir. *Journal of Heredity, 79*(6), 453–458.

Strong, W. L. (2010). *Pinus contorta* var. *yukonensis* var. nov. (Pinaceae) from South–Central Yukon, Canada. *Nordic Journal of Botany, 28*(4), 448–452.

Strong, W. L., & Hills, L. V. (2006). Taxonomy and origin of present-day morphometric variation in *Picea glauca* × *P. engelmannii* seed-cone scales in North America. *Botany, 84*(7), 129–1141.

Strong, W. L., & Hills, L. V. (2013). Holocene migration of lodgepole pine (*Pinus contorta* var. *latifolia)* in Southern Yukon, Canada. *The Holocene, 23*(9), 1340–1349.

Stuart-Rogers, C., & Flavell, A. J. (2001). The evolution of Ty1-copia group retrotransposons in gymnosperms. *Molecular Biology and Evolution, 18*(2), 155–163.

Sturrock, R. (2005). Investigations of Douglas-fir resistance to *Phellinus weirii*. *TICtalk*, a publication of the Forest Genetics Council of British Columbia, *6*(1), 3–7. Available at : www.fgcouncil.bc.ca/tictalk-jun05-web.pdf

Sturrock, R. N., Islam, M. A., & Ekramoddoullah, A. K. M. (2007). Host-pathogen interactions in Douglas-fir seedlings infected by Phellinus sulphurascens. *Phytopathology, 97*, 1406–1414.

Sun, H., Paulin, L., Alatalo, E., & Asiegbu, F. O. (2011). Response of living tissues of *Pinus sylvestris* to the saprotrophic biocontrol fungus *Phlebiopsis gigantea*. *Tree Physiology, 31*(4), 438–451.

Sun, Y., Abbott, R. J., Li, L., Li, L., Zou, J., & Liu, J. (2014). Evolutionary history of Purple cone spruce (*Picea purpurea*) in the Qinghai–Tibet Plateau: Homoploid hybrid origin and Pleistocene expansion. *Molecular Ecology, 23*(2), 343–359.

Sundås-Larsson, A., Svenson, M., Liao, H., & Engström, P. (1998). A homeobox gene with potential developmental control function in the meristem of the conifer *Picea abies*. *Proceedings of the National Academy of Sciences, 95*(25), 15118–15122.

Sutton, B. C. S., Flanagan, D. J., & El-Kassaby, Y. A. (1991). A simple and rapid method for estimating representation of species in spruce seedlots using chloroplast DNA restriction fragment length polymorphism. *Silvae genetica, 40*(3–4), 119–123.

Sutton, B. C. S., Pritchard, S. C., Gawley, J. R., Newton, C. H., & Kiss, G. K. (1994). Analysis of Sitka spruce-interior spruce introgression in British Columbia using cytoplasmic and nuclear DNA probes. *Canadian Journal of Forest Research, 24*(2), 278–285.

Swedjemark, G., Stenlid, J., & Karlsson, B. (1997). Genetic variation among clones of *Picea abies* in resistance to growth of *Heterobasidium annosum*. *Silvae Genet., 46*, 369–374.

Sweet, G. B. (1965). Provenance differences in Pacific Coast Douglas-fir. 1. Seed and seedling characteristics. *Silvae Genetica, 14*, 46–56.

Swenson, N. G., Iida, Y., Howe, R., Wolf, A., Umaña, M. N., Petprakob, K., et al. (2017). Tree co-occurrence and transcriptomic response to drought. *Nature Communications, 8*(1), 1996.

Syring, J., Farrell, K., Businsky, R., Cronn, R., & Liston, A. (2007). Widespread genealogical nonmonomphyly in species of *Pinus* subgenus *Strobus*. *Systematic Biology, 56*(2), 163–181.

Syring, J., Willyard, A., Cronn, R., & Liston, A. (2005). Evolutionary relationships among Pinus (Pinaceae) subsections inferred from multiple low-copy nuclear loci. *American Journal of Botany, 92*, 2086–2100.

Sziklai, O. (1990, August 20–24). Douglas-fir provenance-progeny test at Haney B.C., Canada. P 2.313–2.321. In: *Proceedings of joint meeting Western Forest Genetics Association and IUFRO Working Parties S2.02–05, 06,12 and 14 – Douglas-fir, contorta pine, Sitka spruce, and Abies breeding and genetic resources, Olympia, WA*. Federal Way, WA: Weyerhaeuser Co.

Szmidt, A. E., Aldén, T., & Hällgren, J. E. (1987). Paternal inheritance of chloroplast DNA in *Larix*. *Plant Molecular Biology, 9*(1), 59–64.

Szmidt, A. E., El-Kassaby, Y. A., Sigurgeirsson, A., Alden, T., Lindgren, D., & Hällgren, J. E. (1988b). Classifying seedlots of *Picea sitchensis* and *P. glauca* in zones of introgression using restriction analysis of chloroplast DNA. *Theoretical and Applied Genetics, 76*(6), 841–845.

Szmidt, A. E., Sigurgeirsson, A., Wang, X. R., Haellgren, J. E., & Lindgren, D. (1988a). Genetic relationships among *Pinus* species based on chloroplast DNA polymorphism. In Hällgren, J.-E. (Ed.). *Proc F. Kempe Symposium, Molecular genetics of forest trees*, June 14–16, 1988. Umeå, Sweden, pp. 33–47.

Szmidt, A. E., & Wang, X. R. (1993). Molecular systematics and genetic differentiation of *Pinus sylvestris* (L.) and *P. densiflora* (Sieb. et Zucc.). *Theoretical and Applied Genetics, 86*(2–3), 159–165.

Szmidt, A. E., Wang, X. R., & Changtragoon, S. (1996a). Contrasting patterns of genetic diversity in two tropical pines: *Pinus kesiya* (Royle ex Gordon) and *P. merkusii* (Jungh et De Vriese). *Theoretical and Applied Genetics, 92*(3–4), 436–441.

Szmidt, A. E., Wang, X. R., & Lu, M. Z. (1996b). Empirical assessment of allozyme and RAPD variation in *Pinus sylvestris* (L) using haploid tissue analysis. *Heredity, 76*(4), 412–420.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics, 123*(3), 585–595.

Takhtajan, A. (1986). (T. J. Crovello, Trans.). In A. Cronquist (Ed.). *Floristic regions of the world*. Berkeley: University of California Press.

Tan, W., Blake, T. J., & Boyle, T. J. (1992). Drought tolerance in faster-and slower-growing black spruce (*Picea mariana*) progenies: II. Osmotic adjustment and changes of soluble carbohydrates and amino acids under osmotic stress. *Physiologia Plantarum, 85*(4), 645–651.

Tan, W., Blake, T. J., & Boyle, T. J. (1995). Early selection for drought tolerance and relationship to dry weight partitioning in black spruce families. *Forest Science, 41*(1), 168–180.

Tandre, K., Albert, V. A., Sundås, A., & Engström, P. (1995). Conifer homologues to genes that control floral development in angiosperms. *Plant Molecular Biology, 27*(1), 69–78.

Tandre, K., Svenson, M., Svensson, M. E., & Engström, P. (1998). Conservation of gene structure and activity in the regulation of reproductive organ development of conifers and angiosperms. *The Plant Journal, 15*(5), 615–623.

Tang, S., Dai, W., Li, M., Zhang, Y., Geng, Y., Wang, L., & Zhong, Y. (2008). Genetic diversity of relictual and endangered plant *Abies ziyuanensis* (Pinaceae) revealed by AFLP and SSR markers. *Genetica, 133*(1), 21–30.

Tang, W., Peng, X., & Newton, R. J. (2005). Enhanced tolerance to salt stress in transgenic loblolly pine simultaneously expressing two genes encoding mannitol-1-phosphate dehydrogenase and glucitol-6-phosphate dehydrogenase. *Plant Physiology and Biochemistry, 43*(2), 139–146.

Tang, W., & Tian, Y. (2003). Transgenic loblolly pine (*Pinus taeda* L.) plants expressing a modified delta-endotoxin gene of *Bacillus thuringiensis* with enhanced resistance to *Dendrolimus punctatus* Walker and *Crypyothelea formosicola* Staud. *Journal of Experimental Botany, 54*, 835–844.

Tani, N., Maruyama, K., Tomaru, N., Uchida, K., Araki, M., Tsumura, Y., Yoshimaru, H., & Ohba, K. (2003). Genetic diversity of nuclear and mitochondrial genomes in *Pinus parviflora* Sieb. & Zucc.(Pinaceae) populations. *Heredity, 91*(5), 510–518.

Tauer, C. G., Stewart, J. F., Will, R. E., Lilly, C. J., Guldin, J. M., & Nelson, C. D. (2012). Hybridization leads to loss of genetic integrity in shortleaf pine: Unexpected consequences of pine management and fire suppression. *Journal of Forestry, 110*(4), 216–224.

Taylor, T. N., Taylor, E. L., & Krings, M. (2009). *The biology and evolution of fossil plants* (2nd ed.). Englewood Cliffs: Prentice Hall.

Temel, F., Johnson, G. R., & Adams, W. T. (2005). Early genetic testing of coastal Douglas-fir for Swiss needle cast tolerance. *Canadian Journal of Forest Research, 35*(3), 521–529.

Temel, F., Johnson, G. R., & Stone, J. K. (2004). The relationship between Swiss needle cast symptom severity and level of *Phaeocryptopus gaeumannii* colonization in coastal Douglas-fir (*Pseudotsuga menziesii* var. *menziesii*). *Forest Pathology, 34*, 383–394.

Temesgen, B., Brown, G. R., Harry, D. E., Kinlaw, C. S., Sewell, M. M., & Neale, D. B. (2001). Genetic mapping of expressed sequence tag polymorphism (ESTP) markers in loblolly pine (*Pinus taeda* L.). *Theoretical and Applied Genetics, 102*(5), 664–675.

Terada, K. (1992). Developing strains resistance to the pine needle gall midge (*Thecodiplosis japonensis* Uchida et Inouye) in *Pinus thunbergii* Parl. *Bulletin of the Forest Tree Breeding Institute, 10*, 1–32.

Terrab, A., Paun, O., Talavera, S., Tremetsberger, K., Arista, M., & Stuessy, T. F. (2006). Genetic diversity and population structure in natural populations of Moroccan Atlas cedar (*Cedrus atlantica*; Pinaceae) determined with cpSSR markers. *American Journal of Botany, 93*(9), 1274–1280.

Terry, R. G. (2010). Re-evaluation of morphological and chloroplast DNA variation in *Juniperus osteosperma* Hook and *Juniperus occidentalis* Torr. Little (Cupressaceae) and their putative hybrids. *Biochemical Systematics and Ecology, 38*(3), 349–360.

Terry, R. G., & Adams, R. P. (2015). A molecular re-examination of phylogenetic relationships among *Juniperus, Cupressus,* and the *Hesperocyparis-Callitropsis-Xanthrocyparis* clades of Cupressaceae. *Phytologia, 97*(1), 67–75.

Terry, R. G., Nowak, R. S., & Tausch, R. J. (2000). Genetic variation in chloroplast and nuclear ribosomal DNA in Utah juniper (*Juniperus osteosperma*, Cupressaceae): Evidence for interspecific gene flow. *American Journal of Botany, 87*(2), 250–258.

Terry, R. G., Bartel, J. A., & Adams, R. P. (2012). Phylogenetic relationships among the New World cypresses (Hesperocyparis; Cupressaceae): Evidence from noncoding chloroplast DNA sequences. *Plant Systematics and Evolution, 298*(10), 1987–2000.

Teyssier, C., Maury, S., Beaufour, M., Grondin, C., Delaunay, A., Le Metté, C., et al. (2014). In search of markers for somatic embryo maturation in hybrid larch (*Larix* × *eurolepis*): Global DNA methylation and proteomic analyses. *Physiologia Plantarum, 150*(2), 271–291.

Thavamanikumar, S., Southerton, S. G., Bossinger, G., & Thumma, B. R. (2013). Dissection of complex traits in forest trees—Opportunities for marker-assisted selection. *Tree Genetics & Genomes, 9*(3), 627–639.

The Angiosperm Phylogeny Website. (n.d.). http://www.mobot.org/MOBOT/research/APweb/.

Thiel, D., Nagy, L., Beierkuhnlein, C., Huber, G., Jentsch, A., Konnert, M., & Kreyling, J. (2012). Uniform drought and warming responses in *Pinus nigra* provenances despite specific overall performances. *Forest Ecology and Management, 270*, 200–208.

Thoday, J. M. (1961). Location of polygenes. *Nature, 191*(4786), 368.

Thomas, B. R., Macdonald, S. E., Hicks, M., Adams, D. L., & Hodgetts, R. B. (1999). Effects of reforestation methods on genetic diversity of lodgepole pine: An assessment using microsatellite and randomly amplified polymorphic DNA markers. *Theoretical and Applied Genetics, 98*(5), 793–801.

Thomas, J. C., & Bond, W. J. (1997). Genetic variation in an endangered cedar (*Widdringtonia cedarbergensis*) versus two congeneric species. *South African Journal of Botany, 63*(3), 133–140.

Thomson, A. M., Riddell, C. L., & Parker, W. H. (2009). Boreal forest provenance tests used to predict optimal growth and response to climate change: 2. Black spruce. *Canadian Journal of Forest Research, 39*(1), 143–153.

Tian, S., Luo, L. C., Ge, S., & Zhang, Z. Y. (2008). Clear genetic structure of *Pinus kwangtungensis* (Pinaceae) revealed by a plastid DNA fragment with a novel minisatellite. *Annals of Botany, 102*(1), 69–78.

Toda, R. (1974). Vegetative propagation in relation to Japanese forest tree improvement. *New Zealand J For Sci, 4*, 410–417.

Toda, T., Kurinobu, S., & Sasaki, M. (2002). Growth and survival rate at the seventh year after selecting the progenies of pine trees for resistance to pine wood nematodes by inoculation test. *Journal of the Japanese Forestry Society, 84*(3), 188–192.

Tollefsrud, M. M., Sønstebø, J. H., Brochmann, C., Johnsen, Ø., Skrøppa, T., & Vendramin, G. G. (2009). Combined analysis of nuclear and mitochondrial markers provide new insight into the genetic structure of North European *Picea abies*. *Heredity, 102*(6), 549–562.

Tomback, D. F., Arno, S. F., & Keane, R. E. (Eds.). (2001). *Whitebark pine communities: Ecology and restoration*. Washington, DC: Island Press.

Townrow, J. A. (1967). On *Rissikia* and *Mataia* – Podocarpaceous conifers from the lower Mesozoic of southern lands. *Papers and Proceedings of the Royal Society of Tasmania, 101*, 103–136.

Tsudzuki, J., Nakashima, K., Tsudzuki, T., Hiratsuka, J., Shibata, M., Wakasugi, T., & Sugiura, M. (1992). Chloroplast DNA of black pine retains a residual inverted repeat lacking rRNA genes: nucleotide sequences of trnQ, trnK, psbA, trnI and trnH and the absence of rps16. *Molecular and General Genetics MGG, 232*(2), 206–214.

Tsumura, Y., Suyama, Y., & Yoshimura, K. (2000). Chloroplast DNA inversion polymorphism in populations of *Abies* and *Tsuga*. *Molecular Biology and Evolution, 17*(9), 1302–1312.

Tsumura, Y., Uchiyama, K., Moriguchi, Y., Ueno, S., & Ihara-Ujino, T. (2012). Genome scanning for detecting adaptive genes along environmental gradients in the Japanese conifer, *Cryptomeria japonica*. *Heredity, 109*(6), 349.

Tsutsui, K., Suwa, A., Sawada, K. I., Kato, T., Ohsawa, T. A., & Watano, Y. (2009). Incongruence among mitochondrial, chloroplast and nuclear gene trees in *Pinus* subgenus *Strobus* (Pinaceae). *Journal of Plant Research, 122*(5), 509–521.

Turesson, G. (1922). The species and the variety as ecological units. *Hereditas, 3*, 100–113.

Turnbull, J. W., & Griffin, A. R. (1986). Concept of provenance and its relationship to infraspecific classification in forest trees. In B. T. Styles (Ed.), *Infraspecific classification of wild and cultivated plants*. Oxford: Clarendon Press.

Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science, 313*(5793), 1596–1604.

Uchida, K., Tomaru, N., Tsumura, Y., Takahashi, C., & Ohba, K. (1993). Allozyme variation in plus-tree of Hinoki, *Chamaecyparis obtusa*, selected from artificial stands. *Japan J Breed, 43*(4), 485–494.

Uchiyama, K., Iwata, H., Moriguchi, Y., Ujino-Ihara, T., Ueno, S., Taguchi, Y., et al. (2013). Demonstration of genome-wide association studies for identifying markers for wood property and male strobili traits in *Cryptomeria japonica*. *PLoS One, 8*(11), e79866.

Uchiyama, K., Ujino-Ihara, T., Ueno, S., Taguchi, Y., Futamura, N., Shinohara, K., & Tsumura, Y. (2012). Single nucleotide polymorphisms in *Cryptomeria japonica*: Their discovery and validation for genome mapping and diversity studies. *Tree Genetics & Genomes, 8*(6), 1213–1222.

Ueki, C. (2004). Studies on selection of sugi resistant to the sugi bark borer in the Kansai breeding region. *Bulletin of the Forest Tree Breeding Institute, 20*, 219–292.

Ujino-Ihara, T., Iwata, H., Taguchi, Y., & Tsumura, Y. (2012). Identification of QTLs associated with male strobilus abundance in *Cryptomeria japonica*. *Tree Genetics & Genomes, 8*(6), 1319–1329.

Ukrainetz, N. K., O'Neill, G. A., & Jaquish, B. (2011). Comparison of fixed and focal point seed transfer systems for reforestation and assisted migration: A case study for interior spruce in British Columbia. *Canadian Journal of Forest Research, 41*(7), 1452–1464.

Ukrainetz, N. K., Ritland, K., & Mansfield, S. D. (2008). Identification of quantitative trait loci for wood quality and growth across eight full-sib coastal Douglas-fir families. *Tree Genetics & Genomes, 4*(2), 159–170.

United States Dept. of Agriculture, Forest Service, Forest Health Protection. (2017). https://foresthealth.fs.usda.gov/portal (Last checked April, 2017).

Valcu, C. M., Lalanne, C., Müller-Starck, G., Plomion, C., & Schlink, K. (2008a). Protein polymorphism between 2 *Picea abies* populations revealed by 2-dimensional gel electrophoresis and tandem mass spectrometry. *Journal of Heredity, 99*(4), 364–375.

Valcu, C. M., Lalanne, C., Plomion, C., & Schlink, K. (2008b). Heat induced changes in protein expression profiles of Norway spruce (*Picea abies*) ecotypes from different elevations. *Proteomics, 8*(20), 4287–4302.

Valledor, L., Cañal, M. J., Lenz, C., Rodríguez, R., & Jorrín-Novo, J. V. (2010). Proteome regulation and epigenetic code during *Pinus radiata* needle maturation. *Proteómica, 5*, 153–155.

Valledor, L., Castillejo, M. A., Lenz, C., Rodríguez, R., Canal, M. J., & Jorrín, J. (2008). Proteomic analysis of *Pinus radiata* needles: 2-DE map and protein identification by LC/MS/MS and substitution-tolerant database searching. *Journal of Proteome Research, 7*(7), 2616–2631.

Van Pelt, R. (2001). *Forest giants of North America*. Seattle: University of Washington Press.

van Zonneveld, M., Jarvis, A., Dvorak, W., Lema, G., & Leibing, C. (2009). Climate change impact predictions on *Pinus patula* and *Pinus tecunumanii* populations in Mexico and Central America. *Forest Ecology and Management, 257*, 1566–1576.

Vander Wall, S. B., & Balda, R. P. (1977). Coadaptation of the Clark's Nutcracker and the pinon pine for efficient seed harvest and dispersal. *Ecological Monographs, 47*, 89–111.

Vangestel, C., Vázquez-Lobo, A., Martínez-García, P. J., Ćalić, I., Wegrzyn, J. L., & Neale, D. B. (2016). Patterns of neutral and adaptive genetic diversity across the natural range of sugar pine (*Pinus lambertiana* Dougl.). *Tree Genetics & Genomes, 12*(3), 51.

Vázquez-Lobo, A., Amanda, R., Martínez-García, P. J., Vangestel, C., Wegzryn, J. L., Ćalić, I., Burton, D., Davis, D., Kinloch, B., Vogler, D., & Neale, D. B. (2017). Finding loci associated to partial resistance to white pine blister rust in sugar pine (*Pinus lambertiana* Dougl.). *Tree Genetics & Genomes, 13*(5), 108.

Velasco-Conde, T., Yakovlev, I., Majada, J. P., Aranda, I., & Johnsen, Ø. (2012). Dehydrins in maritime pine (*Pinus pinaster*) and their expression related to drought stress response. *Tree Genetics & Genomes, 8*(5), 957–973.

Vendramin, G. G., Degen, B., Petit, R. J., Anzidei, M., Madaghiele, A., & Ziegenhagen, B. (1999). High level of variation at *Abies alba* chloroplast microsatellite loci in Europe. *Molecular Ecology, 8*(7), 1117–1126.

Verkaik, E., Gardner, R. O., & Braakhekke, W. G. (2007). Site conditions affect seedling distribution below and outside the crown of kauri trees (*Agathis australis*). *New Zealand J Ecol, 31*(1), 13–21.

Verne, S., Jaquish, B., White, R., Ritland, C., & Ritland, K. (2011). Global transcriptome analysis of constitutive resistance to the white pine weevil in spruce. *Genome Biology and Evolution, 3*, 851–867.

Viard, F., El-Kassaby, Y. A., & Ritland, K. (2001). Diversity and genetic structure in populations of *Pseudotsuga menziesii* (Pinaceae) at chloroplast microsatellite loci. *Genome, 44*(3), 336–344.

Vieira, L. N., Faoro, H., Rogalski, M., de Freitas Fraga, H. P., Cardoso, R. L. A., de Souza, E. M., et al. (2014). The complete chloroplast genome sequence of *Podocarpus lambertii*: Genome structure, evolutionary aspects, gene content and SSR detection. *PLoS One, 9*(3), e90618.

Villalobos, D. P., Diaz-Moreno, S. M., Said el, S. S., Canas, R. A., Osuna, D., Van Kerckhoven, S. H., et al. (2012). Reprogramming of gene expression during compression wood formation in pine: Coordinated modulation of S-adenosylmethionine, lignin and lignan related genes. *BMC Plant Biology, 12*, 100.

Vischi, M., Jurman, I., Bianchi, G., & Morgante, M. (2003). Karyotype of Norway spruce by multicolor FISH. *Theoretical and Applied Genetics, 107*(4), 591–597.

Vivas, M., Zas, R., & Solla, A. (2012). Screening of Maritime pine (*Pinus pinaster*) for resistance to *Fusarium circinatum*, the causal agent of *Pitch Canker* disease. *Forestry, 85*, 185–192.

Vogan, P. J., & Schoettle, A. W. (2015). Selection for resistance to white pine blister rust affects the abiotic stress tolerances of limber pine. *Forest Ecology and Management, 344*, 110–119.

Vogel, B. S., Wildung, M. R., Vogel, G., & Croteau, R. (1996). Abietadiene synthase from grand fir (*Abies grandis*) cDNA isolation, characterization, and bacterial expression of a bifunctional diterpene cyclase involved in resin acid biosynthesis. *Journal of Biological Chemistry, 271*(38), 23262–23268.

Volkova, P., Shipunov, A., Borisova, P., Moseng, R., & Ivens, R. (2014). In search of hybridity: The case of Karelian spruces. *Silva Fennica, 48*(2), 1072–1085.

Vornam, B., Arkhipov, A., & Finkeldey, R. (2012). Nucleotide diversity and gene expression of Catalase and Glutathione peroxidase in irradiated Scots pine (*Pinus sylvestris* L.) from the Chernobyl exclusion zone. *Journal of Environmental Radioactivity, 106*, 20–26.

Wachowiak, W., Balk, P. A., & Savolainen, O. (2009). Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (*Pinus sylvestris* L.). *Tree Genetics & Genomes, 5*(1), 117.

Wachowiak, W., Boratyńska, K., & Cavers, S. (2013). Geographical patterns of nucleotide diversity and population differentiation in three closely related European pine species in the *Pinus mugo* complex. *Botanical Journal of the Linnean Society, 172*(2), 225–238.

Wachowiak, W., Odrzykoski, I., Myczko, Ł., & Prus-Głowacki, W. (2006). Lack of evidence on hybrid swarm in the sympatric population of *Pinus mugo* and *P. sylvestris*. *Flora-Morphology, Distribution, Functional Ecology of Plants, 201*(4), 307–316.

Wachowiak, W., Palme, A. E., & Savolainen, O. (2011a). Speciation history of three closely related pines *Pinus mugo* (T.), *P. uliginosa* (N.) and *P. sylvestris* (L.). *Molecular Ecology, 20*(8), 1729–1721.

Wachowiak, W., Salmela, M. J., Ennos, R. A., Iason, G., & Cavers, S. (2011b). High genetic diversity at the extreme range edge: Nucleotide variation at nuclear loci in Scots pine (*Pinus sylvestris* L.) in Scotland. *Heredity, 106*(5), 775.

Wagner, D. B., Dong, J., Carlson, M. R., & Yanchuk, A. D. (1991a). Paternal leakage of mitochondrial DNA in *Pinus*. *Theoretical and Applied Genetics, 82*(4), 510–514.

Wagner, D. B., Furnier, G. R., Saghai-Maroof, M. A., Williams, S. M., Dancik, B. P., & Allard, R. W. (1987). Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. *Proceedings of the National Academy of Sciences, 84*(7), 2097–2100.

Wagner, D. B., Govindaraju, D. R., Yeatman, C. W., & Pitel, J. A. (1989). Paternal chloroplast DNA inheritance in a diallel cross of jack pine (*Pinus banksiana* Lamb.). *Journal of Heredity, 80*(6), 483–485.

Wagner, D. B., Nance, W. L., Nelson, C. D., Li, T., Patel, R. N., & Govindaraju, D. R. (1992). Taxonomic patterns and inheritance of chloroplast DNA variation in a survey of *Pinus echinata*, *Pinus elliottii*, *Pinus palustris*, and *Pinus taeda*. *Canadian Journal of Forest Research, 22*(5), 683–689.

Wagner, D. B., Sun, Z. X., Govindaraju, D. R., & Dancik, B. P. (1991b). Spatial patterns of chloroplast DNA and cone morphology variation within populations of a *Pinus banksiana – Pinus contorta* sympatric region. *American Naturalist, 138*, 156–170.

Wahid, N., González-Martínez, S. C., El Hadrami, I., & Boulli, A. (2004). Genetic structure and variability of natural populations of maritime pine (*Pinus pinaster* Aiton) in Morocco. *Silvae Genetica, 53*(3), 93–98.

Wakamiya, I., Newton, R. J., Johnston, J. S., & Price, H. J. (1993). Genome size and environmental factors in the genus *Pinus*. *American Journal of Botany, 80*, 1235–1241.

Wakasugi, T., Tsudzuki, J., Ito, S., Nakashima, K., Tsudzuki, T., & Sugiura, M. (1994). Loss of all ndh genes as determined by sequencing the entire chloroplast genome of the black pine *Pinus thunbergii*. *Proceedings of the National Academy of Sciences, 91*(21), 9794–9798.

Wallace, A. R. (1876). *The geographical distribution of animals*. London: Harper and Brothers.

Wallis, C., Eyles, A., Chorbadjian, R., McSpadden Gardener, B., Hansen, R., Cipollini, D., et al. (2008). Systemic induction of phloem secondary metabolism and its relationship to resistance to a canker pathogen in Austrian pine. *New Phytologist, 177*(3), 767–778.

Wallis, C. M., Reich, R. W., Lewis, K. J., & Huber, D. P. (2010). Lodgepole pine provenances differ in chemical defense capacities against foliage and stem diseases. *Canadian Journal of Forest Research, 40*(12), 2333–2344.

Walter, R., & Epperson, B. K. (2001). Geographic pattern of genetic variation in *Pinus resinosa*: Area of greatest diversity is not the origin of postglacial populations. *Molecular Ecology, 10*(1), 103–111.

Walter, R., & Epperson, B. K. (2005). Geographic pattern of genetic diversity in *Pinus resinosa*: Contact zone between descendants of glacial refugia. *American Journal of Botany, 92*(1), 92–100.

Wan, L. C., Wang, F., Guo, X., Lu, S., Qiu, Z., Zhao, Y., et al. (2012a). Identification and characterization of small non-coding RNAs from Chinese fir by high throughput sequencing. *BMC Plant Biology, 12*(1), 146.

Wan, L. C., Zhang, H., Lu, S., Zhang, L., Qiu, Z., Zhao, Y., et al. (2012b). Transcriptome-wide identification and characterization of miRNAs from *Pinus densata*. *BMC Genomics, 13*(1), 132.

Wang, B., Mao, J. F., Gao, J. I. E., Zhao, W. E. I., & Wang, X. R. (2011). Colonization of the Tibetan Plateau by the homoploid hybrid pine *Pinus densata*. *Molecular Ecology, 20*(18), 3796–3811.

Wang, C., Perlin, M. H., Stockum, V., Jr, R. R., Hamilton, C. H., & Wagner, D. B. (1997). Chloroplast DNA polymorphisms in *Tsuga canadensis* and *Tsuga caroliniana*. *Canadian Journal of Forest Research, 27*(9), 1329–1335.

Wang, D., Eyles, A., Mandich, D., & Bonello, P. (2006b). Systemic aspects of host–pathogen interactions in Austrian pine (*Pinus nigra*): A proteomics approach. *Physiological and Molecular Plant Pathology, 68*(4), 149–157.

Wang, H. W., & Ge, S. (2006). Phylogeography of the endangered *Cathaya argyrophylla* (Pinaceae) inferred from sequence variation of mitochondrial and nuclear DNA. *Molecular Ecology, 15*(13), 4109–4122.

Wang, J., Abbott, R. J., Ingvarsson, P. K., & Liu, J. (2014). Increased genetic divergence between two closely related fir species in areas of range overlap. *Ecology and Evolution, 4*(7), 1019–1029.

Wang, Q. M., et al. (1998). Selecting provenances and individuals for resistance to die-back and brown-spot needle blight disease and growth in loblolly and slash pine in Hanjing, Jiangsu. *Journal of Jiansu Forestry Science and Technology, 25*(4), 1–7.

Wang, T., Hamann, A., Yanchuk, A., O'Neill, G. A., & Aitken, S. N. (2006a). Use of response functions in selecting lodgepole pine populations for future climate. *Global Change Biology, 12*, 2404–2416.

Wang, T., O'Neill, G. A., & Aitken, S. N. (2010a). Integrating environmental and genetic effects to predict responses of tree populations to climate. *Ecological Applications, 20*(1), 153–163.

Wang, X., Liu, Z., Niu, L., & Fu, B. (2013). Long-term effects of simulated acid rain stress on a staple forest plant, *Pinus massoniana* Lamb: A proteomic analysis. *Trees, 27*(1), 297–309.

Wang, X., Lu, J., Yue, W., Li, L., Zou, J., Li, X., He, X., Duan, B., & Liu, J. (2016). Determining the extent and direction of introgression between three spruce species based on molecular markers from three genomes with different rates of gene flow. *Plant Systematics and Evolution, 302*(6), 691–701.

Wang, X. R., Lindgren, D., Szmidt, A. E., & Yazdani, R. (1991). Pollen migration into a seed orchard of *Pinus sylvestris* L and the methods of its estimation using allozyme markers. *Scandinavian Journal of Forest Research, 6*(1–4), 379–385.

Wang, X. R., & Szmidt, A. E. (1994). Hybridization and chloroplast DNA variation in a *Pinus* species complex from Asia. *Evolution, 48*, 1020–1031.

Wang, X.-R., Tsumura, Y., Yoshimaru, H., Nagasaka, K., & Szmidt, A. E. (1999). Phylogenetic relationships of Eurasia pines (*Pinus*, Pinaceae) based on chloroplast *rbcL, matK, rpl20-rps18* spacer, and *trnV* intron sequences. *American Journal of Botany, 86*, 1742–1753.

Wang, Y., Jia, Q., Zhang, L., Zhang, Z., & Zhang, H. (2015). Allelic variation in Cinnamyl alcohol dehydrogenase (LoCAD) associated with wood properties of *Larix olgensis*. *Forests, 6*(5), 1649–1665.

Wang, Z. M., & Nagasaka, K. (1997). Allozyme variation in natural populations of *Picea glehnii* in Hokkaido, Japan. *Heredity, 78*(5), 470–475.

Wang, Z. S., Sun, H. Q., Wang, H. W., & Ge, S. (2010b). Isolation and characterization of 50 nuclear microsatellite markers for *Cathaya argyrophylla*, a Chinese endemic conifer. *American Journal of Botany, 97*(11), e117–e120.

Warren, R. L., Keeling, C. I., Yuen, M. M. S., Raymond, A., Taylor, G. A., Vandervalk, B. P., et al. (2015). Improved white spruce (*Picea glauca*) genome assemblies and annotation of large gene families of conifer terpenoid and phenolic defense metabolism. *The Plant Journal, 83*(2), 189–212.

Watano, Y., Imazu, M., & Shimizu, T. (1995). Chloroplast DNA typing by PCR-SSCP in the *Pinus pumila – P. parviflora* var. *pentaphylla* complex (Pinaceae). *Journal of Plant Research, 108*(4), 493–499.

Watano, Y., Kanai, A., & Tani, N. (2004). Genetic structure of hybrid zones between *Pinus pumila* and *P. parviflora* var. *pentaphylla* (Pinaceae) revealed by molecular hybrid index analysis. *American Journal of Botany, 91*(1), 65–72.

Webby, R. F., Markham, K. R., & Molloy, B. P. (1987). The characterisation of New Zealand Podocarpus hybrids using flavonoid markers. *New Zealand Journal of Botany, 25*(3), 355–366.

Wegrzyn, J. L., Liechty, J. D., Stevens, K. A., Wu, L. S., Loopstra, C. A., Vasquez-Gross, H. A., et al. (2014). Unique features of the loblolly pine (*Pinus taeda* L.) megagenome revealed through sequence annotation. *Genetics, 196*(3), 891–909.

Wegrzyn, J. L., Lin, B. Y., Zieve, J. J., Dougherty, W. M., Martínez-García, P. J., Koriabine, M., et al. (2013). Insights into the loblolly pine genome: Characterization of BAC and fosmid sequences. *PLoS One, 8*(9), e72439.

Wehenkel, C., del Rocío Mariscal-Lucero, S., Jaramillo-Correa, J. P., López-Sánchez, C. A., Vargas-Hernández, J. J., & Sáenz-Romero, C. (2017). Genetic diversity and conservation of Mexican forest trees. In *Biodiversity and conservation of woody plants* (pp. 37–67). Cham: Springer.

Wehenkel, C., & Sáenz-Romero, C. (2012). Estimating genetic erosion using the example of *Picea chihuahuana* Martínez. *Tree Genetics & Genomes, 8*(5), 1085–1094.

Weidman, E. H. (1939). Evidences of racial influences in a 25-year test of ponderosa pine. *Journal of Agricultural Research, 59*, 855–888.

Wellendorf, H., & Thomsen, I. M. (2008). Genetic variation in resistance against Heterobasidion annosum (Fr.) Bref. in Picea abies (L.) Karst. expressed after inoculation of neighboring stumps. *Silvae Genetica, 57*(1–6), 312–324.

Wells, O. O. (1964). Geographic variation in ponderosa pine. *Silvae Genetica, 13*(4), 89–103.

Wells, O. O., & Wakeley, P. C. (1966). Geographic variation in survival, growth, and fusiform rust infection of planted loblolly pine. *Forest Science, 12*(suppl_2), a0001–z0001.

Weng, C., Kubisiak, T., Nelson, C., & Stine, M. (2002). Mapping quantitative trait loci controlling early growth in a (longleaf pine× slash pine)× slash pine BC 1 family. *Theoretical and Applied Genetics, 104*(5), 852–859.

Westbrook, J. W., Chhatre, V. E., Wu, L. S., Chamala, S., Neves, L. G., Muñoz, P., et al. (2014). An annotated consensus genetic map for *Pinus taeda* L. and extent of linkage disequilibrium in three genotype-phenotype discovery populations. *bioRxiv*, 012625.

Westbrook, J. W., Resende, M. F., Munoz, P., Walker, A. R., Wegrzyn, J. L., Nelson, C. D., et al. (2013). Association genetics of oleoresin flow in loblolly pine: Discovering genes and

predicting phenotype for improved resistance to bark beetles and bioenergy potential. *New Phytologist, 199*(1), 89–100.

Westbrook, J. W., Walker, A. R., Neves, L. G., Munoz, P., Resende, M. F., Neale, D. B., et al. (2015). Discovering candidate genes that regulate resin canal number in *Pinus taeda* stems by integrating genetic analysis across environments, ages, and populations. *New Phytologist, 205*(2), 627–641.

Wheeler, N., & Neale, D. B. (2014). Landscape genomics: An emerging discipline that can aid forest land managers with planting stock decisions. *Mountain Views (CIRMOUNT), 7*(2), 10–14.

Wheeler, N., Wegrzyn, J., & Harry, D. (2011). *Genetic markers*. [Online Learning Module]. Genomics in Tree Breeding and Forest Ecosystem Management, Conifer Translational Genomics Network. eXtension Foundation. Available at: www.extension.org/pages/60385/conifer-translational-genomics-network-online-module-6:-genetic-markers (verified 11/17/2017).

Wheeler, N. C. (1989). Characterization of select pedigrees for genetic load. Weyerhaeuser For Res Tech Rep 050-3210/26) 13 pp.

Wheeler, N. C., & Bramlett, D. (1990). Operational flower stimulation treatments in a young loblolly pine (*Pinus taeda* L.) seed orchard. *Southern Journal of Applied Forestry, 15*(1), 44–50.

Wheeler, N. C., & Critchfield, W. B. (1985). The distribution and botanical characteristics of lodgepole pine. In D. M. Baumgartner, R. G. Krebil, J. T. Arnott, & G. F. Weetman (Eds.), *Symposium proceedings. Lodgepole pine: The species and its management, Spokane, Washington, 1984* (pp. 1–14). Pullman, WA: Washington State University, Cooperative Extension Service.

Wheeler, N. C., & Guries, R. P. (1982a). Biogeography of lodgepole pine. *Canadian Journal of Botany, 60*(9), 1805–1814.

Wheeler, N. C., & Guries, R. P. (1982b). Population structure, genic diversity, and morphological variation in *Pinus contorta* Dougl. *Canadian Journal of Forest Research, 12*(3), 595–606.

Wheeler, N. C., & Guries, R. P. (1987). A quantitative measure of introgression between lodgepole and jack pines. *Canadian Journal of Botany, 65*(9), 1876–1885.

Wheeler, N. C., Guries, R. P., & O'Malley, D. M. (1983). Biosystematics of the genus *Pinus*, subsection Contortae. *Biochemical Systematics and Ecology, 11*(4), 333–340.

Wheeler, N. C., Jech, K., Masters, S., Brobst, S. W., Alvarado, A. B., Hoover, A. J., & Snader, K. M. (1992). Effects of genetic, epigenetic, and environmental factors on taxol content in *Taxus brevifolia* and related species. *Journal of Natural Products, 55*(4), 432–440.

Wheeler, N. C., & Jech, K. S. (1992). The use of electrophoretic markers in seed orchard research. *New Forests, 6*, 311–328.

Wheeler, N. C., Jech, K. S., Masters, S. A., O'Brien, C. J., Timmons, D. W., Stonecypher, R. W., & Lupkes, A. (1995a). Genetic variation and parameter estimates in *Taxus brevifolia* (Pacific Yew). *Canadian Journal of Forest Research, 25*(12), 1913–1927.

Wheeler, N. C., Jech, K. S., & Rose, P. (1995b). *Genetic variation in Weyerhaeuser's natural and domesticated Douglas-fir populations: I Allozyme diversity* (pp. 1–27). Weyerhaeuser Co For Res Tech. Rep. No. 51-1206.

Wheeler, N. C., Jermstad, K. D., Krutovsky, K., Aitken, S. N., Howe, G. T., Krakowski, J., & Neale, D. B. (2005). Mapping of quantitative trait loci controlling adaptive traits in coastal Douglas-fir IV. Cold-hardiness QTL verification and candidate gene mapping. *Molecular Breeding, 15*(2), 145–156.

Wheeler, N. C., Kriebel, H. B., Lee, C. H., Read, R. A., & Wright, J. W. (1976). 15-year performance of European black pine in provenance tests in North Central United States. *Silvae Genetica, 25*(1), 1–6.

Wheeler, N. C., Masters, C. J., Cade, S. C., Ross, S. D., Keeley, J. W., & Hsin, L. Y. (1985). Girdling: An effective and practical treatment for enhancing seed yields in Douglas-fir seed orchards. *Canadian Journal of Forest Research, 15*(3), 505–510.

Wheeler, N. C., & Neale, D. B. (2013). Landscape genomics: An emerging discipline that can aid forest land managers with planting stock decisions. *Mountain Views: Cirmount, 7*(1), 10–14.

Wheeler, N. C., Steiner, K. C., Schlarbaum, S. E., & Neale, D. B. (2015). The evolution of forest genetics and tree improvement research in the United States. *Journal of Forestry, 113*(5), 500–510.

Wheeler, N. C., Wample, R. L., & Pharis, R. P. (1980). Promotion of flowering in the Pineaceae by gibberellins. IV. Seedlings and sexually mature grafts of lodgepole pine. *Physiologia Plantarum, 50*, 340–346.

Whetten, R. W., & Sederoff, R. R. (1992). Phenylalanine ammonia-lyase from loblolly pine purification of the enzyme and isolation of complementary DNA clones. *Plant Physiology, 98*(1), 380–386.

White, E. E. (1990). Chloroplast DNA in *Pinus monticola*. *Theoretical and Applied Genetics, 79*(1), 119–124.

White, T. L., Adams, W. T., & Neale, D. B. (2007). *Forest genetics*. Cambridge, MA: CABI Publishing.

White, T. L., & Ching, K. K. (1985). Provenance study of Douglas-Fir in the Pacific northwest region. *Silvae Genetica, 34*, 2–3.

Whitehill, J. G. A., Henderson, H., Strong, W., Jaquish, B., & Bohlmann, J. (2016). Function of Sitka spruce stone cells as a physical defense against white pine weevil. *Plant, Cell and Environment*. https://doi.org/10.1111/pce.12810.

Whitlock, M. C., & McCauley, D. E. (1999). Indirect measures of gene flow and migration: FST ≠ 1/(4 nm + 1). *Heredity, 82*(2), 117–125.

Whittall, J. B., Syring, J., Parks, M., Buenrostro, J., Dick, C., Liston, A., & Cronn, R. (2010). Finding a (pine) needle in a haystack: chloroplast genome sequence divergence in rare and widespread pines. *Molecular Ecology, 19*, 100–114.

Whittle, C. A., & Johnston, M. O. (2002). Male-driven evolution of mitochondrial and chloroplastidial DNA sequences in plants. *Molecular Biology and Evolution, 19*(6), 938–949.

Wikipedia contributors (2017) Árbol del Tule. Wikipedia, The Free Encyclopedia, 19 May 2017. https://en.wikipedia.org/wiki/%C3%81rbol_del_Tule. Accessed 21 Oct 2017.

Wilcox, M. D. (1983). Inbreeding depression and genetic variances estimated from self- and cross-pollinated families of *Pinus radiata*. *Silvae Genetica, 32*, 89–96.

Wilcox, M. D., & Miller, J. T. (1975). *Pinus nigra* provenance variation and selection in New Zealand. *Silvae Genetica, 24*(5/6), 132–140.

Wilcox, P. L. (1982). Genetic variation and inheritance of resistance to *Dothistroma* needle blight in *Pinus radiata*. *New Zealand Journal of Forestry Science, 12*(1), 14–35.

Wilcox, P. L., Amerson, H. V., Kuhlman, E. G., Liu, B. H., O'Malley, D. M., & Sederoff, R. R. (1996). Detection of a major gene for resistance to fusiform rust disease in loblolly pine by genomic mapping. *Proceedings of the National Academy of Sciences, 93*(9), 3859–3864.

Wilcox, P. L., Echt, C. E., & Burdon, R. D. (2007). Gene-assisted selection applications of association genetics for forest tree breeding. In *Association Mapping in Plants* (pp. 211–247). New York, NY: Springer.

Wiley, E. O., & Lieberman, B. S. (2011). *Phylogenetics: the theory of phylogenetic systematics* (2nd ed.). Hoboken: Wiley-Blackwell.

Wilkinson, R. C. (1983). Seed source variation in susceptibility of eastern white pine to white pine weevil attack. Proceedings, Twenty-eighth Northeastern Forest Tree Improvement Conference; July 7-9, 1983; Durham, NH. Radnor, PA: USDA Forest Service, Northeastern Forest Experiment Station; and Durham, NH: University of New Hampshire, Institute of Natural and Environmental Resources. pp. 134–139.

Williams, C. G. (2009). *Conifer reproductive biology*. Dordrecht/Heidelberg/London/New York: Springer.

Williams, C. G., & Neale, D. B. (1992). Conifer wood quality and marker-aided selection: A case study. *Canadian Journal of Forest Research, 22*(7), 1009–1017.

Williams, M. I., & Dumroese, R. K. (2014). Planning the future's forests with assisted migration. In V. A. Sample & R. P. Bixler (Eds.), *Forest conservation and management in the Anthropocene: Conference proceedings. RMRS-P-71* (pp. 133–144). Fort Collins: US Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Williamson, B., Mitchell, C. P., & Millar, C. S. (1976). Histochemistry of Corsican Pine Needles Infected by *Lophodermella sulcigena* (Rostr.) v. Höhn. *Annals of Botany, 40*(2), 281–288.
Willyard, A., Cronn, R., & Liston, A. (2009). Reticulate evolution and incomplete lineage sorting among the ponderosa pines. *Molecular Phylogenetics and Evolution, 52*(2), 498–511.
Willyard, A., Gernandt, D. S., Potter, K., Hipkins, V., Marquardt, P., Mahalovich, M. F., Langer, S. K., Telewski, F. W., Cooper, B., Douglas, C., & Finch, K. (2017). *Pinus ponderosa*: a checkered past obscured four species. *American Journal of Botany, 104*(1), 161–181.
Willyard, A. J. S., Gernandt, D. S., Liston, A., & Cronn, R. (2007). Fossil calibration of molecular divergence infers a moderate mutation rate and recent radiations for *Pinus*. *Molecular Biology and Evolution, 24*(1), 90–101.
Wingfield, M. J., Hammerbacher, A., Ganley, R. J., Steenkamp, E. T., Gordon, T. R., Wingfield, B. D., & Coutinho, T. A. (2008). Pitch canker caused by *Fusarium circinatum*–a growing threat to pine plantations and forests worldwide. *Australasian Plant Pathology, 37*(4), 319–334.
Wingfield, M. J., Jacobs, A., Coutinho, T. A., Ahumada, R., & Wingfield, B. D. (2002). First report of the pitch canker fungus, *Fusarium circinatum*, on pines in Chile. *Plant Pathology, 51*(3), 397–397.
Wolfe, J. A. (1969). Neogene floristic and vegetational history of the Pacific Northwest. *Madrono, 20*, 83–110.
Woo, K. S., Fins, L., McDonald, G. I., & Wiese, M. V. (2001). Differences in needle morphology between blister rust resistant and susceptible western white pine stocks. *Canadian Journal of Forest Research, 31*(11), 1880–1886.
Wright, J. W. (1944). Ecotypic differentiation in red ash. *Journal of Forestry, 42*, 591–597.
Wright, J. W. (1953). A survey of forest genetics research. *Journal of Forestry, 51*, 330–333.
Wright, J. W. (1955). Species crossability in spruce in relation to distribution and taxonomy. *Forest Science, 1*(4), 319–349.
Wright, J. W. (1959). Species hybridisation in the White Pines. *Forest Science, 5*(3), 210–222.
Wright, J. W. (1976). *Introduction to forest genetics*. New York: Academic Press, 463 pp.
Wright, J. W., & Gabriel, W. J. (1958). Species hybridization in the hard pines, series *Sylvestres*. *Silvae Genetica, 7*, 109–115.
Wright, J. W., Kung, F. H., Read, R. A., Lemmien, W. A., & Bright, J. N. (1971). Genetic variation in Rocky Mountatin Douglas-fir. *Silvae Genetica, 20*, 54–60.
Wright, J. W., Lemmien, W. A., & Bright, J. N. (1969). Early growth of ponderosa pine ecotypes in Michigan. *Forest Science, 15*(2), 121–129.
Wright, S. (1929). The evolution of dominance. *American Naturalist, 63*(689), 556–561.
Wright, S. (1951). The genetical structure of species. *Annals of Eugenics, 15*, 323–354.
Wu, C. I. (2001). The genic view of the process of speciation. *Journal of Evolutionary Biology, 14*(6), 851–865.
Wu, C. S., & Chaw, S. M. (2014). Highly rearranged and size-variable chloroplast genomes in conifers II clade (cupressophytes): Evolution towards shorter intergenic spacers. *Plant Biotechnology Journal, 12*(3), 344–353.
Wu, C. S., & Chaw, S. M. (2016). Large-scale comparative analysis reveals the mechanisms driving plastomic compaction, reduction, and inversions in conifers II (cupressophytes). *Genome Biology and Evolution, 8*(12), 3740–3750.
Wu, C. S., Wang, Y. N., Hsu, C. Y., Lin, C. P., & Chaw, S. M. (2011b). Loss of different inverted repeat copies from the chloroplast genomes of Pinaceae and cupressophytes and influence of heterotachy on the evaluation of gymnosperm phylogeny. *Genome Biology and Evolution, 3*, 1284–1295.
Wu, H., Echt, C. S., Popp, M. P., & Davis, J. M. (1997). Molecular cloning, structure and expression of an elicitor-inducible chitinase gene from pine trees. *Plant Molecular Biology, 33*(6), 979–987.
Wu, H. X., & Ying, C. C. (1997). Genetic parameters and selection efficiencies in resistance to western Gall Rust, Stalactiform blister rust, needle cast, and sequoia pitch moth in Lodgepole pine. *Forest Science, 43*, 571–581.

Wu, H. X., Ying, C. C., & Muir, J. A. (1996). Effect of geographic variation and jack pine introgression on disease and insect resistance in lodgepole pine. *Canadian Journal of Forest Research, 26*(5), 711–726.

Wu, J., Krutovskii, K. V., & Strauss, S. H. (1998). Abundant mitochondrial genome diversity, population differentiation and convergent evolution in pines. *Genetics, 150*(4), 1605–1614.

Wu, J., Krutovskii, K. V., & Strauss, S. H. (1999). Nuclear DNA diversity, population differentiation, and phylogenetic relationships in the California closed-cone pines based on RAPD and allozyme markers. *Genome, 42*(89), 893–908.

Wu, Q., Sun, C., Luo, H., Li, Y., Niu, Y., Sun, Y., et al. (2011a). Transcriptome analysis of *Taxus cuspidata* needles based on 454 pyrosequencing. *Planta Medica, 77*(04), 394–400.

Wu, X., Chen, T., Zheng, M., Chen, Y., Teng, N., Šamaj, J., et al. (2008). Integrative proteomic and cytological analysis of the effects of extracellular Ca2+ influx on *Pinus bungeana* pollen tube development. *Journal of Proteome Research, 7*(10), 4299–4312.

Xiang, Q. P., Wei, R., Shao, Y. Z., Yang, Z. Y., Wang, X. Q., & Zhang, X. C. (2015). Phylogenetic relationships, possible ancient hybridization, and biogeographic history of *Abies* (Pinaceae) based on data from nuclear, plastid, and mitochondrial genomes. *Molecular Phylogenetics and Evolution, 82*, 1–14.

Xie, C., & Ying, C. C. (1995). Genetic architecture and adaptive landscape of interior lodgepole pine (*Pinus contorta* ssp. *latifolia*) in Canada. *Canadian Journal of Forest Research, 25*, 2010–2021.

Xing, F., Mao, J. F., Meng, J., Dai, J., Zhao, W., Liu, H., Xing, Z., Zhang, H., Wang, X. R., & Li, Y. (2014). Needle morphological evidence of the homoploid hybrid origin of *Pinus densata* based on analysis of artificial hybrids and the putative parents, *Pinus tabuliformis* and *Pinus yunnanensis*. *Ecology and Evolution, 4*(10), 1890–1902.

Xiong, J. S., McKeand, S. E., Isik, F., Wegrzyn, J., Neale, D. B., Zeng, Z. B., et al. (2016). Quantitative trait loci influencing forking defects in an outbred pedigree of loblolly pine. *BMC Genetics, 17*(1), 138.

Xu, S., Tauer, C. G., & Nelson, C. D. (2008). Natural hybridization within seed sources of shortleaf pine (*Pinus echinata* mill.) and loblolly pine (*Pinus taeda* L.). *Tree Genetics & Genomes, 4*(4), 849–858.

Xue, B., Charest, P. J., Devantier, Y., & Rutledge, R. G. (2003). Characterization of a MYBR2R3 gene from black spruce (*Picea mariana*) that shares functional conservation with maize C1. *Molecular Genetics and Genomics, 270*(1), 78–86.

Yakovlev, I. A., Asante, D. K., Fossdal, C. G., Junttila, O., & Johnsen, Ø. (2011). Differential gene expression related to an epigenetic memory affecting climatic adaptation in Norway spruce. *Plant Science, 180*(1), 132–139.

Yakovlev, I. A., Lee, Y., Rotter, B., Olsen, J. E., Skroppa, T., Johnsen, O., et al. (2014). Temperature-dependent differential transcriptomes during formation of an epigenetic memory in Norway spruce embryogenesis. *Tree Genetics & Genomes, 10*, 355e366.

Yamamoto, N., Kojima, K., & Matsuoka, M. (1993). The presence of two types of gene that encode the chlorophyll a/b-binding protein (LHCPII) and their light-independent expression in pine (*Pinus thunbergii*). *Plant and Cell Physiology, 34*(3), 457–463.

Yamamoto, N., Mukai, Y., Matsuoka, M., Kano-Murakami, Y., Tanaka, Y., Ohashi, Y., et al. (1991). Light-independent expression of cab and rbcS genes in dark-grown pine seedlings. *Plant Physiology, 95*(2), 379–383.

Yamamoto, N., Tada, Y., & Fujimura, T. (1994). The promoter of a pine photosynthetic gene allows expression of a β-glucuronidase reporter gene in transgenic rice plants in a light-independent but tissue-specific manner. *Plant and Cell Physiology, 35*(5), 773–778.

Yanchuk, A., & Wheeler, N. (2008). Selection and breeding for insect and disease resistance. FAO Forestry Department Review: http://www.fao.org/forestry/26445/en/ (last checked 20 March 2017).

Yanchuk, A. D., Wallin, K., & Murphy, J. C. (2008). Evaluation of genetic variation and resistance in lodgepole pine in the early stages of a mountain pine beetle outbreak. *Tree Genetics and Genomes, 4*, 171–180.

Yanchuk, A. D., Yeh, F. C., & Dancik, B. P. (1988). Variation of stem rust resistance in a lodgepole pine provenance-family plantation. *Forest Science, 34*, 1067–1075.

Yang, J., Pedlar, J. H., McKenney, D. W., & Weersink, A. (2015). The development of universal response functions to facilitate climate-smart regeneration of black spruce and white pine in Ontario, Canada. *Forest Ecology and Management, 339*, 34–43.

Yang, R. C., Ye, Z., & Hiratsuka, Y. (1999). Susceptibility of *Pinus contorta* - *Pinus banksiana* complex to *Endocronartium harknessii*: Host-pathogen interactions. *Canadian Journal of Botany, 77*(7), 1035–1043.

Yang, R. C., & Yeh, F. C. (1995). Patterns of gene flow and geographic structure in *Pinus contorta* Dougl. *Forest Genetics, 2*, 65–75.

Yang, S. H. S., & Loopstra, C. C. (2005). Seasonal variation in gene expression for loblolly pines (*Pinus taeda*) from different geographical regions. *Tree Physiology, 25*(8), 1063.

Yang, Y., Zhang, D., Luscombe, D., Liao, W., Farjon, A., Katsuki, T., Xiang, Q., & Li, N. (2013). Abies beshanzuensis. The IUCN red list of threatened species 2013: e.T32318A2814360. https://doi.org/10.2305/IUCN.UK.2013-1.RLTS.T32318A2814360.en. Accessed 19 Oct 2017.

Yang, Z.-Y., Ran, J.-H., & Wang, X.-Q. (2012). Three genome-based phylogeny of Cupressaceae s.l.: Further evidence for the evolution of gymnosperms and Southern Hemisphere biogeography. *Molecular Phylogenetics and Evolution, 64*(3), 452–470.

Yazdani, R., Nilsson, J. E., Plomion, C., & Mathur, G. (2003). Marker trait association for autumn cold acclimation and growth rhythm in *Pinus sylvestris*. *Scandinavian Journal of Forest Research, 18*(1), 29–38.

Ye, T. Z., Yang, R. C., & Yeh, F. C. (2002). Population structure of a lodgepole pine (*Pinus contorta*) and jack pine (*P. banksiana*) complex as revealed by random amplified polymorphic DNA. *Genome, 45*(3), 530–540.

Yeaman, S., Hodgins, K. A., Lotterhos, K. E., Suren, H., Nadeau, S., Degner, J. C., et al. (2016). Convergent local adaptation to climate in distantly related conifers. *Science, 353*(6306), 1431–1433.

Yeaman, S., Hodgins, K. A., Suren, H., Nurkowski, K. A., Rieseberg, L. H., Holliday, J. A., & Aitken, S. N. (2014). Conservation and divergence of gene expression plasticity following c. 140 million years of evolution in lodgepole pine (*Pinus contorta*) and interior spruce (*Picea glauca*× *Picea engelmannii*). *New Phytologist, 203*(2), 578–591.

Yeh, F. C., & Arnott, J. T. (1986). Electrophoretic and morphological differentiation of *Picea sitchensis*, *Picea glauca*, and their hybrids. *Canadian Journal of Forest Research, 16*(4), 791–798.

Yeh, F. C., & El-Kassaby, Y. A. (1980). Enzyme variation in natural populations of Sitka spruce (*Picea sitchensis*). 1. Genetic variation patterns among trees from 10 IUFRO provenances. *Canadian Journal of Forest Research, 10*(3), 415–422.

Yeh, F. C., & Hu, X. S. (2005). Genetic structure and migration from mainland to island populations in *Abies procera* Rehd. *Genome, 48*(3), 461–473.

Yeh, F. C., Khalil, M. A. K., El-Kassaby, Y. A., & Trust, D. C. (1986). Allozyme variation in *Picea mariana* from Newfoundland: Genetic diversity, population structure, and analysis of differentiation. *Canadian Journal of Forest Research, 16*(4), 713–720.

Yeh, F. C., & Layton, C. (1979). The organization of genetic variability in central and marginal populations of lodgepole pine (*Pinus contorta* spp. *latifolia*). *Canadian Journal of Genetics and Cytology, 21*(4), 487–503.

Yi, N. J., Hang, Z. M., et al. (2000). Genetic variation of RAPD markers in a disease resistant seed orchard of *Pinus elliottii*. *Scientia Silvae Sinicae, 36*(sp. 1), 51–55.

Yi, X., Gao, L., Wang, B., Su, Y. J., & Wang, T. (2013). The complete chloroplast genome sequence of *Cephalotaxus oliveri* (*Cephalotaxaceae*): Evolutionary comparison of *Cephalotaxus* chloroplast DNAs and insights into the loss of inverted repeat copies in gymnosperms. *Genome Biology and Evolution, 5*(4), 688–698.

Ying, C. C., Illingworth, K., & Carlson, M. (1985). Geographic variation of lodgepole pine and its implications for tree improvement in British Columbia. In D. M. Baumgartner, R. G. Krebil, J. T. Arnott, & G. F. Weetman (Eds.), *Symposium proceedings. Lodgepole pine: The species*

and its management, Spokane, Washington, 1984 (pp. 55–60). Pullman: Washington State University, Cooperative Extension Service.

Yoshimaru, H., Ohba, K., Tsurumi, K., Tomaru, N., Murai, M., Mukai, Y., et al. (1998). Detection of quantitative trait loci for juvenile growth, flower bearing and rooting ability based on a linkage map of sugi (*Cryptomeria japonica* D. Don). *Theoretical and Applied Genetics, 97*(1–2), 45–50.

Young, C. H. (2003). USDA Forest Service forest health protection resistance screening center. In Proc. 27th South For Tree Improv Conf. Stillwater, OK. CR McKinley (Ed.), p. 201.

Yu, H., Ge, S., & Hong, D. Y. (2000). Allozyme diversity and population genetic structure of *Pinus densata* Master in northwestern Yunnan, China. *Biochemical Genetics, 38*(5–6), 138–146.

Yu, Q., & Yang, D.-Q. (2003). Genetic variation in decay resistance and its correlation to wood density and growth in white spruce. *Canadian Journal of Forest Research, 33*(11), 2177–2183.

Yu, X. T. (2000). A summary of the studies on Chinese fir in the 1990s. II. Summary of research and development on disease and insect control, wood properties and future research directions. *Journal of Fujian College of Forestry, 20*(2), 283–288.

Zabinski, C. (1992). Isozyme variation in eastern hemlock. *Canadian Journal of Forest Research, 22*(12), 1838–1842.

Zas, R., Solla, A., & Sampedro, L. (2007). Variography and kriging allow screening *Pinus pinaster* resistant to *Armillaria ostoyae* in field conditions. *Forestry, 80*(2), 201–209.

Zeng, W., Sun, S., Li, H., O'Malley, D., Amerson, H., & Li, B. (2004). Search for major genes to improve productivity and rust resistance of loblolly pine. 2004 TAPPI Paper Summit - Spring Technical and International Environmental Conference. 2004, pp. 661–668.

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., et al. (2014b). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science, 346*(6215), 1311–1320.

Zhang, J., Wu, T., Li, L., Han, S., Li, X., Zhang, S., & Qi, L. (2013). Dynamic expression of small RNA populations in larch (*Larix leptolepis*). *Planta, 237*(1), 89–101.

Zhang, J. W., Klopfenstein, N. B., & Peterson, G. W. (1997). Genetic variation in disease resistance of *Juniperus virginiana* and *J. scopulorum* grown in eastern Nebraska. *Silvae Genetica, 46*(1), 11–16.

Zhang, X. H., & Chiang, V. L. (1997). Molecular cloning of 4-coumarate: Coenzyme A ligase in loblolly pine and the roles of this enzyme in the biosynthesis of lignin in compression wood. *Plant Physiology, 113*(1), 65–74.

Zhang, Y., Ma, J., Yang, B., Li, R., Zhu, W., Sun, L., et al. (2014a). The complete chloroplast genome sequence of *Taxus chinensis* var. *mairei* (*Taxaceae*): Loss of an inverted repeat region and comparative analysis with related species. *Gene, 540*(2), 201–209.

Zhao, G., Sun, M., Wilde, S. A., & Li, S. Z. (2004). A Paleo-Mesoproterozoic supercontinent: Assembly, growth and breakup. *Earth-Science Reviews, 67*, 91–123.

Zhao, H. B., & Li, S. J. (1989). The selection of plus trees among *Pinus tabulaeformis* with resistance to *Matsucoccu matsumurae*. *Scientia Silvae Sinicae, 25*(1), 7–13.

Zhao, Y. J. (2000). Genetic test for introduced loblolly pine family. *Journal of Fujian College of Forestry, 20*(3), 276–279.

Zhen, Y., Zhao, Z. Z., Zheng, R. H., & Shi, J. (2012). Proteomic analysis of early seed development in *Pinus massoniana* L. *Plant Physiology and Biochemistry, 54*, 97–104.

Zhou, L., Creech, D. L., Krauss, K. W., Yunlong, Y., & Kulhavy, D. L. (2010b). Can we improve the salinity tolerance of genotypes of *Taxodium* by using varietal and hybrid crosses? *Hortscience, 45*(12), 1773–1778.

Zhou, T. X., & Ma, C. G. (2000). The clonal selection and breeding of Chinese fir: Kaitian series. *Scientia Silvae Sinicae, 36*(3), 40–45.

Zhou, Y., Zhang, L., Liu, J., Wu, G., & Savolainen, O. (2014). Climatic adaptation and ecological divergence between two closely related pine species in Southeast China. *Molecular Ecology, 23*(14), 3504–3522.

Zhou, Y. F., Abbott, R. J., Jiang, Z. Y., Du, F. K., Milne, R. I., & Liu, J. Q. (2010a). Gene flow and species delimitation: A case study of two pine species with overlapping distributions in Southeast China. *Evolution, 64*(8), 2342–2352.

Zimin, A. V., Stevens, K. A., Crepeau, M. W., Puiu, D., Wegrzyn, J. L., Yorke, J. A., et al. (2017). An improved assembly of the loblolly pine mega-genome using long-read single-molecule sequencing. *Gigascience, 6*, 1–4.

Zobel, B. (1978). Gene conservation—As viewed by a forest tree breeder. *Forest Ecology and Management, 1*, 339–344.

Zobel, B., & Talbert, J. (1984). *Applied forest tree improvement*. New York: Wiley.

Zobel, B. J. (1953). Are there natural loblolly-shortleaf pine hybrids? *Journal of Forestry, 51*(7), 494–495.

Zobel, D. B., Roth, L. F., & Hawk, G. M. (1985). Ecology, pathology and management of Port-Orford-cedar (*Chamaecyparis lawsoniana*). Gen. Tech. Rep. PNW 184. Portland, OR. US Department of Agriculture, Forest Service, Pac. NW For. And Range Exp. Sta. 161 p.

Zonneveld, B. J. M. (2012). Conifer genome sizes of 172 species, covering 64 of 67 genera, range from 8 to 72 picogram. *Nordic Journal of Botany, 30*(4), 490–502.

Zou, J., Sun, Y., Li, L., Wang, G., Yue, W., Lu, Z., Wang, Q., & Liu, J. (2013). Population genetic evidence for speciation pattern and gene flow between *Picea wilsonii, P. morrisonicola* and *P. neoveitchii*. *Annals of Botany, 112*(9), 1829–1844.

Zulak, K. G., Lippert, D. N., Kuzyk, M. A., Domanski, D., Chou, T., Borchers, C. H., & Bohlmann, J. (2009). Targeted proteomics using selected reaction monitoring reveals the induction of specific terpene synthases in a multi-level study of methyl jasmonate-treated Norway spruce (*Picea abies*). *The Plant Journal, 60*(6), 1015–1030.

Index

D. B. Neale, N. C. Wheeler, *The Conifers: Genomes, Variation and Evolution*,
https://doi.org/10.1007/978-3-319-46807-5

Q