CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Genome-wide Identification and Characterization of Nodule Inception Like protein *(NLP)* Gene Family in Mungbean *(Vigna radiata)*

by

Sualeha Maheen

A thesis submitted in partial fulfillment for the degree of Master of Science

in the

Faculty of Health and Life Sciences
Department of Bioinformatics and Biosciences

2024

I dedicate my efforts to my mother Farina Birjees Mir, the love of my life, whom I hold close to my heart

CERTIFICATE OF APPROVAL

Genome-wide Identification and Characterization of Nodule Inception Like protein *(NLP)* Gene Family in Mungbean *(Vigna radiata)*

by

Sualeha Maheen

(MBS221013)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Ikram Ullah	Hazara University, Mansehra
(b)	Internal Examiner	Dr. Sania Riaz	CUST, Islamabad
(c)	Supervisor	Dr. Sohail Ahmad Jan	CUST, Islamabad

Dr. Sohail Ahmad Jan
Thesis Supervisor
April, 2024

Dr. Syeda Marriam Bakhtiar
Head
Dept. of Bioinformatics & Biosciences
April, 2024

Dr. Sahar Fazal
Dean
Faculty of Health and Life Sciences
April, 2024

Author's Declaration

I, **Sualeha Maheen** hereby state that my MS thesis titled "**Genome-wide Identification and Characterization of Nodule Inception Like protein *(NLP)* Gene Family in Mungbean *(Vigna radiata)***" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, Islamabad or anywhere else in the country/abroad.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my MS Degree.

(Sualeha Maheen)

Registration No: MBS221013

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Genome-wide Identification and Characterization of Nodule Inception Like protein *(NLP)* Gene Family in Mungbean *(Vigna radiata)***" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS Degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized work.

(Sualeha Maheen)

Registration No: MBS221013

Acknowledgement

All the praises and thanks be to Allah, Who has guided me to this, and never could have been found guidance, where it not but Allah had guided me. In times of darkness, You, Allah are my guiding light. Thank you for illuminating my path.

I would like to express my deepest gratitude to the following individuals without whom this thesis would not have been possible: My supervisor, Sir Sohail Ahmad Jan, for their unwavering guidance, invaluable insights, and continuous support throughout the entire research process. Their expertise and encouragement have been instrumental in shaping the direction of this study. I am grateful for their valuable feedback and constructive criticism, which greatly contributed to the improvement of this thesis. A special thanks to Sir Sami Ullah, whose moral support made this research project possible. His help not only facilitated the execution of the study but also allowed me to attend conferences and engage with the academic community. I am thankful for his assistance and resources, which played a crucial role in the successful completion of this research.

I would also like to thank my mother Farina Birjees Mir for their unwavering support and understanding during the challenging moments of this academic journey. Their encouragement and belief in me were my constant motivators. Lastly, to Mian Sufyan Rauf who generously contributed his time, in formatting my work. I am deeply thankful to everyone.

(Sualeha Maheen)

Abstract

Transcription factors known as NODULE-INCEPTION-LIKE Proteins *(NLPs)* play a crucial role in governing the nitrogen response in plants. Although extensive research has been conducted on *NLPs* in various plants, there is limited information available for mungbean. In this study, in-silico tools were employed to identify and characterize *NLP* in the vascular plant *Vigna radiata*. Seven Vr*NLPs* were identified, sharing physical and chemical attributes with At*NLPs*. Selected Vr*NLP*sequences exhibited a common domain with scale At*NLPs*, confirming their membership in the same gene family. Interestingly, the average gene lengths of Vr*NLPs* were significantly higher than those of At*NLPs*, while the average protein lengths of Vr*NLPs* showed a similar trend. This was also observed in the GRAVY (Grand Average of Hydropathicity) values, with Vr*NLPs* having higher values than At*(NLP)*. However, the average molecular weight (MW) and Theoretical Iso-electric point (pI) of At*(NLP)* were found to be higher than those of Vr*NLPs*. Bioinformatics tools indicated that all Vr*NLPs* are hydrophilic and localized in the nucleus, sharing a substantial degree of homology in their gene structures and protein motifs with At*(NLP)*. Phylogenetic analysis revealed evolutionary divergence and variation among Vr*NLPs*, while demonstrating significant evolutionary connections with *(NLP)* from various other vascular plants, supporting a common ancestry. Vr*NLPs* were found to be closely related to *NLPs* of *Arabidopsis thaliana*, *Zea mays*, *Brassica napus*, *Physcomitrella paten*, and *Oryza sativa L.* Protein-protein interaction analysis indicated that the interacting proteins of the majority of Vr*NLPs* were associated with the transportation of biomolecules. Notably, the protein-protein interaction study suggested a strong coordination between Vr*NLPs* and nitrogen-responsive genes such as nitrate reductase, emphasizing the significant regulation of *NLPs* by nitrogen supply.

Contents

List of Figures

List of Tables

Abbreviations

GRAVY	Grand Average of Hydropathicity
GS (GOGAT)	Gutamine Synthetase (Glutamine oxoglutarate aminotransferase)
GSDS	Gene Structure Display Server
ITOL	Interactive Tree of Life software v6
MEME	Multiple Em foe Motif Elicitation
NCBI	National Centre for Biotechnology Information
NH4+	Ammonium
NLP	NIN-LIKE PROTEIN
NO3-	Nitrate
NRE	Nitrate Responsive Cis-element
NUE	N itrogen Use Efficiency
NUpE	Nitrogen Uptake Efficiency
NUtE	Nitrogen Utilization Efficiency
NiR	Nitrate Reductase
PTM	Post Translational Modifications
Plant TFDB	Plant Transcription Factor Database
TAIR	The Arabidopsis Information Resource
TFs	Transcription Factors
UTRs	Un-translated Regions

Chapter 1

Introduction

1.1 Background

Nitrogen, represented as "N," holds significant importance as a colorless and odorless element. Despite its unassuming nature, nitrogen constitutes 78% of Earth's atmosphere and plays a central role in various plant structures, influencing both internal and external metabolic functions [1]. Studies reveal that only 50% of applied nitrogen (N) is efficiently utilized by crops, with the remaining 50% lost through various environmental pathways. This loss not only contaminates soil, water, and air but also adversely affects the farmer's return on investment. Therefore, enhancing nitrogen use efficiency (NUE) is crucial in crop development programs and agronomic management systems. Factors such as volatilization, surface runoff, and leaching contribute to low nitrogen consumption [2]. Nitrogen is a vital component of the chlorophyll molecule, responsible for giving plants their green color. This element is indispensable for photosynthesis, the process by which plants produce food. Nitrogen is also a primary constituent of plant protoplasm [3]. Plants have developed effective nitrogen absorption and metabolism systems, particularly in well-aerated soil. Nitrate, mainly transported by nitrate transporters, serves as the primary nitrogen source absorbed by plants. Cells take up nitrate, which is

then reduced to ammonium by nitrate reductase and nitrite reductase. Subsequently, glutamine synthetase and glutamate synthetase facilitate the absorption of ammonium into amino acids [4].

However, plants absorb only a limited percentage (30-40%) of the applied nitrogen, with a significant portion (60-70%) being lost in the soil. This phenomenon results in substantial soil and water pollution [6]. Approximately 40% of nitrous oxide emissions are attributed to human activities. Livestock manure presents a dual emissions challenge, especially in larger farming operations. Ben Lilliston, Director of Rural Strategies and Climate Change at the Institute for Agriculture and Trade Policy, explains, "When manure lacks access to oxygen, especially at the pit's bottom, it transforms into nitrous oxide." Furthermore, nitrous oxide poses a secondary concern as exposure to sunlight and oxygen in the stratosphere leads to its conversion into nitrogen oxides. Nitrogen oxides have the potential to deplete the ozone layer, crucial for shielding the Earth's surface from the majority of the sun's UV radiation [5]. The potency of this gas stems from its dual-threat effect. Over a 100-year period, one pound of N2O warms the atmosphere approximately 300 times more than one pound of CO2. N2O emerges as a significant contributor to climate change due to its potency and relatively extended atmospheric lifespan [6].

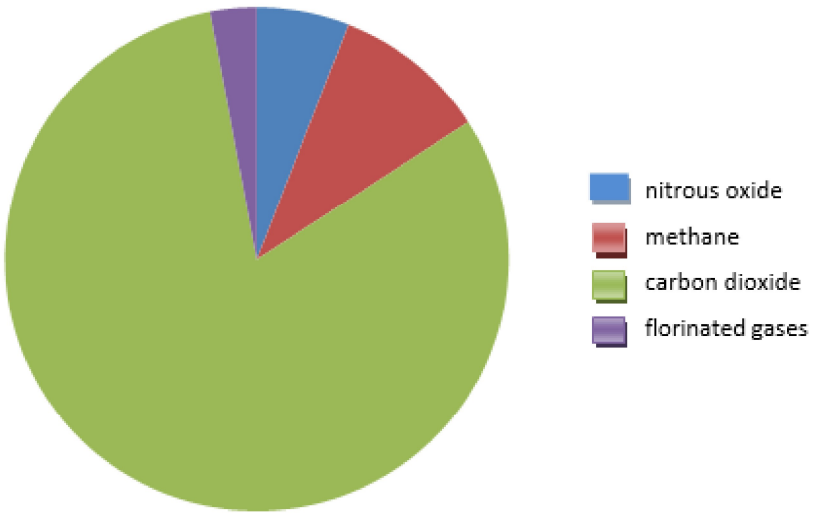


FIGURE 1.1: Concentration of gases in atmosphere [6]

1.2 Nitrogen Use Efficiency (NUE)

NUE is commonly defined as the product of the yield and the total amount of nitrogen (N) provided, with further distinction into nitrogen utilization efficiency and nitrogen uptake efficiency. The term "N Utilization Efficiency" (NUtE) specifically refers to a plant's ability to absorb and redistribute nitrogen within itself. In agricultural systems, there is an internal nitrogen cycle driven by crops, soil, and animals. This cycle experiences permanent nitrogen losses through N export with agricultural products and unavoidable losses to the environment, such as leaching outside the growing season. Agriculture is expected to significantly enhance overall productivity, leading to an increased need for nitrogen. To compensate for exported, lost, and additional nitrogen requirements, mineral nitrogen fertilizer is commonly applied. Nitrogen Uptake Efficiency (NUpE) pertains to the amount of nitrogen a plant can acquire from nitrogen sources [7]. Improving crop NUE is widely acknowledged as an economical, practical, and desirable approach to mitigate nitrogen-related agricultural and environmental issues. Even a modest 1% increase in a crop's NUE is anticipated to have a substantial impact on agricultural production, potentially resulting in annual savings of up to 1.1 billion US dollars [7].

$$NUE = \frac{N \text{ removal with harvest}}{\text{input from minerals}} \times 100$$

With,

- N removal=arable and permanent crop output $\times$ average N content
- Mineral N input = N fertilizer usage

However, the complete molecular mechanism governing NUE remains insufficiently understood.

1.3 Plant Transcription Factors (TFs)

Several transcription factors play a vital role in nitrate signaling by regulating the expression of genes involved in nitrate uptake, assimilation, and related metabolic

pathways. These transcription factors aid plants in adapting to fluctuations in soil nitrate levels and in orchestrating gene expression to optimize nitrogen absorption.

The *NLP* (NIN-like protein) family plays a critical role in regulating genes associated with nitrogen metabolism and nitrate signaling in plants. Originally identified for their involvement in legume nodulation, these transcription factors, known as NIN-like proteins (*NLP*s), have been subsequently recognized for their role in nitrate signaling pathways. *NLP*s typically possess a conserved DNA-binding domain called the RWP-RK domain, enabling them to bind to specific DNA sequences in the promoters of target genes. This binding capability facilitates the regulation of gene expression in response to nitrate availability and other signaling cues. Through this mechanism, *NLP*s exert direct control over the expression of genes involved in the absorption, transport, and assimilation of nitrates.

The molecular processes involved in nitrogen uptake, absorption, transport, and metabolism in plants have evolved through the action of various transcription factors (TFs) and gene families [8]. Plants primarily acquire inorganic nitrate (NO3-) from the soil, transported into cells by nitrate transporters like NRT1 and NRT2, facilitated by channels such as the chloride channel and slow anion-channel-related homologues. Subsequently, nitrate reductases (NIA1, NIA2) and nitrite reductase (NiR) convert the absorbed inorganic nitrate into ammonium (NH4+). Enzymes like GOGAT and GS assist in further absorption of ammonium into organic amino acids like glutamate and glutamine. These absorbed amino acids play crucial roles as nitrogen donors in the synthesis of macromolecules in plants, including chlorophyll, essential amino acids, and nucleic acids [8]. Additionally, assimilated amino acids and absorbed nitrate act as signaling molecules, influencing related transcription factors and cellular functions. Therefore, these discussions underscore the significance of nitrogen and nitrogen-responsive TFs in shaping plant structure, function, and overall NUE. In a study on *Chlamydomonas reinhardtii* under nitrogen starvation, it was found that a protein named MID (minus dominant protein) controls the differentiation of vegetative cells into gametes. This protein, in response to nitrogen signals, either activates or deactivates the minus or plus

gametic differentiation program. MID possesses a conserved sequence called RWPYRK after the leucine zipper motif, initially unrecognized but later identified as the first member of a novel TF family known as the RWP-RK gene family. All vascular plants, slime molds, green algae, and members of the RWP-RK gene family are included in plants. Subsequently, the first nodule inception gene (NIN), with an RWP-RK domain controlling nitrogen-mediated symbiotic root-nodule development, was discovered in the leguminous plant Lotus japonicas. It was revealed that other non-leguminous plants, including Arabidopsis, rice, wheat, and maize, contain additional NIN proteins and *NLP*s (Nodule-Inducing Proteins), but this feature is not present in animals [8].

Two sub-families of the RWP-RK gene family have been identified through thorough research:(a) RKD, which stands for the RWP-RK domain-containing gene family, and (b) *NLP*, which represents for the RKD with an additional domain at the C-terminus called Phox and Beml (PB1). NIN-Like Proteins (*NLP*s) are proteins that have been identified as having structural similarities to NIN genes. While *NLP*s are found in both non-leguminous and leguminous plants, NIN is only found in legumes. Plant-specific transcription factors called *NLP*s are important in regulating the nitrogen response. Protein-protein interactions are mediated by the PB1 domain (PFOOS64) of *NLP*s, whereas DNA binding is carried out by RWP-RK (PF02042) and transcriptional activation of genes is carried out by the N-terminal region. Nitrate-regulated genes' transcription is activated when *NLP*s bind to the nitrate- responsive cis-element (NRE) in the promoter region [9].

Since the *NLP* gene family is a reliable regulator of N-responsive genes, NUE may be enhanced [9]. So far, 9 *NLP* genes have been discovered in Arabidopsis thaliana, 6 in rice, 9 in maize, 31 in Brassica napus, 18 in wheat, and 9 in *Arabidopsis thaliana* through genome- wide research. Similar research on *NLP*s in mungbean hasn't yet been reported, though. Since identification, the study of the complete structural and functional characterization of *NLP* genes for NUE improvement has been focused on vascular plants.

Our research is centered on the vascular plant commonly known as mung bean, green gram, or moong bean, a legume widely cultivated for its edible seeds. Mung

bean plants are generally small, erect, and bushy, characterized by compound leaves with three leaflets that are green and alternately arranged on the stem. The plant produces small, fragrant, pale yellow flowers with a papilionaceous corolla, displaying butterfly-shaped petals. Mung beans find cultivation in various parts of the world, particularly in Asia, thriving in warm, tropical, and subtropical climates. They are often chosen as a summer crop. Notably, these plants possess the ability to fix nitrogen, contributing to the enhancement of the soil in which they are cultivated. Mung bean seeds are edible and can be consumed in diverse forms, including sprouted, boiled, or ground into flour. When fresh, the seeds are green, transforming to yellow upon drying. With a mild, nutty flavor, mung beans are rich in protein, fiber, vitamins (especially B vitamins like folate), and minerals (such as potassium and magnesium). Additionally, they are low in fat and calories, making them a nutritious addition to various dishes and dietary plans. Recognized for their health benefits, mung beans are considered a healthy food choice, providing a source of plant-based protein and serving as a valuable component in vegetarian and vegan diets. Their potential advantages include aiding in digestion, promoting heart health, and supporting weight management [9].

1.4 Problem Statement

NLP gene family or its member genes have been extensively examined in number of plants, however less is known about the *NLP* gene family in mungbeans.

1.5 Aim

The main aim of this study is to identify and characterize *NLP* gene family in Mungbeans *(Vigna radiata)*.

1.6 Objectives

- To identify *NLP* gene family in mungbeans.
- To identify the NUE (Nitrogen Use Efficiency) role in *Vigna radiata*.
- To assess the evolutionary relationship of *NLP* gene family.

1.7 Scope of a Study

The study of mung beans (*Vigna radiata*) presents a broad scope, covering various scientific, agricultural, culinary, and economic dimensions. Mung beans stand out as an exceptional source of plant-based protein, making them a valuable dietary option for individuals following vegetarian and vegan lifestyles. Protein, crucial for muscle growth, repair, and overall bodily functions, is abundantly present in mung beans. With low levels of fat and calories, mung beans prove to be a suitable addition to weight management and calorie-conscious diets. Given their high concentration of grain protein, mung beans show significant promise in tropical conditions. An additional noteworthy aspect is their potential to harness biological nitrogen fixation (BNF) through interactions with natural rhizobia in the nodule microbiome. This capability provides a substantial degree of nitrogen independence from fertilizers. Notably, mung bean seeds boast the highest nitrogen concentration (7.029%) among all seeds [10].

1.8 Impact on Society

Through genome-wide identification and functional characterization of the *NLP* gene family in *Vigna radiata*, efforts are directed at enhancing the Nitrogen Use Efficiency (NUE) of this plant. The resulting benefits extend to farmers by reducing the necessity for expenditures on fertilizers. Overuse of fertilizers is linked to various health issues, and minimizing their usage not only eliminates associated

health risks but also allows the medical sector to harness the advantages of the pulse plant more effectively for the benefit of humanity [10]

Chapter 2

Literature Review

2.1 Mungbean

Mungbean, a widely consumed pulse globally, especially in Asian nations, boasts a rich history of traditional medicinal use. Recognized as an excellent source of protein, dietary fiber, minerals, and vitamins, it also contains significant levels of bioactive substances like polyphenols, polysaccharides, and peptides. These nutritional qualities position mungbean as a popular functional food, contributing to overall health and well-being. Studies have shown that mungbean can have positive effects on hyperglycemia, hyper-lipidemia, and hypertension. Additionally, it has demonstrated potential in cancer prevention, melano-genesis inhibition, and exhibits hepato-protective and immuno-modulatory properties. The health benefits associated with mungbean are attributed to the concentration and characteristics of active chemicals present in the pulses. Notably, key polyphenols identified include vitexin and isovitexin, while peptides featuring hydrophobic amino acid residues with low molecular weight contribute to increased bioactivity in mungbeans [9]. The ongoing study aimed at identifying the full genome profile of mungbean is expected to enhance our understanding of its economic importance as a pulse crop. This, in turn, can contribute to improving the Nitrogen Use Efficiency (NUE) of the crop, further highlighting its significance in agricultural practices.

2.2 Production of Mungbean in Pakistan

Pakistan is one of the countries that cultivates mungbeans (*Vigna radiata*), commonly known as "moong dal" in Pakistan. Mungbeans are an essential part of Pakistani cuisine and are grown in various regions of the country. Mungbeans are grown in several provinces of Pakistan, with Punjab, Sindh, and Khyber Pakhtunkhwa (KP) being the major producing areas.

The specific regions within these provinces can vary based on climate and soil conditions. Mungbeans are well-suited to Pakistan's climate, as they thrive in warm, tropical, and subtropical regions. The crop is typically sown during of pulse cthe Kharif season (summer) when temperatures are conducive to its growth. Mungbeans are drought-tolerant and may be grown in both rainfed and irrigated environments.

Farmers cultivate mungbeans using a variety of agricultural practises, including seedbed preparation, seeding, and pest management. Climate, soil quality, and agricultural practises all have an impact on mungbean production.

Under favourable conditions, yields per acre or hectare can range from several hundred kilogrammes to more than a tonne. if the present growth rates of area allocated to mungbean remain the same then the area of mungbean in Pakistan would predict to decline from 165,000t in 2012 to 116,000t in 2016. Similarly, yield analysis revealed that that if the present growth rates of mungbean yield remains the same then its yield would be increase from 703.75 kg/h in 2012 to 807.82 kg/h [10].

Forecasted values show that there is an increasing trend in yield of mungbean in Pakistan, however predicted values of area under this crop show decreasing trend.

TABLE 2.1: Comparative profitability analysis of recommended mungbean varieties at NARC experimental station, Islamabad, Pakistan [10]

Particular	**Punjab**	**Sindh**	**KPK**	**Balochistan**	**Pakistan**
Area (%)	76.8	7.86	6.18	9.17	100
Production (%)	78.74	5.77	6.69	8.79	100
Yield ($kgha^{-1}$)	569.00	407.00	600.00	532.00	555

2.3 Marketing of Mungbeans in Pakistan

Mungbean output Pakistan in for 2022-23 is predicted to reach 183.7 thousand tonnes from an area of 213.2 thousand hectares, representing a 12.58% and 2.29% growth in area and production, respectively, above the previous objectives. India is the major mungbean grower, accounting for over 60% of total global production, followed by China and Mayanmar [11]. Government policies, subsidies, and support programs for farmers and traders can impact the marketing of mungbeans. Initiatives that improve infrastructure, provide market information, and offer financial support can enhance the efficiency of the mungbean market [11].

2.4 Worldwide Export

India is one of the largest exporters of mungbeans globally. It produces a substantial quantity of mungbeans and exports them to various countries, including neighboring nations and international markets. Myanmar is another major exporter of mungbeans, particularly to countries in Southeast Asia and the Middle East. Thailand is known for its mungbean exports ,often shipping to countries in Asia and beyond. Australia exports mungbeans to several countries, primarily in Asia and the Middle East [12].

Mungbeans are exported to a variety of countries, with significant demand in Asian countries such as China, Bangladesh, Pakistan, Vietnam, and Indonesia. Middle Eastern countries like Saudi Arabia and the United Arab Emirates are also substantial importers of mungbeans. Western countries, including the United States and Canada, import mungbeans to meet the demand of their multicultural populations and the growing interest in plant-based diets. The global mungbean area is around 7.3 million hectares, with an average yield of 721 kg/ha Each contribute for 30% of the worldwide output of 5.3 million t. China, Indonesia, Thailand, Kenya, and Tanzania are other significant producers. Dry grains (important in South Asia and Kenya), sprouts (important in East and Southeast Asia), transparent

noodles/starch (important in East and Southeast Asia), and paste (important in East Asia) are the four main divisions of the mungbean market [12].

2.5 Worldwide Import

The import of mungbeans (also known as green gram or moong dal) is a global trade, and it can vary in terms of volume and destination depending on the year and market demand. Mungbeans are commonly imported by countries for various purposes, including food consumption, processing, and cultivation. Imports In 2021, the major importers of dried black or green gramme beans were India ($632 million), China ($318 million), Vietnam ($232 million), Indonesia ($118 million), and Japan ($73 million) [13]. In January 2022, specific real-time percentages for mung bean imports might fluctuate and vary from year to year based on factors such as production, demand, market conditions, and government policies. Import percentages are subject to change due to global market dynamics and country-specific agricultural and trade situations [13].

In any case, it is significant to recognize that the continuous rates for mungbean imports in January 2022 can go through changes, differing from one year to another. These variances are dependent upon variables, for example, creation levels, shopper interest, winning economic situations, and legislative strategies. Import rates stay dynamic because of the always changing scene of worldwide business sectors and country-explicit impacts, including farming practices and exchange approaches [13]. Subsequently, a complete comprehension of these factors is fundamental for partners engaged with the worldwide mungbean exchange to settle on informed choices and adjust to developing business sector elements.

2.6 Industrial Use of Mungbean

Mungbeans, primarily known as a food crop, also have some industrial uses, although these applications are relatively limited compared to their use as a food

source. Mungbeans contain starch, and this starch can be extracted and used in various industrial processes. Mungbean starch is used in the production of noodles, vermicelli, and other food products. It can also be utilized in the textile and paper industries for sizing and finishing. Mungbeans are sometimes grown as a cover crop or green manure in agricultural practices. When they are grown and then plowed back into the soil, they can enrich the soil with nitrogen and organic matter, improving soil fertility and structure. This can be especially valuable in sustainable farming systems. Mungbean starch has been explored as a potential source for the production of biodegradable plastics and packaging materials. Researchers are investigating the use of plant-based starches, including mung bean starch, to create more environmentally friendly alternatives to traditional plastics [14].

2.7 Medicinal Use of Mungbean

Mungbeans, scientifically known as *Vigna radiata*, have been used in traditional medicine for centuries in various cultures around the world. While they are primarily consumed as a food source, mungbeans are also recognized for their potential medicinal properties [12].

2.8 Digestive Health

Mungbeans are considered easy to digest and are often recommended for individuals with digestive issues. They are known to soothe the stomach lining, reduce acidity, and alleviate symptoms of indigestion [12].

2.9 Antioxidant

Mungbeans contain antioxidants such as vitamin C, vitamin E, and various phytonutrients. Antioxidants help protect cells from damage caused by free radicals

and oxidative stress [12].

2.10 Weight Management

Mungbeans are low in calories and high in fiber, making them a good addition to a weight loss or management plan. They can help promote a feeling of fullness and reduce overall calorie intake [13].

By consolidating mungbeans into a fair eating routine, people can bridle their capability to help weight the board objectives, offering a nutritious and fulfilling elective chasing a sound way of life.

2.11 Anti-inflammatory

Mungbeans contain compounds with anti-inflammatory properties. They may help reduce inflammation in the body, making them potentially beneficial for conditions like arthritis[13].

The consideration of mungbeans in the eating routine hence presents a charming road for investigating regular calming systems, advancing by and large prosperity and possibly adding to the administration of fiery circumstances.

2.12 Respiratory Health

Mungbeans are used in some traditional remedies to alleviate respiratory issues like coughs and congestion [14].

2.13 *Vigna radiata Cultivation*

Mungbeans are widely cultivated in tropical and subtropical regions of Asia. They thrive in warm climates and are commonly grown in countries like India, China, Myanmar, and Thailand. They are known for their adaptability to various soil types. Mungbeans are highly nutritious. They are a good source of protein, dietary fiber, vitamins (such as B vitamins and vitamin C), and minerals (including potassium, magnesium, and iron). Mungbeans are often considered a healthy food choice due to their nutrient content [15].

TABLE 2.2: Botanical *Vigna radiata* [15]

Botanical Name	*Vigna radiata*
Common Name	Mungbean
Family	Legumes
Origin	India
Foliage	The leaves are alternate, trifoliolate with elliptical to ovate leaflets.
Benefits	They may protect against heat stroke, aid digestive health, promote weight loss, and lower "bad" LDL cholesterol, blood pressure, and blood sugar levels.

2.14 Environmental Factors Affecting Plant Growth & Yield

Abiotic stress refers to environmental factors that hinder plant growth and productivity. The escalating global population has posed a threat to global food security, prompting scientists to investigate the factors contributing to declining crop production. In response to the increased demand for food, researchers have assessed the number of crops lost due to environmental factors, some of which have irreversible effects on cultivated areas, impacting both crop quality and quantity. In

an ever-changing environment, various unfavorable elements, including heat, cold, drought, and salinity, can affect both agricultural land and crop productivity [14].

Nitrogen, a crucial nutrient for plant growth and development, is strongly associated with several abiotic stress responses. Insufficient nitrogen inhibits plant growth, leading to stunted growth, reduced leaf size, chlorosis (leaf yellowing), and lower yield. Abiotic stresses such as salinity, drought, and extreme temperatures affect nitrogen absorption and assimilation in plants. Plants, being sessile organisms, have evolved effective mechanisms to regulate nitrogen to promote development when faced with various stresses [15].

However, excess nitrogen can also be detrimental to plants, resulting in overly lush growth with weak stems, increased susceptibility to diseases, delayed blooming, and reduced fruit set. Moreover, nitrogen runoff can lead to environmental issues such as water contamination [15]. The delicate balance of nitrogen in plants highlights the complexity of managing nutrient levels to optimize growth and mitigate the impact of abiotic stresses.

2.15 Role of Nitrogen in Plants

The macronutrient nitrogen is essential for plants because it is a component of amino acids, which are the building blocks of enzymes and proteins in plants [16].

2.15.1 Protein Synthesis

Nitrogen is a critical component of proteins, which are essential for plant structure and function. Proteins are involved in various biological processes, including enzymes that facilitate biochemical reactions, structural proteins in cells, and transport proteins that move molecules within plants [16].

Moreover, proteins act as fundamental underlying components in cells, giving the system to cell design. The significance of nitrogen in protein blend is additionally underlined by its part in the arrangement of transport proteins, working with

the development of atoms inside plants and adding to the multifaceted trap of natural pathways that oversee plant development, improvement, and physiological capabilities [16].

2.15.2 Nucleic Acid Synthesis

Nitrogen is a component of nucleic acids (DNA and RNA), which carry genetic information and are vital for cell division, growth, and reproduction [16].

Nitrogen's presence in the construction of nucleic acids highlights its importance in the hereditary outline of plants. By adding to the piece of DNA and RNA, nitrogen turns into a vital participant in guaranteeing the uprightness and usefulness of these nucleic acids, subsequently impacting the fundamental components that oversee the many-sided dance of cell life and generation in the plant realm.

2.15.3 Chlorophyll Production

Nitrogen is a component of chlorophyll, the pigment essential for photosynthesis. Chlorophyll captures light energy used in photosynthesis to produce sugars and other organic compounds necessary for plant growth [16].

Nitrogen's presence in chlorophyll features its importance in working with the change of sun oriented energy into substance energy, a central cycle that energizes the plant's metabolic exercises and fills in as the foundation for the complex snare of life inside the plant realm.

2.15.4 Enzyme Function

Many enzymes involved in metabolic reactions require nitrogen for their structure and function. These enzymes facilitate various biochemical processes, including respiration, nitrogen metabolism, and synthesis of hormones [16].

By adding to the sub-atomic organization of these chemicals, nitrogen turns into a fundamental player in controlling and working with the complicated organization of metabolic pathways that support the plant's physiological reactions, development, and in general metabolic homeostasis.

2.15.5 Plant Growth and Development

Adequate nitrogen levels promote healthy and vigorous plant growth. It influences leaf and stem development, root growth, and overall plant architecture [16].

By adding to the blend of fundamental biomolecules, nitrogen turns into a critical determinant in the mind boggling dance of cell processes that oversee plant development. The effect of nitrogen reaches out past simple food, effectively molding the morphology and underlying elements of the plant, in this way affecting its general wellbeing and formative direction.

2.15.6 Regulation of Metabolic Processes

Nitrogen regulates numerous metabolic pathways, including the synthesis of amino acids, which are building blocks of proteins, and the production of secondary metabolites involved in defense mechanisms against pathogens and pests [16].

By complicatedly tweaking these metabolic pathways, nitrogen arises as an expert controller, guiding the fragile harmony between development, guard, and by and large metabolic congruity inside the plant realm.

2.15.7 Stress Response

Nitrogen availability influences a plant's ability to withstand environmental stresses, such as drought, salinity, and disease. Proper nitrogen nutrition enhances a plant's resilience to stress conditions [16].

Generally, nitrogen arises as an imperative calculate the complex interaction between a plant's dietary status and its capacity to mount powerful reactions to natural difficulties, in this way advancing a strong pressure reaction system for supported plant wellbeing and essentialness.

2.16 Regulation of Nitrogen in Plants

The most essential nutrient, nitrogen (N), is obtained by roots in the greatest quantities and is necessary for plants to complete their life cycles [17]. Plants obtain N through their roots throughout their complete life cycle, and the total amount and types of this macronutrient's availability have a considerable influence on how plants grow and interact with their environment. The presence of nitrogen (N) in agricultural soils is a crucial factor frequently limiting crop productivity. Therefore, despite the negative consequences on ecosystems and substantial socioeconomic costs, nitrogen fertilizers are widely used worldwide. Therefore, to lessen the global impact of these anthropogenic activities, it is important to improve the existing understanding of N nutrition in plants. Nitrogen is found in soil in both organic and inorganic forms, mostly as free amino acids, urea and short peptides [18].

Inorganic forms of nitrogen include nitrate (NO_3^-) and ammonium (NH_4^+). Due to soilheterogeneity and dynamic microbial conversions both of which are impacted by agronomic methods and environmental factors the accessibility of these nutrients by roots changes greatly across time and place. While NH_4^+ concentrations normally vary between 20 and 200 M, NO3+ is the most prevalent form in aerobic soils, with values between 1 and 5 mM [18].

Moreover, it has been demonstrated that the availability of amino acids, even at extremely low concentrations, can influence plant development and root structure. Similar to this, the finding of urea transporters found in plasma membrane (PM)

of root cells has raised awareness of a direct use of this nutrient by plants. Urea-based formulations make up over 50% of all nitrogen fertilizer used agriculture [19].

However, although NH_4^+ is strongly absorbed by soil particles and slowly released, NO_3^- is quickly leached. On the other hand, urea and free amino acids are often present in concentrations between 1 and 150 M and 70 M, and only make up a small percentage of then that is readily available for crops [18]. It is debatable if organic N is important for crop nutrition. Although there is evidence that plants can obtain proteins, peptides, and amino acids from soil, a significant impact in crop fields has been ignored. But other research in recent years has highlighted many questions.

The first observations demonstrated that availability of just amino acids supports plant growth, and soon after that, root transport mechanisms for amino acid uptake were characterized, their molecular foundations having been partially clarified in *A. thaliana* and verified in crops. Moreover, it has been demonstrated that the availability of amino acids, even at extremely low concentrations, can influence plant development and root structure. Similar to this, the finding of urea transporters found in plasma membrane (PM) of root cells has raised awareness of a directuse of this nutrient by plants. Urea- based formulations make up over 50% of all nitrogen fertilizer used agriculture [19].

Seed germination, plant growth, root and leaf functionality, hormonal balance, and seed production are all impacted by the total amount of N available and the forms delivered. The significance of understanding N nutrition as a composite scenario, taking into consideration each nutrient's input, their reciprocal interactions, and the variety of impacts on plant metabolism, has been emphasized in recent literature. Naturally, the "-omic" techniques seemto sufficiently meet this need given their inherent holistic character. The two main adaptations to N availability in roots are changes in absorption activity and modifications to the root system architecture (RSA), both of which are related to the capacity of N forms to function as nutrients and/or regulatory signals for plant development and metabolism [19].

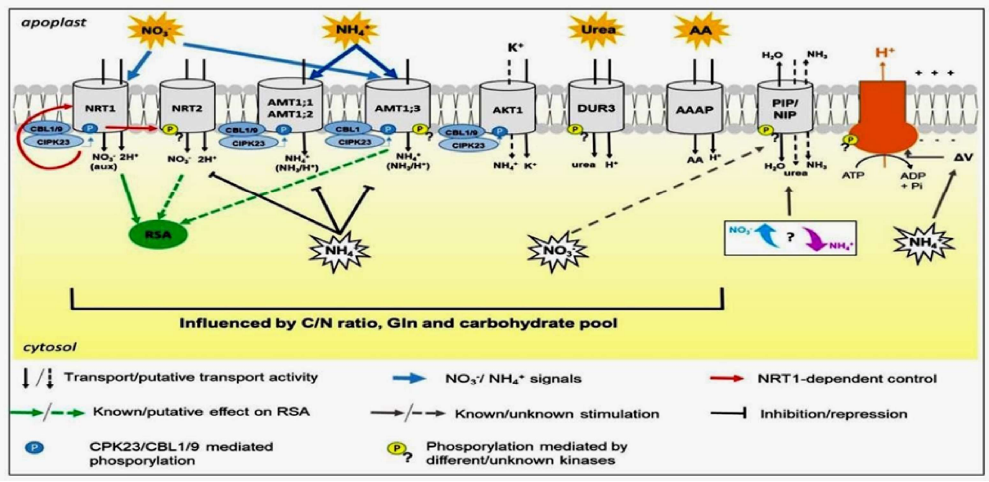


FIGURE 2.1: Using *Arabidopsis thaliana* as a model

The figure summarizes the metabolic relationships and regulatory pathways among the major transporters involved in nitrogen uptake by roots. The same root cell is reported to have transceptors, aquaporins, transporters, and H+-ATPase in order to maintain simplicity, however this is not the typical real-world scenario [19].

Proteins found in the PM of root cells are involved in the transportation and sensing of nitrogen forms. The understanding of the transcriptional mechanisms controlling uptake as well as the components engaged in signalling has significantly increased in recent years.

It was also highlighted that quick adaptations to unexpected changes in nitrogen availability depend heavily on post- translational modifications (PTMs) of transporters, such as phosphorylation events and protein complex formation. N uptake is associated with other important PM functions, such as the electrochemical proton gradient and water homeostasis, from a physiological aspect. The identification of proteins involved in N uptake has been the topic of numerous research, with Arabidopsis getting the most thorough molecular analysis.

Although various parallelisms in crops have been identified, we picked Arabidopsis as our starting point [19]. However, due to their multiple unique characteristics, the description of these aspects in rice (*Oryza sativa*), actinorhizal crops, and legumes (*Fabaceae spp.*) is outside the scope of our research [19].

2.17 Role of Transporters & Transceptors in Nitrogen Uptake by Roots

The availability of NO_3^- and the plant's nutritional condition for N strictly control these proteins' activity. The normal NO_3^- primary reaction (NPR), which includes the rapid induction of NO_3^- transporters and nitrogen absorbing enzymes, is triggered in N-starved plants when NO_3^- becomes available again. After this adaptation, there is a further down regulation of absorption that is connected with the buildup of NO_3^- and its later metabolites i.e.: glutamine. In Arabidopsis and crops, a significant function for the transcriptional regulation of NO_3^- transporters was shown, and the significance of PTMs was emphasized [20]. The Arabidopsis root cells' PM contains three members of the NPF family which is engaged in N03 absorption. While NRT1.1 (AtNPF6.3) and NRT1.2 (AtNPF4.6) both contribute to NO3 influx, with the latter being primarily in charge of the constitutive influx in the low- affinity range (>0.25 mm), NAXT1 (AtNPF2.7) mediates NO3 efflux to the external medium [21].

2.17.1 Role of Transporters & Receivers in Nitrogen Uptake by Roots

Three key protein types, namely transporters, receptors, and transmitter, interact with each other to finely regulate and influence the intake of N nutrients.

2.17.2 Uptake of Nitrate

The availability of NO_3^- and the plant's nutritional condition for N strictly control these proteins' activity. The normal NO_3^- primary reaction (NPR), which includes the rapid induction of NO_3^- transporters and nitrogen absorbing enzymes, is triggered in N-starved plants when NO_3^- becomes available again. After this adaptation, there is a further down regulation of absorption that is connected with the buildup of NO_3^- and its later metabolites i.e.: glutamine. In Arabidopsis and

crops, a significant function for the transcriptional regulation of NO_3^- transporters was shown, and the significance of PTMs was NO3 influx, with the latter being primarily in charge of the constitutive influx in the low- affinity range (>0.25 mm), NAXT1 (AtNPF2.7) mediates NO3 efflux to the external medium [21].

2.17.3 Uptake of Ammonium

Ammonium is known as essential nitrogen nutrient that plants can quickly absorb and assimilate. It is termed as signalling molecule that affects RSA and growth of plant. Nutritional conditions have a significant impact on how much NH_4^+ is taken in by plants and distributed across their organs. NH_4^+ concentrations in the cytosol and vacuole are thought to range from 1 to 10 mM and 1 to 45 mm, respectively, at the cellular level. It's interesting to note that in roots, the apoplastic NH_4^+ concentration is buffered at 1 to 2 mm under both low and high NH_4^+ supplies [21]. Ammonia ions (NH_4^+) may be found in soil and are absorbed by plants through their roots. Transport proteins in the root cell membranes enable this absorption.

2.18 Impact of Nitrogen Use Efficiency (NUE)

The most important nutrient that limits crop output is nitrogen, which is also a part of the structure of DNA and proteins. Therefore, the bulk of farmers rely on nitrogen fertilizers to increase agricultural output and financial returns. However, N can be lost through a number of processes, including as ammonia (NH3) volatilization, nitrate (NO_3^-) leaching, de- nitrification losses as di-nitrogen (N2) gas emissions, and nitrous oxide (N20) emissions, all of which contribute to environmental degradation and climate change. Recent studies highlight the significant contribution of agriculture sector contributes significantly to N loss to the environment. Population growth and rising food costs will increase the demand for mineral fertilizer in the future, which could result in an increase in nitrogen losses

unless significant changes are made to the entire food production-consumption chain and more efficient nitrogen management techniques are developed [22].

Nitrogen utilization efficiency refers to the complex processes of nitrogen intake, translocation, assimilation, and remobilization (NUE). It also displays how much cotton lint and seed output have increased as a result of the application of nitrogen. NUE of cotton is determined by nitrogen uptake efficiency (UpE) and utilization efficiency (UtE). Lint yielded collected data after the nitrogen application served as a proxy for cotton NUE. UtE is calculated by dividing the cotton yield ratio by the total plant N, and UpE is the total N absorbed by plants following N treatment. Plant nitrogen UpE is more significant than nitrogen UtE when there is a lack of nitrogen. Like other nutrient types, nitrogen is also a mobile nutrient in soil and is more susceptible to leach [23].

2.19 Enhancement of Nitrogen Use Efficiency (NUE)

Improved irrigation techniques, better fertilizer administration that takes into account the 4Rs, and the use of hybrids with higher potential yields and fewer N inputs are all connected with improvements in nitrogen use efficiency (NUE). In cropping systems, efficient management, or N source, rate, time, and placement, increases NUE. Overall NUE is improved by nitrogen inhibitors, split nitrogen applications, irrigation timing, and fertilizer placement techniques that take into account the crop variety and the soil. Recently, it is observed that late N application increased N accumulation at the stage of boll-setting. NUE is also considered to be an essential element of fertilizer inputs to any agricultural system since it preserves the N balance between inputs and outputs without harming the economy or the environment. Giving rose plants nitrogen fertilizer at the initial bloom stage is another way to increase NUE since rose plants use N more effectively for reproduction [24].

Several strategies can improve NUE:

2.19.1 Precision Agriculture

Using technology to precisely apply fertilizers according to plant needs can prevent over-application, reducing nitrogen runoff and leaching [24].

2.19.2 Optimizing Fertilizer Management

Employing slow-release or controlled-release fertilizers helps synchronize nitrogen availability with plant demand, minimizing losses [24].

2.19.3 Crop Rotation and Cover Crops

Alternating crops or using cover crops can improve NUE by diversifying root structures, enhancing nutrient uptake, and reducing nitrogen leaching [24].

2.19.4 Nitrogen-Fixing Plants

Introducing leguminous crops or utilizing nitrogen-fixing bacteria in the soil can increase available nitrogen without relying solely on synthetic fertilizers [24].

2.19.5 Genetic Approaches

Breeding crops for improved nitrogen uptake efficiency or tolerance to low nitrogen conditions can enhance NUE [24].

2.19.6 Soil Health Improvement

Practices like no-till farming, mulching, and composting enhance soil structure, microbial activity, and nutrient availability, contributing to better nitrogen uptake by plants [24].

2.19.7 Precision Irrigation

Efficient water management can reduce nitrogen losses by preventing excess leaching caused by overwatering [24].

Improving NUE not only benefits crop productivity and reduces input costs for farmers but also helps mitigate environmental issues such as water pollution (from nitrogen runoff) and greenhouse gas emissions (from nitrogen-based fertilizers). Balancing agricultural productivity with environmental sustainability is crucial, and optimizing nitrogen use is a key aspect of achieving that balance [24].

2.20 Transcription Factors

Plant transcription factors typically consist of several structural components, including a DNA-binding region, an oligomerization site, a transcription-regulation domain, and a nuclear localization signal, with some exceptions. While certain transcription factors may possess two different types of DNA-binding and oligomerization domains, most exhibit only one type, sometimes in multiple copies. The normal arrangement is such that oligomerization sites and DNA-binding areas either overlap or are in close proximity, and the combined tertiary structure of these regions governs crucial elements of transcription factor activity. Many transcription factors feature pairs of nuclear localization signals, and similar to DNA-binding domains, basic amino acid residues are pivotal to their functionality.

Transcription factors are encoded by multigene families, and their members are distributed throughout the genome or grouped on a single chromosome. The evolution of transcription factor families may involve processes such as exon capture, gene duplication and mutation, and translocation, as suggested by distribution and sequence analysis. In plants, the activities of transcription factor genes are regulated during both the transcription process and afterward, whereas the activities of their protein by-products are regulated during the post-translational process [25]. This intricate regulation underscores the complexity of gene expression control in plants.

2.21 Role of Transcription Factors

TFs are regulatory proteins that bind to certain DNA regions to trigger transcription. The control of gene expression by transcription factors is considered to be combinatorial in eukaryotes since it necessitates the coordinated interactions of numerous proteins. Many housekeeping genes are required by practically every kind of cell and appear to be uncontrolled or constitutive. However, the control of gene expression in a tissue-specific way is at the heart of cellular differentiation, as shown in the multitude of cell types found in various species. The same genome is responsible for the production of all cell kinds, each with its unique function—for example, red blood cells exchange oxygen, muscle cells expand and contract, and immune system cells recognize invaders. To guarantee that a cell can execute its intended job, genes that govern cell identity are activated under highly particular temporal, geographical, and environmental circumstances [26]. Regulation at the transcriptional level appears to be the main method for regulating gene expression in eukaryotes. Transcription factors, which either activate or repress transcription, regulate the regulation of gene expression by transcription. These activators and repressors are themselves controlled in a variety of ways and function through a range of processes, including protein-protein interactions, DNA-protein interactions, and alteration of the chromatin structure. Several different mechanisms can be used by a single transcription factorto impact the transcription of several target genes [25].

Plant growth and development are directed by TFs, which influence processes such as cell differentiation, organ creation, blooming, and fruit production. Different TFs control different phases of development in distinct plant tissues [26]. Plants adapt continually to environmental changes such as light, temperature, drought, and infections. TFs play an important part in these responses by activating genes that help the plant adapt and survive in a variety of environments. TFs implicated in stress responses, for example, can influence the expression of genes involved in drought tolerance or pathogen defense. TFs control the expression of genes involved in the formation of secondary metabolites such as pigments, flavonoids,

and defense chemicals. They are also in charge of specialized activities such as hormone signalling pathways and food absorption [25]

TFs can build complex signalling networks with other regulatory proteins to coordinate plant responses. These interactions result in complex regulatory networks that affect gene expression in a coordinated and dynamic fashion [25].

2.22 Genes as Transcription Factor in Regulating Nitrogen Response

A family of transcription factors known as RWP-RK is specific to plants and is involved in nodulation, gametogenesis, and nitrate response. However, very little is known about the phylogeny, regulation, and genome-wide characterization of RWP-RK genes in species of nitrogen-fixing plants that nodulate and species that do not nodulate (NFC). As a result, total292 RWP-RKs, including 278 from 25 NFC species and 14 from outgroup, *Arabidopsis thaliana*, were identified [27]. The NIN-like proteins (*NLP*s) and the RWP-RK domain proteins make up two subfamilies of the 292 RWP-RKs that we categorized (RKDs). According to the RWP-RKs' transcriptome and phylogenetic analyses, the *NLP* genes were only slightly more elevated in nitrate response and nodulation than the RKD genes. Additionally, it's possible that AI*NLP* genes in A.thaliana and nodule-specific *NLP* genes in some nodulating NFC species shared a common ancestor (OG0002084). Moreover, A.thaliana co-expression networks under N-supplementation and - starvation conditions revealed that there is a stronger relationship between the expression of At*NLP* genes and symbiotic genes during N- starvation. It was proven that N-starvation caused nodulation in *P. vulgaris* by regulating the expression of Pv*NLP*2, which is closely related to AI*NLP*6 and AI*NLP* and has a different common ancestor (0G0004041). It was concluded that nodulation in NFC plant species would have formed as a result of distinct origins of the *NLP* genes responsible for both the N- starvation response and the specific expression of nodulation [27].

2.23 *NLP* Gene Role in Nitrogen Signaling

NIN-like proteins (*NLP*s) are important transcription factors for nitrate signalling. *NLP* genes are nitrogen-responsive in the soil. When nitrogen is scarce or plentiful, these genes are activated or repressed, acting as nitrogen availability sensors. The expression of genes involved in nitrogen absorption from the soil and nitrogen assimilation into organic compounds inside the plant is regulated by *NLP* transcription factors. They regulate the expression of nitrogen-metabolizing transporters, enzymes, and regulatory proteins.

According to reports, the nutrient-Ca2+-*NLP* regulatory pathway integrates transport, transcription systemic growth processes and metabolism in plants and plays a crucial part in nitrate signalling.

Nitrate transporter 1.1 (NPF6.3/NRT1.1) has been discovered to function as a nitrate sensor at the plasma membrane in Arabidopsis. The major nitrate response genes are activated by nitrate, and calcium-dependent protein kinases 10/30/32 (CK10/30/32) mediate Ca2+ signals by nitrate and phosphorylate *NLP*6/7 to secure their position in the nucleus [26]. The NIN protein, which controls the production of symbiotic root nodules, was first found in the legume *Lotus japonicus*. Other non-leguminous plants, such as Arabidopsis, rice, wheat, and maize, were shown to have a greater number of NIN proteins and *NLP*s, but not mammals [27]. The RWP-RK domain for DNA binding is present in both NIN proteins and *NLP*s, and *NLP*s also contain an extra PB1 domain for protein-protein interactions. NRG2, PCF (TCP)-domain family protein 20, and nitrate-inducible GARP-type transcriptional repressor 1 (NIGT1) are a few other transcription factors that have been shown to interact with *NLP*s. *NLP* genes affect root development and architecture by encouraging the formation of lateral roots and root hairs, which increase the surface area available for nitrogen uptake from the soil. They contribute to nitrogen utilization efficiency by modulating the expression of nitrogen utilization genes. This assists plants in adapting to fluctuating nitrogen availability, enhancing their capacity to use limited nitrogen resources effectively. *NLP* genes are involved in the plant's response to nitrogen availability signals,

such as nitrogen deficit or excess, and they control the plant's adaptability to these conditions [27].

It has been determined that *NLP*s have additional roles in addition to nitrate signaling, including those in the nitrogen starvation response, nitrogen and phosphate (P) interactions, nitrate-assisted seed germination, nitrate-dependent nodule symbiosis, and root cap cell release [27].

2.24 Characterization of *NLP* Gene Family in *Arabidopsis thaliana*

Among the significant transcription factors (TFs) engaged in the nitrate-related GRN (gene regulatory work) in *Arabidopsis thaliana*, NIN-LIKE PROTEIN 7 (*NLP*7), a member of the NIN-Like protein family, stands out as an orchestrator of the nitrate-regulated transcriptional response [25]. Nine *NLP*s are encoded in the Arabidopsis genome, and since all of them can bind the nitrate responsive cis- element (NRE) in vitro, it is believed that they all play a role in the nitrate response. Through a nuclear retention mechanism that activates early nitrate-responsive genes and includes the phosphorylation of a conserved S205 residue by calcium- dependent protein kinases, nitrate regulates the nuclear accumulation of Arabidopsis *NLP*7 and *NLP*O (CDPKs).

Despite being crucial for nitrate-promoted seed germination, *NLP*8 function is not controlled by nitrate-dependent nuclear retention because the protein is retained in the nucleus independent of nitrate exposure [28]. Recent research revealed that Arabidopsis *NLP*2 and *NLP*7 both play important roles in vegetative growth since they both affect rosette biomass when grown in the presence of nitrate. As a result, *NLP*2 and *NLP*7 share characteristics like nitrate-dependent activation while belonging to different evolutionary clades. *NLP*2's relative position and influence within the nitrate-related GRN, as well as how it influences nitrate-dependent growth, remain unclear [29]. In response to nitrate, *NLP*2 plays a significant function in controlling carbon and energy metabolism.

Additionally, it has been discovered that *NLP*2 has a role in the nitrate-dependent orchestration of N metabolism and other related activities, which partly depend on themolecular interactions between *NLP*2 and *NLP*T.

2.25 Characterization of the *NLP* Gene Family in *B. napus*

The conserved transcription factors known as NODULE-INCEPTION-like proteins are essential in plants' reactions to nitrogen shortage. In *B. napus,* there are 31 *NLP* genes, including 16 in the A subgenome and 15 in the C subgenome. Predictions of subcellular localization showed that the majority of Ba*NLP* proteins are found in the nucleus. According to phylogenetic study, the *NLP* gene family can be divided into three groups, and the ancestorof both monocots and dicots had at least three ancient copies of the *NLP* gene before they diverged [29]. It's possible that the Brassicaceae species underwent multiple duplication events with the origin of group III *NLP* genes. 14 amino acids of the Bna*NLP*7-1 protein were found to be involved in DNA binding, according to three-dimensional structural analysis; nevertheless, no binding sites were found.

2.26 Characterization of the *NLP* Gene Family in Tomato

Tomato (*Solanum lycopersicum*), one of the most important horticultural crops, has a definite preference for nitrate as an inorganic nitrogen source. An extensive analysis of tomato *NLP* genes. *NLP* genes are well conserved in tomato. The main driving force for the evolution of the SI*NLP* gene was segmental duplication. It is possible that some SI*NLP* genes contribute to functional divergence in gene families during evolution by positive selection. The patterns of their expression

lead to the SINIP genes' many physiological functions in controlling nitrate uptake, among other things, in tomato growth and development. More functional analysis will be needed to examine each SINI's regulatory roles, especially those of SI*NLP*3 and SINIP6 [32]. Uncovering the molecular basis of nitrogen utilization and enhancing nitrogen usage efficiency in tomatoes are anticipated to require a thorough understanding of roles of SI*NLP* in varying nutritional situations [33].

According to an examination of the expression profile, BnaNLP genes are expressed in the majority of organs, although they frequently express themselves strongly in just one. For instance, the four members of the BnaNLP7 subfamily are substantially expressed in leaves, whereas the members of theBnaNLP6 subfamily are largely expressed in roots. In response to nitrogen-deficient environments [30], BaNLP genes also displayed various expression patterns. All BnaNLP1/4/5/9 subfamily members were upregulated under nitrogen deprivation, all BaNLP2/6 subfamily members were downregulated, and BnaNLP7/8 subfamily members displayed diverse expression patterns in distinct organs. This study gives a complete evolutionary history of the NLP genes in B. napus as well as information on the biological activities of BaNIP genes in response to the nitrogen shortage [31].

2.27 Characterization of the *NLP* Gene Family in Rice

The NIN-LIKE PROTEIN 1 (OsNLP1) of rice (*Oryza sativa*) is essential for N consumption. ONLP 1 protein localized in the nucleus, and N starvation quickly increases its transcript level. Under various N circumstances, ONLP1 over expression enhances plant development, grain yield, and NUE, whereas ONLP1 deletion reduces grain yield and NUE under N- limiting conditions. By coordinating several nitrogen uptake and assimilation genes, ONLP1 controls how nitrate and ammonium are used. Chromatin immuno-precipitation and yeast one- hybrid experiments showed that ONLP may directly bind to the promoter of these genes to cause their expression. Thus, ONLPI is a viable target for increasing rice yield and

NUE as well as a significant regulator of nitrogen use, according to our findings [34].

2.28 Characterization of the NLP Gene Family in Maize

The identification and characterization of *Zea maize NLPs* (ZmNLPs) investigation showed that a total of 9 ZmNLPs from the maize genome were identified in accordance with conserved domains and gene structure. In eight different tissues at varying embryonic stages, almost all ZmNLPs were constitutively expressed, however the expression patterns of these genes varied between tissues. The benchmark N-responsive gene (ZmNRT2.2) was up- regulated more than fivefold at 0.5 h after treatment, while ZmNLP4 and ZmNLP5 showed the highest up- regulation, indicating they may be involved in the primary nitrogen response. The expression levels of four ZmNIPs (ZmNLP4, 5, 6, and 8) were over-regulated by a factorof two in response to N treatment, according to quantitative real-time PCR (qPCR) data. This indicates that ZmNLPs are important to the maize nitrogen response [33]. Therefore, the above research demonstrated that the NIN-LIKE PROTEIN (NLP) family of transcription factors, which bind to nitrate-responsive cis-elements (NRE), act as transcriptional to understand more about the physiological and molecular activities of *NLP* genes development [34]

Chapter 3

Materials and Methods

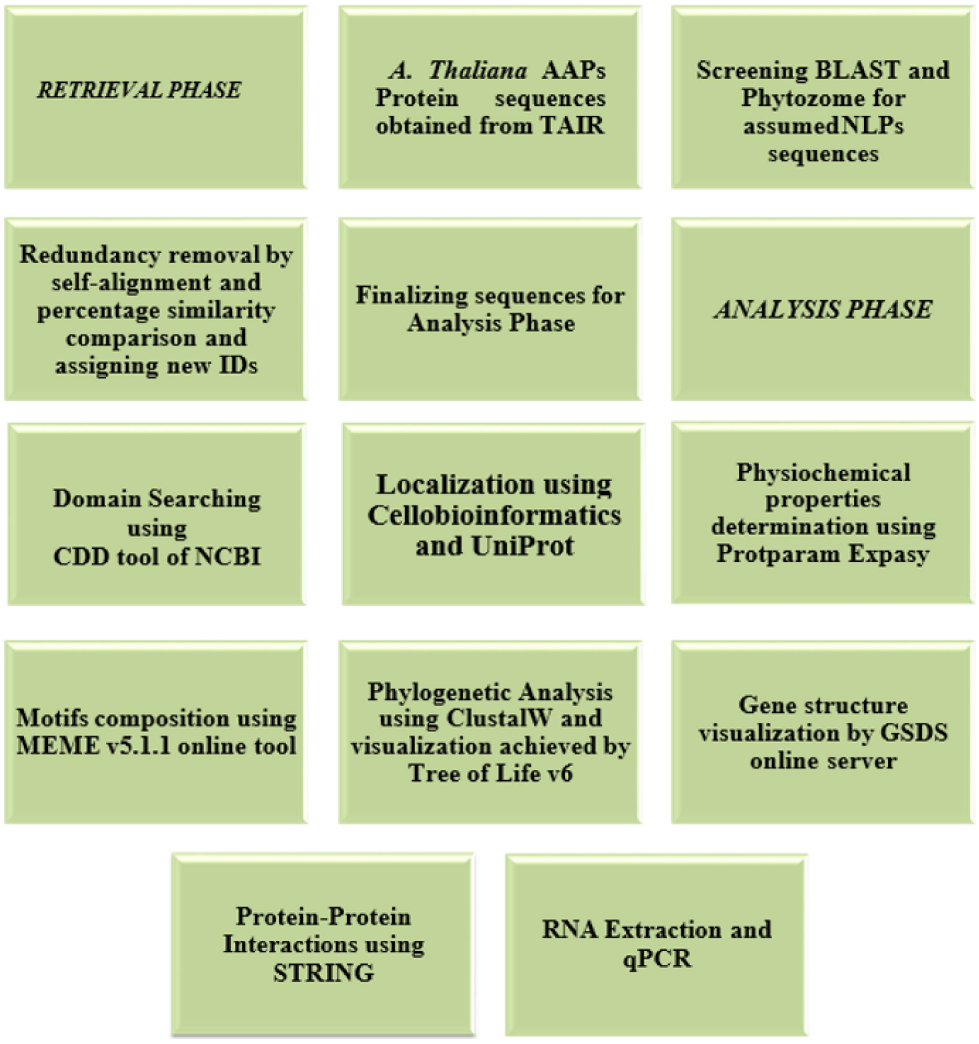


FIGURE 3.1: Overview of methodology

3.1 Screening of Genome and Transcription Factors Databases

The *Arabidopsis* genome resource TAIR (http://arabidopsis.org) was utilized to get the entire gene, amino acid, and coding sequences for each member of the *Arabidopsis thaliana NLP* gene family. The recognition of possible AtNIPs was examined using three protein databases: NCBI (https://www.ncbi.nlm.nih.gov/), TAIR (https://www.arabidopsis.org/), and TFDB (http://planttfdb.gao-lab.org/).

3.2 Vr*NLP* Protein Sequences Retrieval

The AtNLP protein sequences were employed as query sequences in two database i.e NCBI and TFDB to identify *NLP* protein sequences for *NLP* Vr*NLP*. Additionally, TAIR was screened using accession numbers to retrieve Vr*NLP* protein sequences. The sequences were obtained in FASTA format and stored for use in subsequent sampling and analyses. Two databases were screened for the identification and characterization of any putative Vr*NLPs* and all Vr*NLP* sequences were downloaded and aligned for further analysis.

3.3 Removal of Redundant Sequences of Vr*NLP*

The protein sequences of Vr*NLP* were aligned to eliminate redundancy. Repetition among the accession numbers of each sequence, obtained from different databases such as NCBI or TFDB, was removed. The sequences were rearranged, and matching was performed to identify identical sequences from different databases. This process ensures that no sequence is repeated, and the samples are specific and accurate.

3.4 Allocation of Lab IDs and Establishment of Sequence Similarity of Vr*NLP* Protein Sequences

At first, Lab-IDs were given to each sequence after the removal of redundant ,splice and incomplete sequences. Secondly, to determine the sequence similarities between the Vr*NLP* protein sequences a comparison table was generated. Sequences with similarity below the threshold (80%) would not be taken into consideration. Comparison table was generated. Sequences with similarity below the threshold (80%) would not be taken into consideration.

3.5 Physicochemical Properties and Conserved Domain Side Nitrification in VrNLPs

Potential Vr*NLPs* was chosen from the retrieved sequences based on the conserved domains. The chosen genes had both PB1 and RWPRK domains. Protparam Expasy (https://web.expasy.org/protparam/) was used to examine the physical and chemical characteristics of a few selected Vr*NLP*, including their protein molecular weight (MW), hydropathicity (GRAVY), and theoretical isoelectric point (pl), while Genscript their sub- cellular localizations was predicted via CELLO (http://cello.life.nctu.edu.tw/).

3.6 Motif Composition in Vr*NLP* Gene Family

The presence of consensus motif was determined using online tool called Multiple Em for Motif Elicitation or MEME v5.1 (https://meme-suite.org/meme/tools/meme). It is the most accurate online tool for motif elicitation.All parameter setting were kept at default settings with an exception of motif threshold,

which was kept at 15 to ensure specificity and precision. The identification of consensus motif will provide evidence of local regions that are similar between scale and study organism.

3.7 Phylogenetic Analysis of Vr*NLP*

To generate the phylogenetic tree for Vr*NLPs* genes, in which different organisms like *Arabidopsis thaliana*, *Brassica napus*, *Zea mays*, *Physcomitrella patens* and *Oryza sativa*, *L.* were aligned through Clustal Omega click here. Interactive Tree of Life software v6 (ITOL: https://itol.embl.de/) was used to create the phylogenetic tree of *Vigna radiata* with the selected plants to illustrate the evolutionary link between *NLP* and the other plants.

3.8 Gene structure and Motif Composition in Vr*NLP* Gene Family

The coding and full- length gene sequences of Vr*NLP* was used to examine gene structural components using GSDS online server (Gene Structure Display Server: http://gsds.cbi.pku.edu.cn/) The introns, exons, and un-translated regions (UTRs) were identified. This tools assisted in determining exons, introns and untranslated regions (UTRs) present in the sequence.This may also help in comparing gene length between two families

3.9 Protein-Protein Interaction of Vr*NLP*

The cellular proteins in interaction with Vr*NLP* and AtNLPs were predicted using an online tool called STRING (https://string-db.org/), and their figures along with the description, was exported from the tool for comparative study and future reference.The interacting proteins of Vr*NLPs* were compared with those

of AtNLPs to identify any functional homology.The colour lines show known and predicted interactions, the pink and blue colour shows known interactions and dark green, blue and red shows predicted interactions. These balls that are in different colours shows nodes while the coloured lines are the edges means protein sequences. In PPI, the proteins interacting with NLP are those involved in N uptake, assimilation & transport - thus they predict the role in overall NUE.

Chapter 4

Results

4.1 Genome-Wide Identification and Analysis of NLP Gene in *Vigna radiata*

In the current research, the full-length protein-coding sequences of the NLP gene from the *Arabidopsis thaliana* NLP gene family were retrieved by entering the accession number of NIP on the TAIR database (TAIR: http://arabidopsis.org/). There were 9 variants of the NLP gene and the longest sequence was chosen. Then the longest sequence was Protein Blast in the NCBI database (www.ncbi.nlm.nih.gov/genome/gdv/). The results were downloaded in description CSV format in an Excel file.

All the results were saved with name, accession numbers, and lab ids in an Excel file. They were self-aligned to make a percentage similarity table of all downloaded sequences After the similarity table, all the spliced variants, repeated/redundant sequences, and short or incomplete fragments were excluded from retrieved sequences simultaneously validated through conserved domain identification, and selected 8 sequences of *Vigna radiata* were downloaded in the FASTA format and a new Excel sheet of selected sequences with their lab ids, gene names, and accession numbers was made. NLPs in the *Vigna radiata* genome were identified using

the *Arabidopsis thaliana* NLP protein sequences as well as the pfam accessions of RWP-RK (PF02042) and PB1 domain (PF00564) as queries.

4.2 Physicochemical Properties & Conserved Domains Identification in VrNLP

In the final analysis, 7 VrNLPs were identified that contained both RWP-RK and PB1 domains (Table 4.1) and were labelled from 1 to 8 with respect to chromosome numbers. Accession numbers of same or redundant sequences found in selected databases are enlisted. Potential VrNLPs were chosen from the retrieved sequences based on the conserved domains discovered database by the CDD (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). All of these 8 sequences and query sequences were pasted in CD search in CCD and submitted.

The result is summarized in (Table 4.1) which includes query, hit type, position, E-value, bit-score, Accession, and short name of both plants. Whereas, the physical

and chemical properties of A. thaliana and Vigna radiata NLP gene families including protein molecular weight (MW), hydropathicity (GRAVY), and theoretical isolectric point (pl) values of both plants were examined in the Expasy ProtParam tool and are summarized in (Table 4.2). The gene lengths, protein lengths, and molecular weights, pl and GRAVY values of both plants were close to each other. The values of both plants showed negative GRAVY values which showed as hydrophilic proteins. The PI and GRAVY values of both plants are close to each other. While the sub-cellular localizations were examined which proposed A. thaliana and V. radiata of NLPs to be localized in nucleus.

AtNLPs and VrNLPs were found to have average protein lengths of 881 and 916 amino acids, respectively. However, there was significant variation between the gene lengths of AtNLPs and VrNLPs. With the exception of AtNLP3, all AtNLPs and VrNLPs showed pl values below 7, indicating that they are acidic proteins, while AtNLP3 pl value of 8.14 indicates that it is a basic protein.

TABLE 4.1: Conserved domains of NLP gene families of *A. thaliana* and *V. radiata*

Organism	**Query**	**Hit Type**	**Position**		**E-value**	**Bitscore**	**Accession**	**Short Name**
			From	To				
Arabidopsis thaliana	AtNLP1	specific	812	893	6.54E-41	112.802	cd06407	PB1_NLP
			608	656	2.62E-24	54.0518	pfam02042	RWP-RK
	AtNLP2	specific	896	944	1.26E-41	112.802	cd06407	PB1_NLP
			648	696	4.74E-24	54.0518	pfam02042	RWP-RK
	AtNLP3	specific	674	758	1.56E-40	112.802	cd06407	PB1_NLP
			498	546	1.69E-24	54.0518	pfam02042	RWP-RK
	AtNLP4	specific	745	826	7.15E-43	112.802	cd06407	PB1_NLP
			558	606	1.67E-24	54.0518	pfam02042	RWP-RK
	AtNLP5	specific	711	787	3.72E-36	112.802	cd06407	PB1_NLP
			549	597	5.58E-25	54.0518	pfam02042	RWP-RK
	AtNLP6	specific	742	822	3.01E-34	112.802	cd06407	PB1_NLP
			556	604	8.48E-25	54.0518	pfam02042	RWP-RK
	AtNLP7	specific	866	944	4.32E-34	112.802	cd06407	PB1_NLP
			591	639	4.66E-25	54.0518	pfam02042	RWP-RK
	AtNLP8	specific	848	928	6.25E-39	112.802	cd06407	PB1_NLP
			590	651	1.25E-20	54.0518	pfam02042	RWP-RK
	AtNLP9	specific	793	874	3.37E-34	112.802	cd06407	PB1_NLP
			535	584	5.08E-25	54.0518	pfam02042	RWP-RK
Vigna radiata	VrNLP1	specific	827	907	1.42E-42	112.802	cd06407	PB1_NLP
			607	655	7.44E-24	54.0518	pfam02042	RWP-RK
	VrNLP2	specific	685	766	3.76E-40	112.802	cd06407	PB1_NLP
			493	541	1.48E-23	54.0518	pfam02042	RWP-RK
	VrNLP 3	specific	648	729	6.18E-37	112.802	cd06407	PB1_NLP
			514	562	2.16E-23	54.0518	pfam02042	RWP-RK
	VrNLP 4	specific	945	1025	1.07E-38	112.802	cd06407	PB1_NLP
			635	683	8.88E-25	54.0518	pfam02042	RWP-RK
	VrNLP 5	specific	903	983	1.98E-36	112.802	cd06407	PB1_NLP
			595	643	4.56E-24	54.0518	pfam02042	RWP-RK
	VrNLP6	specific	867	947	1.54E-29	112.802	cd06407	PB1_NLP
			598	646	2.38E-24	54.0518	pfam02042	RWP-RK
	VrNLP 7	specific	879	959	6.83E-31	112.802	cd06407	PB1_NLP
			602	650	4.16E-24	54.0518	pfam02042	RWP-RK

The gene and coding sequences of AtAAPs and VrNLPs were used to analyze their structural features which included the identification of exons, introns and un-translated regions (UTRs). This was done using an online server called GSDS or Gene Structure Display Server (http://gsds.gao-lab.org/).

Table 4.2: Physicochemical properties of NLP gene families of A. thaliana and V. radiata

Plant	Gene name	Chromosome	Position	Gene Length (bp)	Protein Length	Molecular Weight	Isoelectric Point	Gravy	Localization
Arabidosis thaliana	AtNLP1	1	22324437-22327359	2933	909	100886.25	5.09	-0.443	Nuclear
	AtNLP2	4	16777264-16782054	4791	963	107278.54	5.65	-0.476	Nuclear
	AtNLP3	4	17954710-17958063	3354	767	85066.66	8.35	-0.271	Nuclear
	AtNLP4	4	7154410-7158287	3878	844	94231.93	5.45	-0.472	Nuclear
	AtNLP5	5	17367827-17369510	1684	808	90684.13	5.8	-0.467	Nuclear
	AtNLP6	4	950546-953690	3145	841	93863.46	5.95	-0.356	Nuclear
	AtNLP7	5	1698270-1702659	4390	959	105742.04	5.55	-0.42	Nuclear
	AtNLP8	2	18061716-18066708	4993	947	104883.79	5.47	-0.43	Nuclear
	AtNLP9	3	22009008-22012804	3797	894	98712.97	5.43	-0.383	Nuclear
Vigna radiata	VrNLP1	5	23422183-23427061	4879	920	102669.81	5.67	-0.48	Nuclear
	VrNLP2	5	24677902-24680551	2650	783	87014.05	5.99	-0.371	Nuclear
	VrNLP3	6	94900-97518	2619	731	81509.27	5.93	-0.351	Nuclear
	VrNLP4	8	40521825-40526233	4409	963	106382.02	5.33	-0.458	Nuclear
	VrNLP5	1	34948877-34954133	5257	979	108365.47	5.84	-0.351	Nuclear
	VrNLP6	2	3664175-3671010	6836	1002	111677.29	5.82	-0.409	Nuclear
	VrNLP7	Un	12240-20459	8220	1039	115185.97	5.61	-0.323	Nuclear
			At	3662.777778	881.3333333	97557.94	5.86	-0.413111111	
			Vr	4981.428571	916.7142857	101439.6063	5.741428571	-0.391857143	

The gene and coding sequences of AtAAPs and VrNLPs were used to analyze their structural features which included the identification of exons, introns and un-translated regions (UTRs).

4.3 Gene Structure Determination

The gene and coding sequences of AtAAPs and VrNLPs were used to analyze their structural features which included the identification of exons, introns and un-translated regions (UTRs). This was done using an online server called GSDS or Gene Structure Display Server (http://gsds.gao-lab.org/).

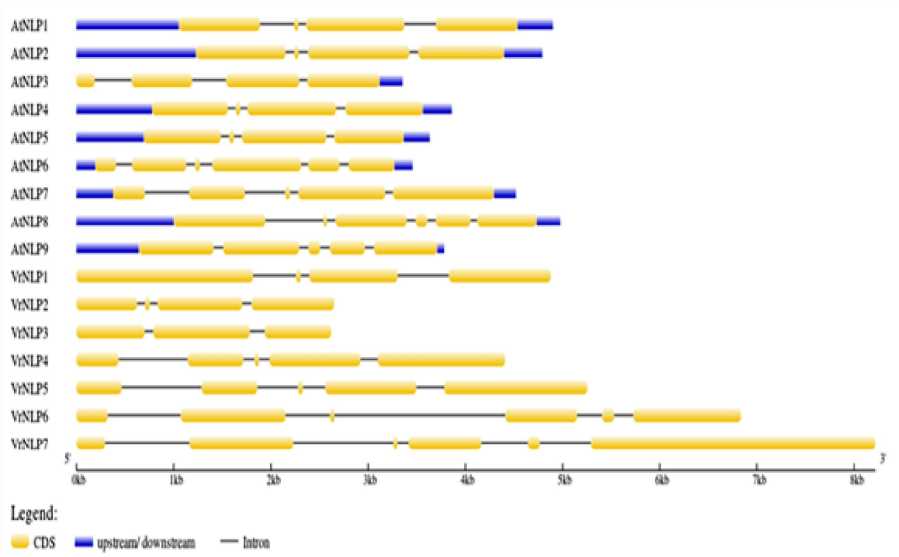


FIGURE 4.1: Displays the result of gene structure determination

4.4 Protein-protein Interaction of VrNLP

Protein don't work in isolated form they work in the form of a group. The network was predicted online through STRING ((https://www.expasy.org/resources/string). It has been proposed that all VrNLP protein interact with several N-interacted genes. The colour lines show known and predicted interactions, the pink and blue colour shows known interactions and dark green, blue and red shows predicted interactions. These balls that are in different colours shows nodes while the coloured lines are the edges means protein sequences. In PPI, the proteins interacting with NLP are those involved in N uptake, assimilation & transport - thus they predict the role in overall NUE. Through STRING (www.expasy.org/resources/string) the interacting NLP protein networks were predicted online. It has been proposed that

all VrNLP proteins interact with several N-related genes. (Table 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 4.10).

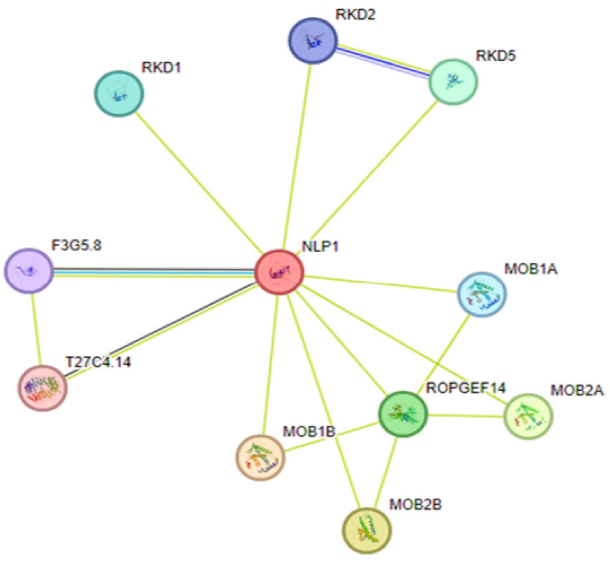


FIGURE 4.2: AtNLP1 protein-protein interaction

TABLE 4.3: AtNLP4 protein-protein interaction

	Predicted functional patterns
MOB1B	MOB kinase activator-like 1B; Belongs to the MOB1/phocein family.
MOB2B	Putative MOB kinase activator-like 2B.
MOB2A	MOB kinase activator-like 2A.
ROPGEF14	Rop guanine nucleotide exchange factor 14; Guanine-nucleotide exchange factor (GEF) that acts as an activator of Rop
NIR1	Ferredoxin–nitrite reductase, chloroplastic; Catalyzes the six-electron reduction of nitrite to ammonium
RFS2	Probable galactinol–sucrose galactosyltransferase 2; Transglycosidase operating by a ping-pong reaction mechanism.
RKD1	Protein RKD1; Putative transcription factor.

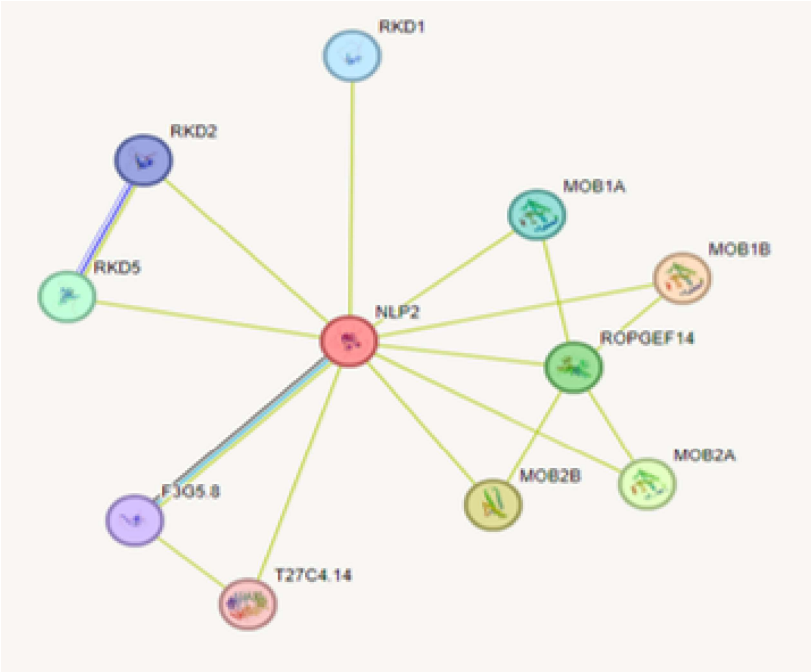


FIGURE 4.3: AtNLP2 protein-protein interaction

TABLE 4.4: AtNLP5 protein-protein interaction

	Predicted functional patterns
MOB1B	MOB kinase activator-like 1B; Belongs to the MOB1/phocein family.
MOB2B	Putative MOB kinase activator-like 2B.
MOB2A	MOB kinase activator-like 2A.
ROPGEF14	Rop guanine nucleotide exchange factor 14; Guanine-nucleotide exchange factor (GEF) that acts as an activator of Rop
MYB4R1	Putative transcription factor.
HRS1	Transcription factor HRS1; Transcription factor involved in nitrate and phosphate signaling in roots
RKD1	Protein RKD1; Putative transcription factor.

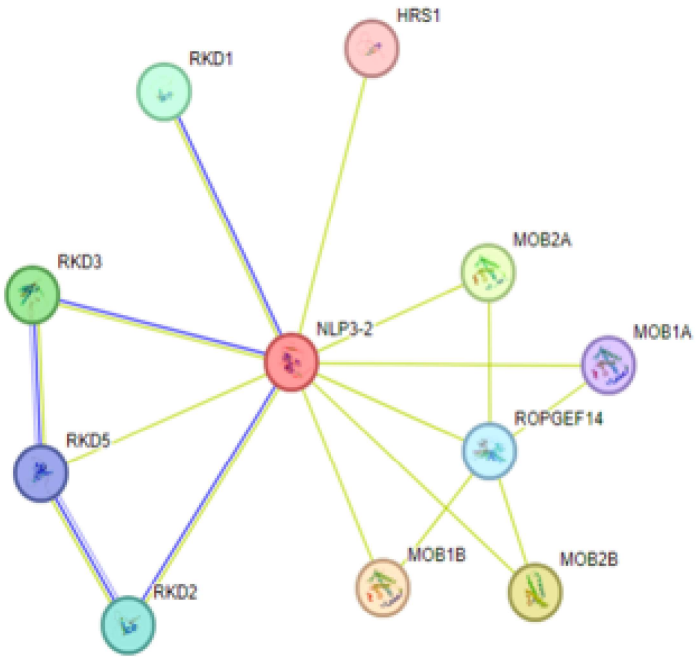


FIGURE 4.4: AtNLP3 protein-protein interaction

TABLE 4.5: AtNLP6 protein-protein interaction

	Predicted functional patterns
MOB1B	MOB kinase activator-like 1B; Belongs to the MOB1/phocein family.
MOB2B	Putative MOB kinase activator-like 2B.
MOB2A	MOB kinase activator-like 2A.
RKD3	Protein RKD3; Putative transcription factor.
RKD5	Protein RKD5; Putative transcription factor.
RKD1	Protein RKD5; Putative transcription factor.
OPGEF14	Rop guanine nucleotide exchange factor 14; Guanine-nucleotide exchange factor (GEF) that acts as an activator of Rop

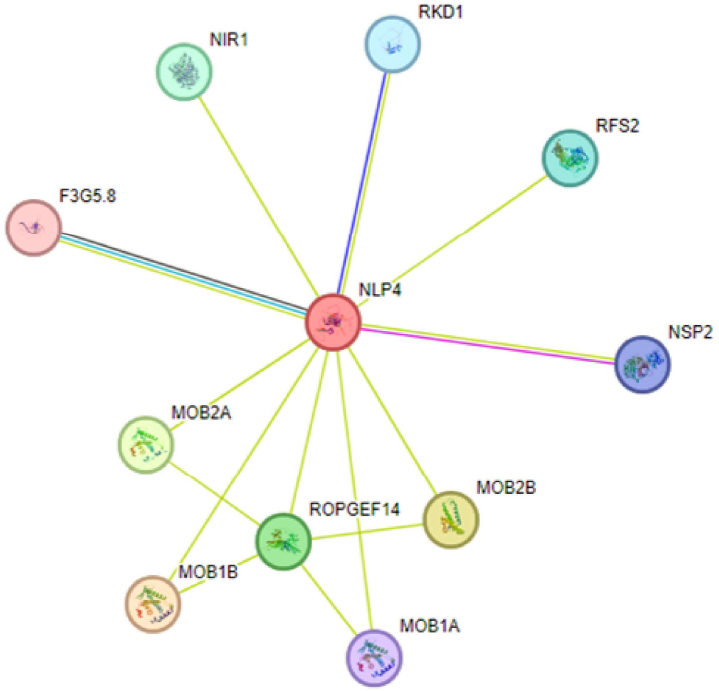


FIGURE 4.5: AtNLP4 protein-protein interaction

TABLE 4.6: AtNLP4 protein-protein interaction

	Predicted functional patterns
MOB1B	MOB kinase activator-like 1B; Belongs to the MOB1/phocein family.
MOB2B	Putative MOB kinase activator-like 2B.
MOB2A	MOB kinase activator-like 2A.
ROPGEF14	Rop guanine nucleotide exchange factor 14; Guanine-nucleotide exchange factor (GEF) that acts as an activator of Rop
NIR1	Ferredoxin–nitrite reductase, chloroplastic; Catalyzes the six-electron reduction of nitrite to ammonium
RFS2	Probable galactinol–sucrose galactosyltransferase 2; Transglycosidase operating by a ping-pong reaction mechanism.
RKD1	Protein RKD1; Putative transcription factor.

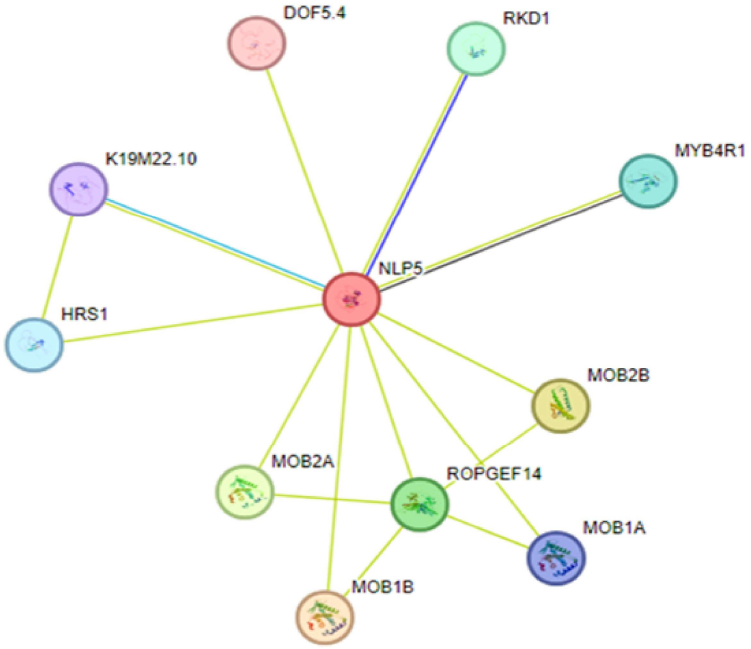


FIGURE 4.6: AtNLP5 protein-protein interaction

TABLE 4.7: AtNLP4 protein-protein interaction

	Predicted functional patterns
MOB1B	MOB kinase activator-like 1B; Belongs to the MOB1/phocein family.
MOB2B	Putative MOB kinase activator-like 2B.
MOB2A	MOB kinase activator-like 2A.
ROPGEF14	Rop guanine nucleotide exchange factor 14; Guanine-nucleotide exchange factor (GEF) that acts as an activator of Rop
MYB4R1	Putative transcription factor.
HRS1	Transcription factor HRS1; Transcription factor involved in nitrate and phosphate signaling in roots
RKD1	Protein RKD1; Putative transcription factor.

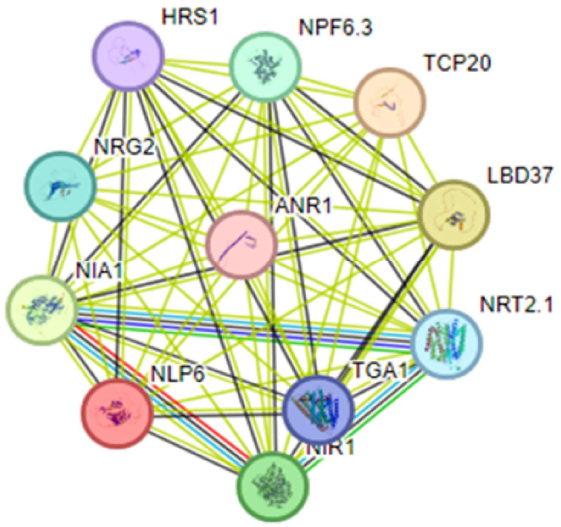


FIGURE 4.7: AtNLP6 protein-protein interaction

TABLE 4.8: AtNLP6 protein-protein interaction

	Predicted functional patterns
TCP20	Transcription factor TCP20; Transcription factor that binds to the site II motif (3'- TGGGCC/T-5') in the promoter of PCNA-2
LBD37	LOB domain-containing protein 37.
NIA1	Nitrate reductase [NADH] 1; Nitrate reductase is a key enzyme involved in the first step of nitrate assimilation in plants, fungi
NIR1	Ferredoxin–nitrite reductase, chloroplastic; Catalyzes the six-electron reduction of nitrite to ammonium.
NRG2	Nitrate regulatory gene2 protein; Required for nitrate signaling.
TGA1	Transcription factor TGA1; Transcriptional activator that binds specifically to the DNA sequence 5'-TGACG-3'.
HRS1	Transcription factor HRS1; Transcription factor involved in nitrate and phosphate signaling in roots.

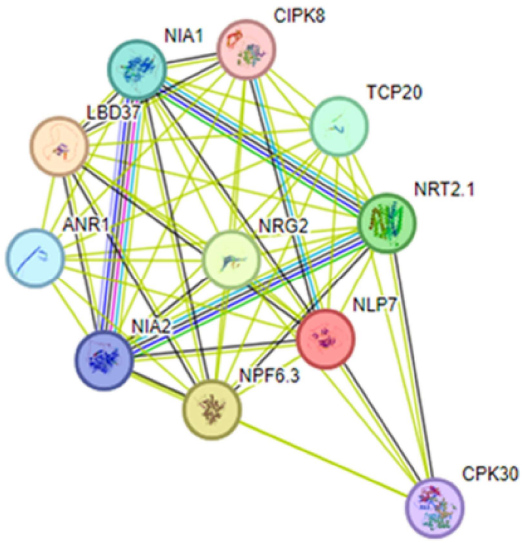


FIGURE 4.8: AtNLP7 protein-protein interaction

TABLE 4.9: AtNLP7 protein-protein interaction

	Predicted functional patterns
LBD37	LOB domain-containing protein 37.
NPF6.3	Protein NRT1/ PTR FAMILY 6.3; Dual affinity nitrate transporter. Involved in proton- dependent nitrate uptake.
NRG2	Nitrate regulatory gene2 protein; Required for nitrate signaling.
NRT2.1	High-affinity nitrate transporter 2.1; Involved in nitrate transport, but does not seem to be able to mediate transport by its own.
TCP20	Transcription factor TCP20; Transcription factor that binds to the site II motif (3'- TGGGCC/T-5') in the promoter of PCNA-2
NIA1	Nitrate reductase [NADH] 1; Nitrate reductase is a key enzyme involved in the first step of nitrate assimilation in plants, fungi.
ANR1	MADS-box transcription factor ANR1; Probable transcription factor. Required for root plasticity in response to nitrate, NO(3)(-)

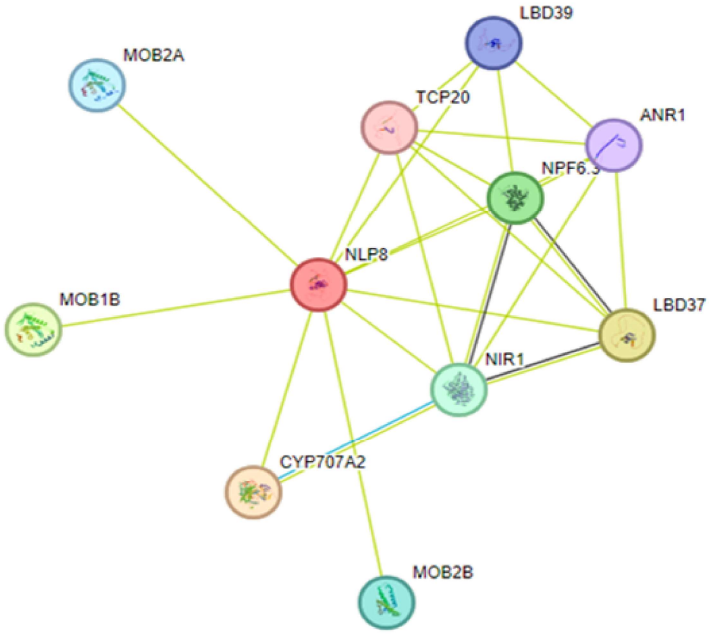


FIGURE 4.9: AtNLP8 protein-protein interaction

TABLE 4.10: AtNLP8 protein-protein interaction

	Predicted functional patterns
CYP707A2	Abscisic acid 8'-hydroxylase 2; Involved in the oxidative degradation of abscisic acid,
LBD37	LOB domain-containing protein 37.
MOB1B	MOB kinase activator-like 1B; Belongs to the MOB1/phocein family.
NPF6.3	Protein NRT1/ PTR FAMILY 6.3; Dual affinity nitrate transporter. Involved in proton- dependent nitrate uptake.
NIR1	Ferredoxin–nitrite reductase, chloroplastic; Catalyzes the six-electron reduction of nitrite to ammonium.
MOB2B	Putative MOB kinase activator-like 2B.
MOB2A	MOB kinase activator-like 2A.

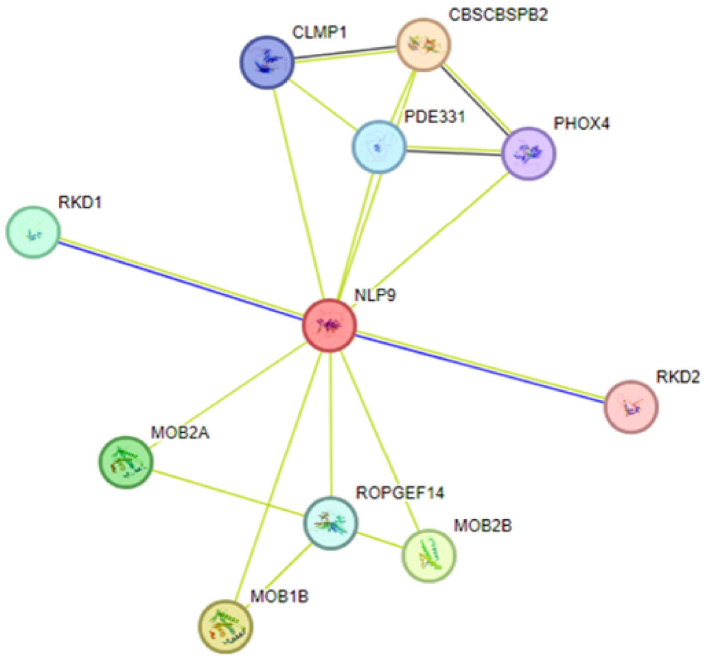


FIGURE 4.10: AtNLP9 protein-protein interaction

TABLE 4.11: AtNLP9 protein-protein interaction

	Predicted functional patterns
BSCBSPB2	CBS domain-containing protein CBSCBSPB2.
MOB1BA	MOB kinase activator-like 1B; Belongs to the MOB1/phocein family.
MOB2B	Putative MOB kinase activator-like 2B.
MOB2A	MOB kinase activator-like 2A.
RKD1	Protein RKD1; Putative transcription factor.
ROPGEF14	Rop guanine nucleotide exchange factor 14; Guanine-nucleotide exchange factor (GEF) that acts as an activator of Rop
PDE331	Octicosapeptide/Phox/Bem1p family protein.

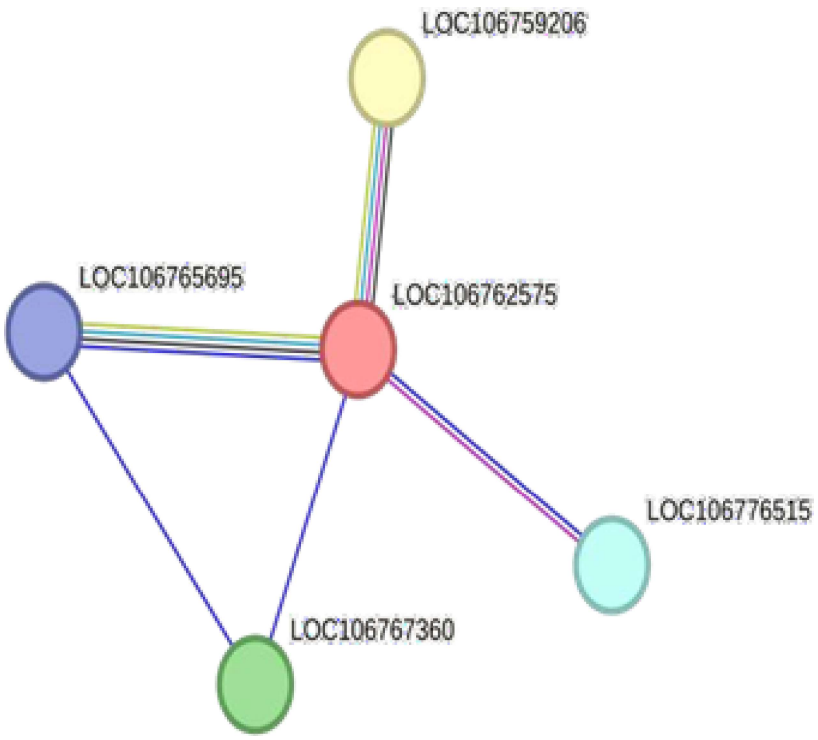


FIGURE 4.11: VrNLP1 protein-protein interaction

TABLE 4.12: Protein-protein interaction of XP_014502050.1

	Predicted functional patterns
LOC106759206	AP-2 complex subunit alpha. Subunit of adaptor protein complex 2(AP-2). Adaptor protein complex function in protein transport via transport vesicles in different membrane traffic pathway.
LOC106767360	Uncharacterized pattern LOC106767360
LOC106776515	Uncharacterized pattern LOC106776515
LOC106765695	Uncharacterized pattern LOC106765695

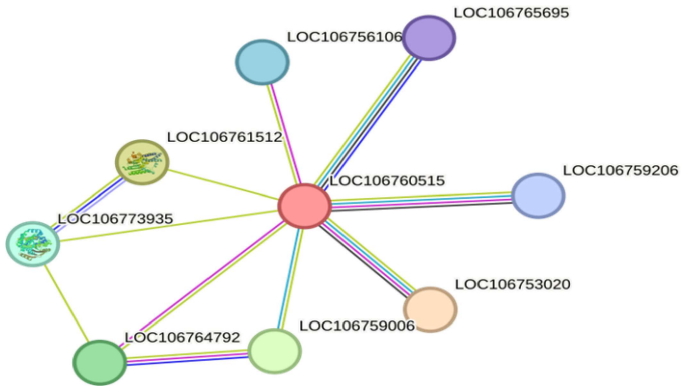


FIGURE 4.12: VrNLP2 protein-protein interaction

TABLE 4.13: Protein-protein interaction of XP_022636040.1

	Predicted functional patterns
LOC106753020	Cellulose synthase A catalytic subunit 1.
LOC106761512	Scarecrow like protein 26, belong to GRAS family.
LOC106759006	NAC domain containing protein 92.
LOC106764792	Auxin response factor as transcription factor that bind to DNA sequence
LOC106773935	Scarecrow like protein 26, belong to GRAS family.
LOC106756106	Receptor protein kinase; belong to super kinase superfamily
LOC106759206	AP-2 complex subunit alpha, subunit of adapter protein complex alpha 2
LOC106765695	Uncharacterized protein LOC106765695 isoform X1

FIGURE 4.13: VrNLP3 protein-protein interaction

TABLE 4.14: Protein-protein interaction of XP_022637353.1

	Predicted functional patterns
LOC106759206	AP-2 complex subunit alpha, subunit of adapter protein complex alpha 2

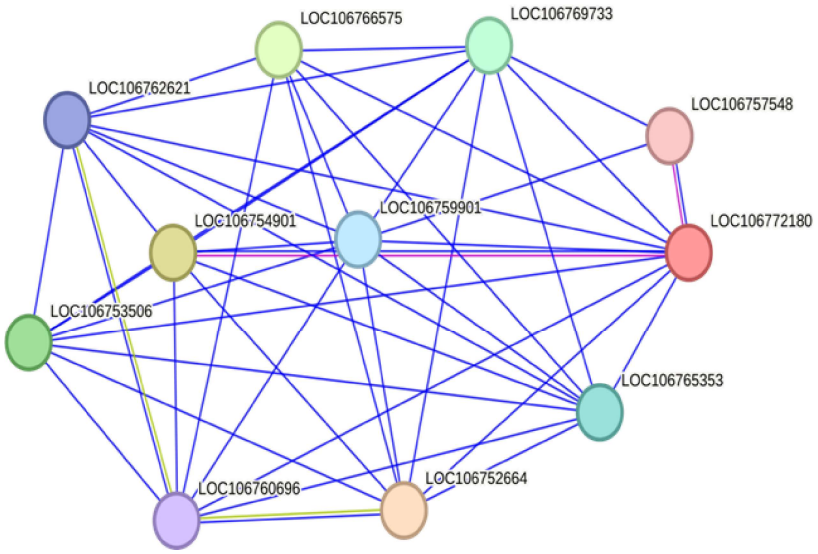


FIGURE 4.14: VrNLP4 protein-protein interaction

TABLE 4.15: Protein-protein interaction of Vradi08g18730

	Predicted functional patterns
LOC106752664	Uncharacterized protein LOC106752664
LOC106754901	F-box At3g23880
LOC106766575	F-box At3g23880 like.
LOC106753506	F-box only protein 6-like isoform X1.
LOC106769733	Uncharacterized protein LOC106769733
LOC106765353	Uncharacterized membrane protein At3g75140
LOC106759901	Myosin-2
LOC106762621	Uncharacterized membrane protein At3g4910
LOC106762621	Uncharacterized protein LOC106762621
LOC106760696	Uncharacterized membrane protein
LOC106757548	F-box At3g23880 like.

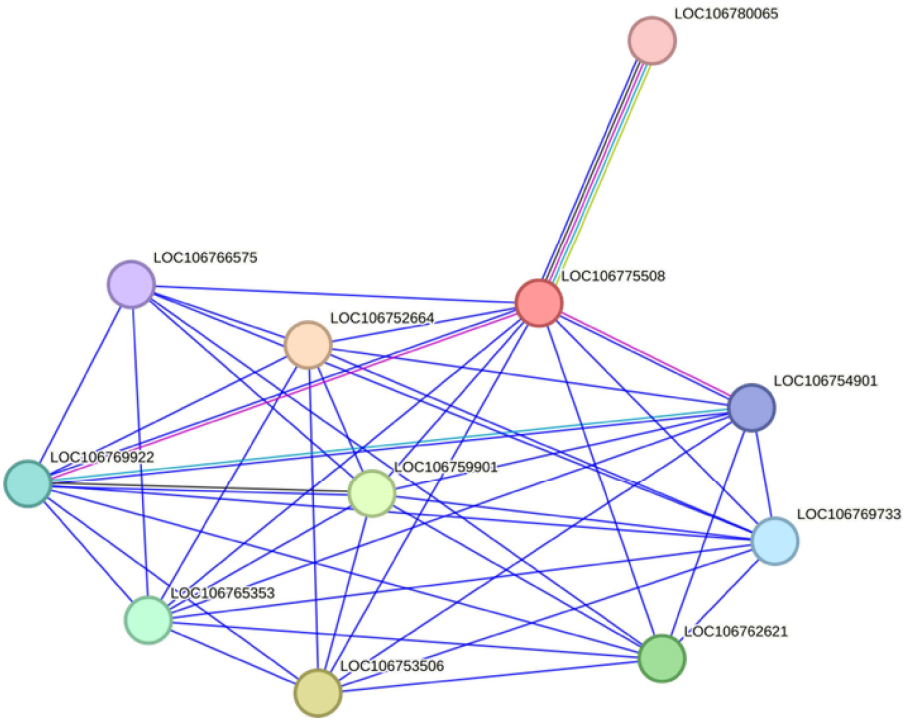


FIGURE 4.15: VrNLP5 protein-protein interaction

TABLE 4.16: Protein-protein interaction of Vradi01g14220.1

	Predicted functional patterns
LOC106752664	Uncharacterized protein LOC106752664
LOC106735306	F-box only protein 6 isoform X1
LOC106759901	Myosin
LOC106762621	Uncharacterized membrane protein At1g04910 isoform X1
LOC106765353	Uncharacterized membrane protein Atg175140
LOC106769922	Uncharacterized membrane protein LOC106769922
LOC106769733	Uncharacterized membrane protein LOC106769733
LOC106754901	F-box At3g23880
LOC106766575	F-box At3g23880
LOC106780065	Auxin related protein 2 isoform

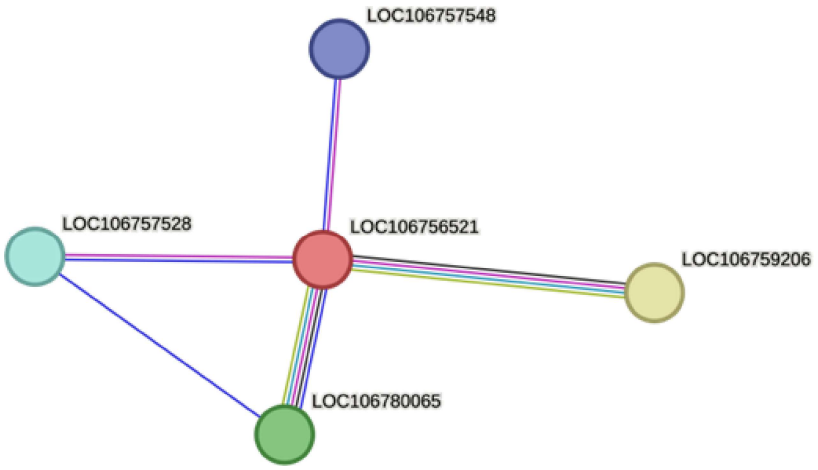


FIGURE 4.16: VrNLP6 protein-protein interaction

TABLE 4.17: Protein-protein interaction of XP_014494469.1

	Predicted functional patterns
LOC106759206	AP-2 complex subunit alpha, subunit of adapter protein complex alpha 2
LOC106780065	Auxin related protein 2 isoform X1
LOC106757528	F-box At3g23880
LOC106757548	F-box At3g23880

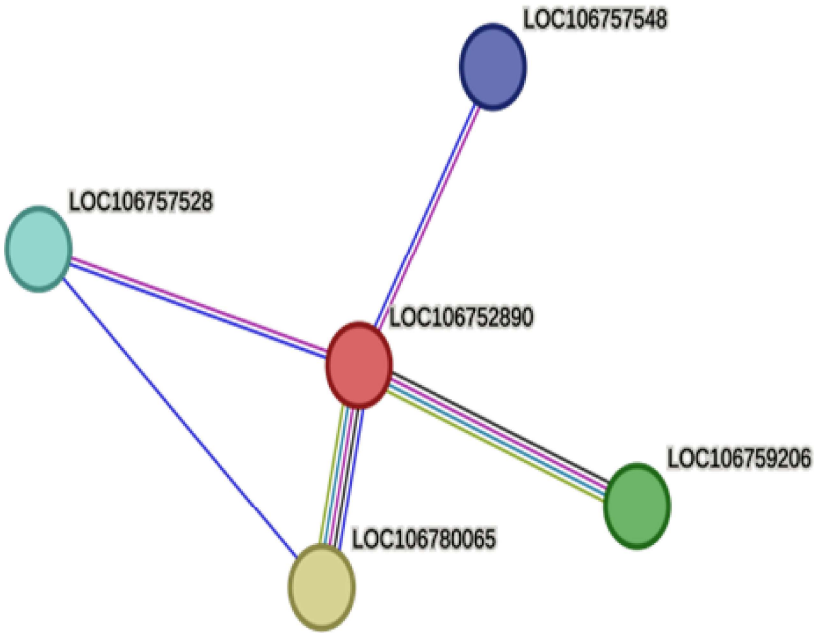


FIGURE 4.17: VrNLP7 protein-protein interaction

TABLE 4.18: Protein-protein interaction of XP_014494469.1

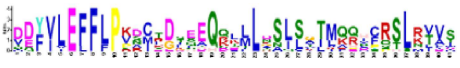
	Predicted functional patterns
LOC106759206	AP-2 complex subunit alpha, subunit of adapter protein complex alpha 2
LOC106780065	Auxin related protein 2 isoform X1
LOC106757528	F-box At3g23880
LOC106757548	F-box At3g23880

4.5 Consensus motifs composition in AtNLPs and VrNLPs

For identification of consensus motifs, AtNLPs and VrNLPs protein sequences were selected for 15 consensus motifs as shown in figure 4.18. All the sequences featured motifs that were remarkably conserved in the proteins of Arabidopsis thaliana and Vigna radiata.

TABLE 4.19: Logos of identified motifs

	Logo	E-Value	Sites	Widths
1		5.6e-463	16	36
2		1.1e-407	16	50
3		2.3e-371	14	50
4		6.10E-275	16	29
5		1.80E-249	15	41
6		4.00E-249	16	30

	Logo	E-Value	Sites	Widths
7		6.60E-225	16	41
8		1.20E-156	14	29
9		3.10E-145	9	50
10		6.60E-93	8	42
11		8.70E-110	15	27
12		1.10E-90	16	21
13		7.50E-73	16	29
14		1.60E-60	4	50
15		1.70E-50	8	21

4.6 Sequence Alignment and Phylogenetic Relationship of NLP Gene Family

Sequence of different strains of NLP genes were taken from different plants: *Arabidopsis thaliana* Maize (*Zea mays*), Rice (*Oryza sativa*), and Rapeseed (*Brassica napus*). Sequences were downloaded and then converted into FASTA file, all sequences were aligned in mega file phylogenetic tree is constructed using neighbourhood joining method.

Phylogenetic tree is constructed using CLUSTAL OMEGA (https://www.ebi.ac.uk/Tools/msa/clustalo/) and then was redesigned using itol (itol.embl.de/upload.egi). The evolutionary relationship of VrNLPs with different

species having NLP gene families were analyzed by the phylogenetic tree. The Vigna radiata are diverged in all phylogenetic tree (Fig 4.6). The NLP gene family of Vigna radiata showed evolutionary relationship with the other selected

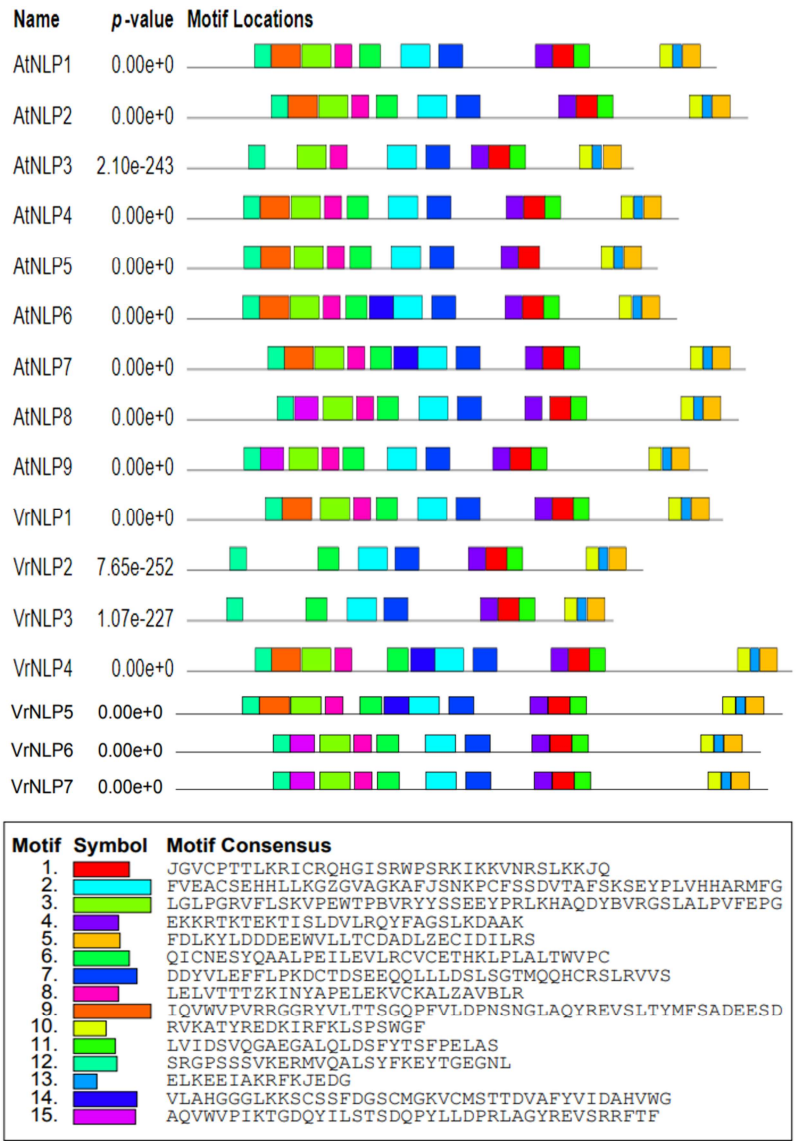


FIGURE 4.18: Consensus motifs of AtNLPs and VrNLPs

5 plants and their structural and functional are predicted to be more similar with Arabidopsis thaliana and Zea mays. According to the phylogenetic tree, the root is common for all NLP so they are in originated from common and sister evolutionary divergence of NB gene is observed. Aryan is observed within the VR NLP Jean family of Vigna Rida so all these are unique structure so their lives structural and functional similarity as the tree has been formed by neighbor joining method so there lies close similarity in VrNLP7 and VrNLP6 as compared to VrNLP7 and VrNLP4. There are two major clades and each clade is divided into two sub-clade and which in turn is divided into two sub sub clade so their lies structured and

functional similarity while Physcomitrella pattens is non-vascular so it's clade is different, while AtNLP3 has different clade too because it is basic while the rest of them are acidic in nature

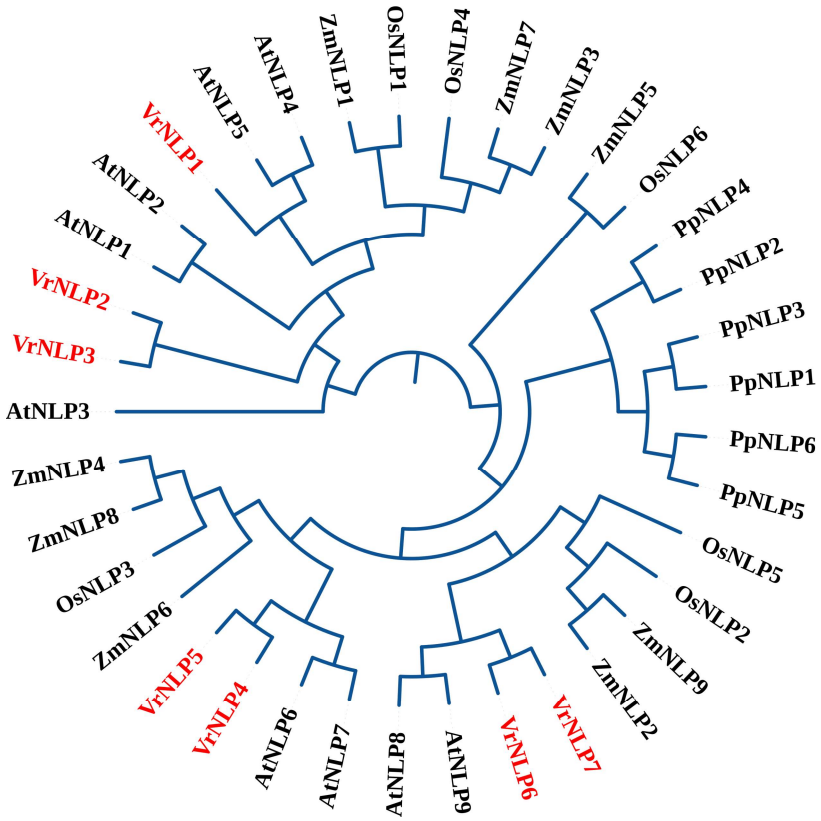


FIGURE 4.19: Phylo-genetic analysis of VrNLPs through neighbour joining method using MEGA-X

Chapter 5

Discussion

Plant growth and development are significantly influenced by transcription factors under both abiotic and biotic stresses. An important class of plant-specific transcription factors is the NODULE-INCEPTION-Like Proteins (NLPs) [29]. Previous researches have shown an established substantial role of NLPs in N uptake, assimilation, and transport controlled by N availability [1]. Although it is well known that the availability of N does not cause the expression of NLPs [7] therefore, NLPs do direct their first response to N by nuclear- retention mechanisms in order to localize NLPs. Due to the increased accumulation of NLPs proteins caused by an increased N supply, N responsive genes are better expressed, allowing plants to consume more N. NLPs are yet to be studied in horticultural plants, despite the fact that such studies have attempted to include thorough characterization of NLPs in a variety of vascular plants.

According to our research, the vascular *Vigna radiata* shows the same effect. The opportunity to study VrNLPs in the current study was made possible by the *Vigna radiata* whole-genome sequence that was published in 2018. Genome-wide studies are very effective at mining a genome database for initial identification as well as preliminary indication of structural and functional characteristics of a particular gene, even though such studies do not confirm the precise molecular mechanisms happening inside a cell. Such genome-wide studies conducted in the past have

been beneficial and have been supported by in-depth studies realized in the future [16].

In the current study, we identified three NLPs genes in the genome-databases of *Vigna radiata* and compared their characteristics to those of NLPs of A. thaliana. As in in silico studies are mostly dependent on comparison algorithms, it is possible to predict how a gene will function based on the similarities found when comparing genomic data. VrNLPs were found to have higher protein lengths, gene lengths, and molecular weights as compared to AtNLPs, but the pl and GRAVY values of both gene families were found close together, suggesting putative functional similarities between the members of the both gene families.

AtNLPs and VrNLPs were grouped into five different clades in a phylogenetic tree as a result of Tesearch on their evolutionary relationships, as shown in Figure 4.6. The *Arabidopsis thaliana* and *Zea mays* were the closest members in the clade of VrNLP genes, while all VrNLPs were clustered in a sub-Cade. This phylogenetic relationship suggests two logical interpretations. First off, the Evolutionary relationship between vascular and non-vascular plants may be the reason why all VrNLPs are placed in a different sub-cluster. Second, the existence of VrNLPs in a sister group that is closely related to NLPs from vascular plants validates the ancestry of NLPs in vascular plants

The characteristics of NLP genes and protein sequence can also be linked to the evolutionary relationship. The study of predicted proteins that interact with a gene family is another fundamental step in directing functional characterization. Further evidence of the NLP gene families' evolutionary divergence in vascular plants and bryophytes, as well as their ancestral relationship, can be found in all of the VrNLPs shared consensus protein motifs [33]. The *Arabidopsis thaliana* and *Vigna radiata* proteins both include consensus motifs that indicate their involvement in plant growth and development. According to our findings, both vascular and non-vascular NLP members exhibit larger phylogenetic divergence as well as higher ancestral linkage. The evolutionary link between the members of AtNLPs and VrNLPs is also determined by the presence of either one or both of the two protein domains (RWP-RK and PB1). Also supporting the ancestral link and

evolutionary divergence of NLP gene families in vascular plants is the fact that all VrNLPs share common protein motifs.

Vigna radiata NLP gene was discovered to be responsive to plant growth and development, which may be connected to plant growth and development influenced by N supply and control. Although it is simply suggested through in silico tools from our study that all VrNLPs are primarily involved in plant growth and development mechanisms and that responses to stress and phytohormones may be their secondary roles, this assertion can be verified through in-depth research guided by advanced molecular techniques. Another basic step in directing functional characterization is the analysis of predicted proteins that interact with a gene family. The predicted proteins that were enlisted, according to our findings, may have a conserved role in the uptake, transport, and assimilation of N. Functional characterization of NLP genes in rice revealed that they are responsive to N and are crucial in enhancing total NUE, as shown in earlier research [1].

Chapter 6

Conclusion and Future Work

In light of our findings in this study and those previously reported, it can be concluded that N availability has a major impact on both the responsiveness and regulation of VrNIPs. NLPS are a promising group of transcriptional regulators that may help increase a crop's N use efficiency (NUE). Our study simply offers a theoretical framework for the study of NLPs, highlighting the need for more in-depth research. To begin with, thorough structural and functional characterization using mutant studies can actually pinpoint their molecular characteristics. Our goal in studying NLPs in *Vigna radiata* was to fill the knowledge gap caused by a lack of relevant reports. *Vigna radiata* will be the primary focus of these investigations, particularly those related to N transport, as it shares a boundary with vascular plants and algae, making it a viable candidate for using detailed mechanisms and crucial elements of N regulation to increase crop NUE.

As, the aim of this study was to identify and analyze, both structurally and functionally, the *Vigna radiata* gene family in a vascular bryophyte in comparison to a vascular plant. For this reason.

To begin, the sequences of NLPs genes of *Arabidopsis thaliana* were retrieved from database, followed by protein-BLAST using NCBI and *Vigna radiata* selected as the organism of choice. The resultant sequences, after thorough screenings and removals, were narrowed down to finalized samples, which were then subjected to various computational tools for genome-wide analysis of NLP gene family of *Vigna*

radiata. Various analyses were performed including conserved domain identification, physicochemical characterization and localization, motif composition, phylogenetic analysis, protein-protein interaction and gene structure determination to structurally and functionally analyze the NLP gene family of *Vigna radiata*.

Following the successful identification of the NLP gene family in, the physicochemical characterization and localization, along with the observation of motifs and the interacting proteins, the NLP of *Vigna radiata* were found quite similar to those of *Arabidopsis thaliana*. The phylogenetic analysis showed evolutionary divergence and variation present between the NLP gene family of *Vigna radiata* when compared to other vascular plants. The method used was, therefore, effective and successful in providing in-depth information about the selected gene family. These genes have theoretically been proven to be a part of the NLP gene family in *Vigna radiata* and predicts their structural and functional conservation in *Vigna radiata* compared to the model plant *Arabidopsis thaliana*.

6.1 Future Recommendations

The technique utilized in this study was found fruitful and very compelling in accomplishing the point of this concentrate hence, this strategy is proposed to be utilized in the investigation of different plants in future. Other quality families might be distinguished utilizing the technique, having modern or business benefits, with low conservative expense and less time utilization. The data uncovered from this study gives a strong hypothetical foundation regarding the matter, nonetheless, the review can be continued in wet-lab for confirmation of the NLP gene family of *Vigna radiata* utilizing sub-atomic research facility methods, for example, the Polymerase Chain Response (PCR).

This study clears way for other quality families to be distinguished and modern strategies to be further developed which can, thusly, upgrade crop creation, yield and proficient utilization of nitrogen in soil. Also, hereditary designing methods can be further developed which can significantly impact the horticulture area.

These improvements might hold promising outcomes for our current circumstance, while refining different modern cycles simultaneously.

Bibliography

[1] A. Alfatiah, J. Wu, Z. Zhang, T. Xia, S.U. Jan, L-H. Yu, and C-B. Xang, "Rice NIN-LIKE PROTEIN 1 rapidly responds to nitrogen deficiency and improves yield and nitrogen use efficiency," *Journal of Experimental Botany*, vol. 71, no. 19, pp. 6032–6042, 2020.

[2] P. Govindasamy *et al.*, "Nitrogen use efficiency—a key to enhance crop productivity under a changing climate," *Frontiers of Plant Science*, vol. 14, 2023.

[3] Phoslab, *Phoslab Environmental Laboratories*, 07-May-2013. click here.

[4] J. M. Alvarez et al., *Systems approach identifies TGA1 and TGA4 transcription factors as important regulatory components of the nitrate response of Arabidopsis thaliana roots*, *Plant Journal*, vol. 80, no. 1,pp. 1–13,2014.

[5] Anas, M., and others, *Fate of nitrogen in agriculture and environment: agronomic, eco-physiological and molecular approaches to improve nitrogen use efficiency*, *Biological Research*, vol. 53, no. 1, 2020.

[6] T. Garnett, V. Conn, and B. N. Kaiser, *Root-based approaches to improving nitrogen use efficiency in plants*, *Plant Cell and Environment*, vol. 32, no. 9, pp. 1272–1283, 2009.

[7] S. Shankman, "What is nitrous oxide and why is it a climate threat?," *Inside Climate News*, Sep. 11,2019.click here.

[8] D. Hou *et al.*, "Mungbean (*Vigna radiata L.*): Bioactive polyphenols, polysaccharides, peptides, and health benefits," *Nutrients*, vol. 11, no. 6, pp. 1238, 2019.

[9] P.-W. Ong *et al.*, "Environment as a limiting factor of the historical global spread of mungbean," *Elife*, vol. 12, 2023.

[10] N. Habib, M. Z. Anwer, S. Siddiqui, S. Batool, and S. Naheed, "Trend analysis of mungbean area and yield in Pakistan," *Asian Journal of Agriculture and Rural Development*, vol. 3, no. 12, pp. 909–913, 2013.

[11] A. Ahmed, "Pakistan becomes self-sufficient in mungbean output," *Dawn*, DAWN.COM, 17-Sep-2021.

[12] D. Tang, Y. Dong, H. Ren, L. Li, and C. He, "A review of phytochemistry, metabolite changes, and medicinal uses of the common food mungbean and its sprouts (*Vigna radiata*)," *Chemistry Central Journal*, vol. 8, no. 1, 2014.

[13] H. E. Khairul Mazed, M. Yeasmin, J. F. Moonmoon, M. Moniruzzaman, and M. H. Tareque, "Seed as influenced by potassium and sulphur application," Jst.go.jp, 2016

[14] T. Huzar, "Mung beans: Health benefits, nutrition, and recipe tips," *Medicalnewstoday.com*, 11-Jan-2019. click here

[15] R. Raman, MS, and RD, "10 impressive health benefits of mung beans," *Healthline*, 15-May-2023. click here

[16] Q. M. Imran, N. Falak, A. Hussain, B.-G. Mun, and B.-W. Yun, "Abiotic stress in plants: Stress perception to molecular response and role of biotechnological tools in stress resistance," *Agronomy*, vol. 11, no. 8, p. 1579, 2021.

[17] J. Y. Ye, W. H. Tian, and C. W. Jin, "Nitrogen in plants: from nutrition to the modulation of abiotic stress adaptation," *Stress Biology*, vol. 2, no. 1, 2022.

[18] G. Yuan *et al.*, "Genome-wide identification and expression analysis of NIN-like Protein *(NLP)* genes reveals their potential roles in the response to nitrate signalling in watermelon," *Horticulture Plant Journal*, vol. 8, no. 5, pp. 602–614, 2022.

[19] A. G. Owen and D. L. Jones, "Competition for amino acids between wheat roots and rhizosphere microorganisms and the role of amino acids in plant N acquisition," *Soil Biology and Biochemistry*, vol. 33, no. 4–5, pp. 651–657, 2001.

[20] W.-H. Wang, B. Köhler, F.-Q. Cao, and L.-H. Liu, "Molecular and physiological aspects of urea transport in higher plants," *Plant Science*, vol. 175, no. 4, pp. 467–477, 2008.

[21] L. Muratore, L. Espen, and B. Prinsi, "Nitrogen uptake in plants: The plasma membrane root transport systems from a physiological and proteomic perspective," *Plants*, vol. 10, no. 4, p. 681, 2021.

[22] N. M. Crawford and A. D. M. Glass, "Molecular and physiological aspects of nitrate uptake in plants," *Trends Plant Sci.*, vol. 3, no. 10, pp. 389–395, 1998.

[23] Segonzac *et al.*, "Nitrate efflux at the root plasma membrane: Identification of an *Arabidopsis* excretion transporter," *Plant Cell*, vol. 19, no. 11, pp. 3760–3777, 2007.

[24] Z. Wu *et al.*, "Genome-wide identification, characterization, and regulation of RWP-RK gene family in the nitrogen-fixing clade," *Plants*, vol. 9, no. 9, p. 1178, 2020.

[25] S. U. Jan *et al.*, "Genome-wide study and expression analysis of Nodule-inception-like protein *(NLP)* gene family in Physcomitrella patens reveal its role in nitrogen response," *Pak. J. Bot.*, vol. 55, no. 1, 2023.

[26] "Regulation of Transcription and Gene Expression," *Nature.com.*click here.

[27] L. Zhang, W. Fang, F. Chen, and A. Song, "The role of transcription factors in the regulation of plant shoot branching," *Plants*, vol. 11, no. 15, p. 1997, 2022.

[28] M. S. Chattha *et al.*, "Enhancement of nitrogen use efficiency through agronomic and molecular-based approaches in cotton," *Frontier Plant Science.*, vol. 13, 2022.

[29] M. Liu *et al.*, "Genome-wide identification and characterization of NODULE-INCEPTION-like protein *(NLP)* family genes in *Brassica napus*," *Int. J. Mol. Sci.*, vol. 19, no. 8, p. 2270, 2018.

[30] L. Liu, M. J. White, and T. H. MacRae, "Transcription factors and their genes in higher plants. Functional domains, evolution and regulation," *European Journal of Biochemistry.*, vol. 262, no. 2, pp. 247–257, 1999.

[31] L. Zhang, W. Fang, F. Chen, and A. Song, "The role of transcription factors in the regulation of plant shoot branching," *Plants*, vol. 11, no. 15, p. 1997, 2022.

[32] L. Yuan *et al.*, "The organization of high-affinity ammonium uptake in *Arabidopsis* roots depends on the spatial arrangement and biochemical properties of AMT1-type transporters," *Plant Cell*, vol. 19, no. 8, pp. 2636–2652, 2007.

[33] B. Jagadhesan *et al.*, "Genome-wide analysis of *NLP* transcription factors reveals their role in nitrogen stress tolerance of rice," *Scientific Report.*, vol. 10, no. 1, pp. 1–16, 2020.

[34] M. Ge *et al.*, "Genome-wide analysis of maize *NLP* transcription factor family revealed the roles in nitrogen response," *Plant Growth Regulation.*, vol. 84, no. 1, pp. 95–105, 2021.