CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Comparative Genomic Analysis to Explore Key Genetic Factors Associated with Probiotic Capabilities of *Akkermansia muciniphila*

by

Shabeen Fatima

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

2020

I dedicate this thesis to my parents and my teachers.

CERTIFICATE OF APPROVAL

Comparative Genomic Analysis to Explore Key Genetic Factors Associated with Probiotic Capabilities of *Akkermansia muciniphila*

by

Shabeen Fatima

(MBS183015)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. S. Babar Jamal Bacha	NUMS, Islamabad
(b)	Internal Examiner	Dr. Erum Dilshad	CUST, Islamabad
(c)	Supervisor	Dr. Syeda Marriam Bakhtiar	CUST, Islamabad

Dr. Syeda Marriam Bakhtiar
Thesis Supervisor
December, 2020

Dr. Sahar Fazal
Head
Dept. of Biosciences & Bioinformatics
December, 2020

Dr. Muhammad Abdul Qadir
Dean
Faculty of Health & Life Sciences
December, 2020

Author's Declaration

I, **Shabeen Fatima** hereby state that my MS thesis titled "**Comparative Genomic Analysis to Explore Key Genetic Factors Associated with Probiotic Capabilities of *Akkermansia muciniphila*** " 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.

(Shabeen Fatima)

Registration No: MBS183015

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Comparative Genomic Analysis to Explore Key Genetic Factors Associated with Probiotic Capabilities of *Akkermansia muciniphila***" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been dully 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.

(Shabeen Fatima)

Registration No: MBS183015

Acknowledgements

In the name of **Allah**, the Most Gracious and the Most Merciful Alhamdulillah, all praises to Allah for giving me strength and for His blessings in completing my MS thesis. First, I would like to express my gratitude to Capital University of Science and Technology (CUST) Islamabad my Alma matter for providing me an opportunity to do MS Biosciences and achieving my goal to pursue higher studies. I would like to start with a special appreciation that goes to my Supervisor, **Dr. Syeda Marriam Bakhtiar**, Assistant Professor, CUST for her support, encouragement and guidance throughout this Pandemic Situation. She helped me a lot and answered my all queries on time. I want to extend special thanks to her. I would like to acknowledge the contributions of all my teachers Dr.Shaukat Iqbal, Dr.Erum Dilshad, Dr. Arshia Amin and our beloved HoD Dr. Sahar Fazal. Finally, I express my gratitude to my parents and siblings for providing me support and help me throughout my years of study and through the process of researching and writing this thesis.

(Shabeen Fatima)

Registration No: MBS183015

Abstract

In contrast to antibacterial drugs, probiotics are gaining interest as an alternative to treat and control digestive malfunctions including functional gastrointestinal disorders. Probiotics comes with a property to not only support a gut barrier but also enhancing health by supporting immune system. This ability of probiotics in supporting and enhancing the activities of immune system have been also utilized to control inflammatory diseases *Akkermansia muciniphila* is a species of bacteria that helps to maintain our gut lining and possess many health benefits. *Akkermansia muciniphila* (*A. muciniphila*), an intestinal symbiont colonizing in the mucosal layer, is considered to be a promising candidate as probiotics. *Akkermansia muciniphila* is known to have an important value in improving the host metabolic functions and immune responses. Moreover, *Akkermansia muciniphila* may have a value in modifying cancer treatment. However, most of the current researches focus on the correlation between *Akkermansia muciniphila* and diseases, and little is known about the causal relationship between them. This study was designed to analyse genomic features of *Akkermansia muciniphila* so analyse its safety to be used as probiotic and also to evaluate its probiotic potentials. Pangenome analysis COG and phylogenetic analysis revealed that *Akkermansia muciniphila*shows a stable genome character. The antibiotic resistance pattern was analysed and only intrinsic resistant genes necessary of propbiotics were present and no multidrug resistance was found. It was also found that no pathogenicity islands or virulent genes are present in any of the selected strains. Hence, *Akkermansia muciniphila* could be considered safe to be used as probiotic, for further validation, genomic islands of each strain were separately analysed. Bacteriocin producing genes of each strain were also analysed proposing the conclusion that *Akkermansia muciniphila*is safe and has potential to be used as probiotic against inflammatory diseases especially obesity.

Keywords: *A. muciniphila*, probiotics, metabolic disorders

Contents

Author's Declaration iv

Plagiarism Undertaking v

Acknowledgements vi

Abstract vii

List of Figures xi

List of Tables xiv

Abbreviations xvi

1 Introduction 1
- 1.1 Background . . . 1
- 1.2 Characteristics of Probiotics . . . 2
- 1.3 *Akkermansia muciniphila* as Probiotics . . . 3
- 1.4 Aim and Objectives . . . 7

2 Literature Review 9
- 2.1 Probiotics and their Role in Gut Health . . . 9
- 2.2 Significance of Probiotics . . . 10
- 2.3 Mechanisms of Probiotic Action . . . 11
- 2.4 Probiotics as Barriers . . . 13
- 2.5 Production of Antimicrobial Substances . . . 13
- 2.6 Competition for an Attachment to Intestinal Cells . . . 14
- 2.7 Immune Modulation . . . 15
- 2.8 Intervention with Quorum Sensing Signaling . . . 15
- 2.9 Role of Probiotics Against Gastrointestinal Diseases . . . 16
  - 2.9.1 Antibiotics-Associated Diarrhea . . . 18
  - 2.9.2 Infections of *Clostridium difficile* . . . 18
  - 2.9.3 *Helicobacter Pylori* infections . . . 20
  - 2.9.4 Irritable Bowel Syndrome (IBS) . . . 20
  - 2.9.5 Ventilator-Associated Pneumonia . . . 21

2.9.6 Allergy and Enhancement of Immune Response 22
2.10 Safety Concerns with Use of Bacteria as Probiotics 22
2.11 Acid and Bile Tolerance 23
2.12 Competitive Exclusion of Pathogens 24
2.13 Adhesion 24
2.14 Bacteriocin Production 25
2.15 Immuno Modulation 25
2.16 *Akkermansia muciniphila* and its Potential as Probiotic 26

3 Material and Methods **28**
3.1 Genomic Data Collection 29
3.1.1 Selection of Strains 29
3.1.2 Retreival of Bacterial Sequences 29
3.1.3 Selection of Reference Strain 30
3.1.4 Quality Assessment 30
3.2 Bacterial Pan-Genome Analysis 30
3.2.1 Comparative Analysis of Orthologous Genes 31
3.2.2 Phylogenetic Analysis 31
3.3 Antibiotic Resistance (Resistome) Determinants 32
3.4 Virulence Factors 32
3.4.1 Genomic Islands Determination 33

4 Result and Analysis **34**
4.1 Genomic Data Collection 34
4.1.1 Selection Of Strains 34
4.1.2 Selection of Reference Genome 37
4.1.3 Gene Prediction and Annotation 38
4.2 Pan Genome Analysis 40
4.2.1 Exclusive Gene Family Analysis 45
4.2.2 Sequence Extraction 46
4.2.3 Phylogenetic Analysis 46
4.2.4 Clusters of Orthologous Genes 47
4.2.5 Determination of Genetic Islands 49
4.2.6 *Akkermansia muciniphila* ATCC BAA-835 52
4.2.7 *Akkermansia muciniphila* CBA5201 54
4.2.8 *Akkermansia muciniphila* DSM 22959 58
4.2.9 *Akkermansia muciniphila* JCM 30893 61
4.2.10 *Akkermansia muciniphila* AMDK-7 63
4.2.11 *Akkermansia muciniphila* AMDK-8 66
4.2.12 *Akkermansia muciniphila* AMDK-10 68
4.2.13 *Akkermansia muciniphila* AMDK-11 72
4.2.14 *Akkermansia muciniphila* AMDK-12 75
4.2.15 *Akkermansia muciniphila* AMDK-13 78

4.2.16 *Akkermansia muciniphila* AMDK-14 81
4.2.17 *Akkermansia muciniphila* AMDK-15 84
4.2.18 *Akkermansia muciniphila* AMDK-16 87
4.2.19 *Akkermansia muciniphila*AMDK-17 89
4.2.20 *Akkermansia muciniphila* AMDK-18 91
4.2.21 *Akkermansia muciniphila* AMDK-19 93
4.2.22 *Akkermansia muciniphila* AMDK-20 95
4.2.23 *Akkermansia muciniphila* AMDK-21 98
4.2.24 *Akkermansia muciniphila* AMDK-22 100

5 Conclusions and Recommendations **104**

Bibliography **106**

List of Figures

1.1 Mechanism of Action of Probiotics in Gut where (a) Adhesion of Microorganisms and then their Colonization (b) Indicates Activation and Enhancement of Immune system (c) Creat an Epithelial Barrier (d) Competition with Pathogens (e) Bacteriocin Production 4

1.2 *Akkermansia muciniphila* ATCC BAA-835 Micrograph taken by the Scanning Electron Microscope [1]. 5

1.3 Mechanisms of Action, by which *Akkermansia muciniphila* Helps in Health Maintenance[2]. 6

2.1 Role of Probiotics in Human Health [13]. 11

3.1 Summarizes the Methodological Steps used to Analyze Probiotic Potentials of *Akkermansia muciniphila*. 28

4.1 Circular Genomic View of *Akkermansia muciniphila* Strain ATCC BAA-835 along with Major Genomic Regions. 37

4.2 Gene Accumulation Curve of All Selected Strains, the Curve in Purple Indicates Core Genome and the Trend Indicates that the Number of Core Genes Decrease with Addition of New Strain. The curve in Orange Color Indicates that Pangenome which Show the Trend of Increase in Pangenome with Addition of Each Strain. . . . 42

4.3 The Pan Genome Profile Trends Obtained using BPGA, Pangenome is Indicated by Cyan Color and Show an Increasing Trend, While Pink Color Show Core Genome Depicted a Decrease with Addition of Each Strain. 42

4.4 Information About Core Accessory Unique Genes from Roary . . . 45

4.5 Number of New Genes or Unique Genes Added to Pangenome with Addition of Each New Strain in Pan Genomic Analysis 46

4.6 Phylogenetic Tree of All Selected Strains 47

4.7 Clusters of Orthologous Groups (COG) and Distribution of the Core Genes, Accessory Genes and Unique Genes in *Akkermansia muciniphila* Strains. 47

4.8 Distribution of the Core genes, Accessory Genes and Unique Genes Involve in Different Processes in *Akkermansia muciniphila* Strains. . 48

4.9 Genetic Islands in *Akkermansia muciniphila* ATCC BAA-835 as Predicted by Island Viewer. 52

4.10 Bacteriocin of *Akkermansia muciniphila* ATCC- BAA 835 from BAGEL4 . . . 54
4.11 Genetic Islands in *Akkermansia muciniphila* CBA5201 as Predicted by Island Viewer. . . . 55
4.12 Bacteriocin of *Akkermansia muciniphila* CBA5201 from BAGEL4 . 58
4.13 Genetic Islands in *Akkermansia muciniphila* DSM 22959 as Predicted by Island Viewer. . . . 59
4.14 Bacteriocin of *Akkermansia muciniphila* DSM22959 from BAGEL . 61
4.15 Genetic islands in *Akkermansia muciniphila*JCM 30893 as Predicted by Island Viewer. . . . 62
4.16 Bacteriocin of *Akkermansia muciniphila* JCM 30893 from BAGEL4 63
4.17 Bacteriocin of *Akkermansia muciniphila* AMBK-7 from BAGEL4 . 63
4.18 Genetic Islands in *Akkermansia muciniphila* AMDK-7 as Predicted by Island Viewer. . . . 64
4.19 Genetic Islands in *Akkermansia muciniphila* AMDK-8.as Predicted by Island Viewer. . . . 68
4.20 Bacteriocin of *Akkermansia muciniphila* AMBK-8 from BAGEL4 . 68
4.21 Genetic Islands in *Akkermansia muciniphila*AMDK-10 as Predicted by Island Viewer. . . . 69
4.22 BAGEL Results . . . 72
4.23 Genetic Islands in *Akkermansia muciniphila*AMDK-11 as Predicted by Island Viewer. . . . 73
4.24 Bacteriocin of *Akkermansia muciniphila* AMBK-11 from BAGEL4 . 75
4.25 Genetic Islands in *Akkermansia muciniphila*AMDK-12 as Predicted by Island Viewer. . . . 76
4.26 Bacteriocin of *Akkermansia muciniphila* AMBK-12 from BAGEL4 . 78
4.27 Genetic Islands in *Akkermansia muciniphila* AMDK-13 as Predicted by Island Viewer. . . . 79
4.28 Bacteriocin of *Akkermansia muciniphila* AMBK-13 from BAGEL4 . 82
4.29 Genetic Islands in *Akkermansia muciniphila* AMDK-14 as Predicted by Island Viewer. . . . 82
4.30 Bacteriocin of *Akkermansia muciniphila* AMBK-14 from BAGEL4 . 84
4.31 Genetic Islands in *Akkermansia muciniphila* AMDK-15 as Predicted by Island Viewer. . . . 84
4.32 Bacteriocin of *Akkermansia muciniphila* AMBK-15 from BAGEL4. 87
4.33 Genetic Islands in *Akkermansia muciniphila* AMDK-16 as Predicted by Island Viewer. . . . 89
4.34 Bacteriocin of *Akkermansia muciniphila* AMBK-16 from BAGEL4 . 89
4.35 Genetic Islands in *Akkermansia muciniphila* AMDK-17 as Predicted by Island Viewer. . . . 90
4.36 Bacteriocin of *Akkermansia muciniphila* AMBK-17 from BAGEL4 . 91
4.37 Genetic Islands in *Akkermansia muciniphila* AMDK-18 as Predicted by Island Viewer. . . . 92
4.38 Bacteriocin of *Akkermansia muciniphila* AMBK-18 from BAGEL4 . 93

4.39 Genetic Islands in *Akkermansia muciniphila* AMDK-19 as Predicted by Island Viewer. 95
4.40 Bacteriocin of *Akkermansia muciniphila* AMBK-19 from BAGEL4 . 95
4.41 Genetic Islands in *Akkermansia muciniphila* AMDK-20. as predicted by Island Viewer. 96
4.42 Bacteriocin of *Akkermansia muciniphila* AMBK-20 from BAGEL4 . 98
4.43 Genetic islands in *Akkermansia muciniphila* AMDK-21 as predicted by Island Viewer. 98
4.44 Genetic Islands in *Akkermansia muciniphila* ATCC BAA-835 as Predicted by Island Viewer. 100
4.45 Genetic Islands in *Akkermania muciniphila* AMDK-22 as Predicted by Island Viewer. 101
4.46 Bacteriocin of *Akkermansia muciniphila* AMBK-22 from BAGEL4 . 103

List of Tables

1.1 Characteristics of Probiotics [5]. 2

2.1 Mechanisms of Probiotic Actions [15]. 12
2.1 Mechanisms of Probiotic Actions [15]. 13
2.2 List of Different Strains of Probiotics Against Gastrointestinal Infectious Diseases [36]. 16
2.2 List of Different Strains of Probiotics Against Gastrointestinal Infectious Diseases [36]. 17
2.3 Correlation between *Akkermansia muciniphila* and Disease in Humans [78]. 26
2.3 Correlation between *Akkermansia muciniphila* and Disease in Humans [78]. 27

4.1 List of Selected Strains After Literature Review, Strains were Selected Based on Availability of Complete Genome and Information Related to Source of Isolation. 35
4.2 Summarizes the Findings of Genome Annotation Using RAST Server 38
4.3 Information about Core, Variable and Unique genes. 43
4.4 Results from Roary for Pangenome Analysis 45
4.5 Antibiotic Resistance Genes through CARD Analysis 49
4.6 Islands of *Akkermansia muciniphila* ATCC-BAA 835 Strain 53
4.6 Islands of *Akkermansia muciniphila* ATCC-BAA 835 Strain 54
4.7 Islands of *Akkermansia muciniphila* CBA5201 Strain 55
4.7 Islands of *Akkermansia muciniphila* CBA5201 Strain 56
4.7 Islands of *Akkermansia muciniphila* CBA5201 Strain 57
4.7 Islands of *Akkermansia muciniphila* CBA5201 Strain 58
4.8 Islands of *Akkermansia muciniphila* DSM22959 Strain 59
4.8 Islands of *Akkermansia muciniphila* DSM22959 Strain 60
4.8 Islands of *Akkermansia muciniphila* DSM22959 Strain 61
4.9 Islands of *Akkermansia muciniphila* JCM30893 Strain 62
4.9 Islands of *Akkermansia muciniphila* JCM30893 Strain 63
4.10 Islands of *Akkermansia muciniphila* AMDK-7 Strain 64
4.10 Islands of *Akkermansia muciniphila* AMDK-7 Strain 65
4.10 Islands of *Akkermansia muciniphila* AMDK-7 Strain 66
4.11 Islands of *Akkermansia muciniphila* AMDK-8 Strain 66
4.11 Islands of *Akkermansia muciniphila* AMDK-8 Strain 67

4.12 Islands of *Akkermansia muciniphila* AMDK-10 Strain 69
4.12 Islands of *Akkermansia muciniphila* AMDK-10 Strain 70
4.12 Islands of *Akkermansia muciniphila* AMDK-10 Strain 71
4.12 Islands of *Akkermansia muciniphila* AMDK-10 Strain 72
4.13 Islands of *Akkermansia muciniphila* AMDK-11 Strain 73
4.13 Islands of *Akkermansia muciniphila* AMDK-11 Strain 74
4.13 Islands of *Akkermansia muciniphila* AMDK-11 Strain 75
4.14 Islands of *Akkermansia muciniphila* AMDK-12 Strain 76
4.14 Islands of *Akkermansia muciniphila* AMDK-12 Strain 77
4.14 Islands of *Akkermansia muciniphila* AMDK-12 Strain 78
4.15 Genetic islands in *Akkermansia muciniphila* AMDK-13 79
4.15 Genetic islands in *Akkermansia muciniphila* AMDK-13 80
4.15 Genetic islands in *Akkermansia muciniphila* AMDK-13 81
4.16 Islands of *Akkermansia muciniphila* AMDK-14 Strain 83
4.17 Islands of *Akkermansia muciniphila* AMDK-15 Strain 85
4.17 Islands of *Akkermansia muciniphila* AMDK-15 Strain 86
4.18 Islands of *Akkermansia muciniphila* AMDK-16 Strain 87
4.18 Islands of *Akkermansia muciniphila* AMDK-16 Strain 88
4.19 Genetic Islands in *Akkermansia muciniphila* AMDK-17 90
4.19 Genetic Islands in *Akkermansia muciniphila* AMDK-17 91
4.20 Islands of *Akkermansia muciniphila* AMDK-18 Strain 92
4.21 Islands of *Akkermansia muciniphila* AMDK-19 Strain 93
4.21 Islands of *Akkermansia muciniphila* AMDK-19 Strain 94
4.22 Islands of *Akkermansia muciniphila* AMDK-20 Strain 96
4.22 Islands of *Akkermansia muciniphila* AMDK-20 Strain 97
4.23 Genetic islands in *Akkermansia muciniphila* AMDK-21 99
4.23 Genetic islands in *Akkermansia muciniphila* AMDK-21 100
4.24 Islands of *Akkermansia muciniphila* AMDK-22 Strain 101
4.24 Islands of *Akkermansia muciniphila* AMDK-22 Strain 102
4.24 Islands of *Akkermansia muciniphila* AMDK-22 Strain 103

Abbreviations

BPGA	Ultra-Fast Bacterial Pan-Genome Analysis Pipeline
CARD	Comprehensive Antibiotic Resistance Database
CDS	Coding DNA Sequence
GI	Genomic Island
HGT	Horizontal Gene Transfer
MGEs	Mobile Genetic Elements
VFDB	Virulence Factor of Bacterial Pathogens Database

Chapter 1

Introduction

1.1 Background

Probiotics are live nonpathogenic microorganisms widely used in pharmaceutical, medicinal and food industries. In past few years focus has been shifted towards not only the characteristics of well-defined and newly discovered probiotics but also on their capabilities. Generally, probiotics are desired to have resistance to acid and bile salts in order to avoid dysbiosis and inflammatory responses [1].

In contrast to antibacterial drugs, probiotics are gaining interest as an alternative to treat and control digestive malfunctions including functional gastrointestinal disorders. Probiotics comes with a property to not only support a gut barrier but also enhancing health by supporting immune system [8]. This ability of probiotics in supporting and enhancing the activities of immune system have been also utilized to control inflammatory diseases such as rheumatoid arthritis [1] ,type 1 diabetes [4], multiple sclerosis [4], atopic dermatitis [4], and myasthenia gravis [4].

Not only against infections and inflammations, probiotics have also been reported to have a significant role in treating cancers, neurodegenerative diseases, metabolic syndrome and psychiatric illnesses, as well as for the patients who are on mechanical ventilators in intensive care units [8]. Despite all these promising application

of probiotics in control and treatment of diseases, the major challenge still remains the selection of suitable probiotic strain [3].

1.2 Characteristics of Probiotics

Probiotic bacteria are simply defined in terms of live cultures or living bacterial species which can help in health maintenance of digestive tract i. e. they have capability to maintain balance in the gut microbiota which could be disturbed as an outcome of infection or use of antibiotics [5]. This basic definition helps us to understand why the features or parameters on which a probiotic is analyzed are usually focused on their capability to survive in gastrointestinal tract involving tolerance to acids and bile salts [6], microcin and hydrogen peroxide production for competitive advantage and antimicrobial ability [7], and impact or stimulation of immune system [8]. All the researches involving classical research or structural and functional genomics mostly focus on these parameters. Table 1.1 summarizes characteristic properties of probiotics.

Table 1.1: Characteristics of Probiotics [5].

S.#	**Characteristics**
1	They can show maximum viability in the digestive system
2	They don't pretend the toxicity, as well as pathogenicity.
3	They should be capable to colonize in the intestinal epithelial cells.
4	They can consume the maximum nutrients and substrate in normal diets

Based on these parameters few bacterial genera and species are considered at higher rank with respect to probiotic capabilities. These species include *Lacticasei bacilluscasei, Lactobacillus delbruekii, Lactobacillus acidophilus, Lactiplantibacillus plantarum, Limosilactobacillus fermentum, Limosilactobacillusreuteri, Bifidobacterium breve, Bifidobacterium longum, Bifidobacterium lactis, Propionibacterium freudenreichii, Bacillus subtilis, Bacillus cereus, E.coli* and *Entrecoccus*

faecium. All of these species have therapeutic application in prevention and treatment of intestinal disorders, such as diarrhea in newborns [9]. Most of the studies in recent past have been focused on Lactic Acid bacteria and their potential as probiotics among which most of the Lactic acid bacteria under focus were of gut origin.

Whatever is the source of or type of probiotic strain, when it comes to its applications and introduction to living hosts various factors including formulation and dose are also considered so that the strain can impart the desired property and activate immune system [9]. Figure 1.1 summarizes various mechanisms by which probiotics help in health maintenance and enhancement. Therefore, the research focus has now diverted to understand the host bacteria interactions, activation of immune system and many more using the state of art technologies of genomics, proteomics, interactomes and transcriptomics. With advent of Genetic engineering the focus was diverted from probiotic strains to probiotic genes, genomic analysis is performed to identify genes responsible for probiotic features. Transcriptomics and proteomics studies have been utilized to identify differentially expressed genes to differentiate or categories probiotic and non-probiotic strains of same bacterial species. Networks and pathways along with protein-protein interactions further elucidated the mechanisms of probiotic action.

1.3 *Akkermansia muciniphila* as Probiotics

Although many disease conditions are reported to be improved with probiotics use, but yet major source of probiotic strains still remains gut microbiota. Similarly, major applications of probiotics ate also against metabolic dysfunctions and gastrointestinal tract. Inflammatory and metabolic disease which is getting lots of interest is obesity, and potential use of probiotics to reduce body weight is talk of the town these days. *Akkermansia muciniphila*, discovered at Wageningen university of the Netherlands in 2004 in search of mucin-degrading bacteria from human fecal matter [10]. This bacterium is gram negative oval shaped (Figure 1.2),

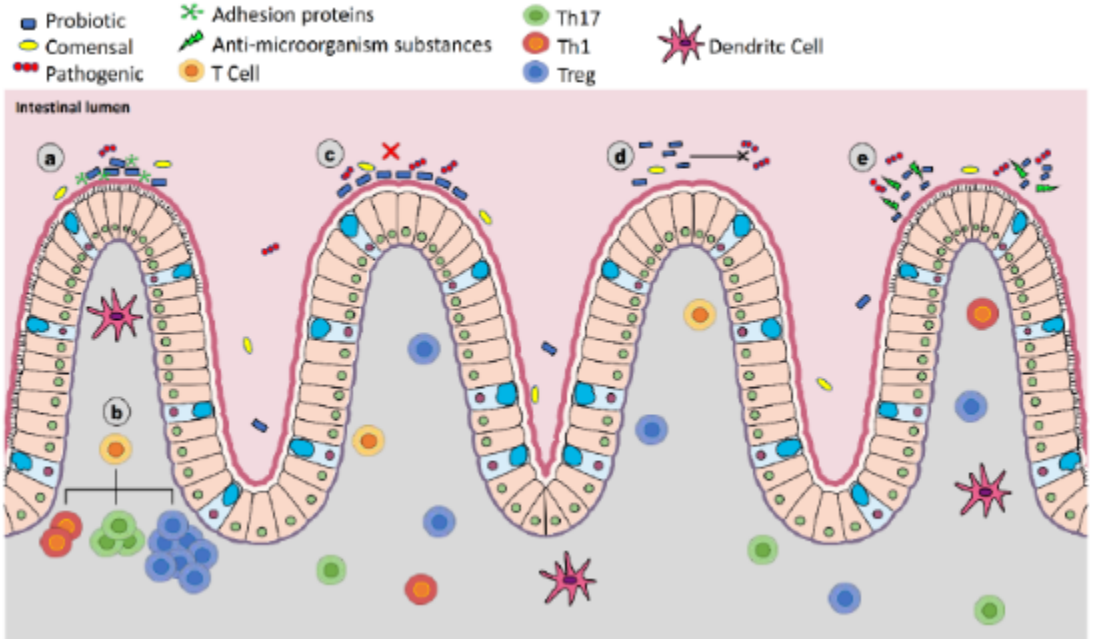

FIGURE 1.1: Mechanism of Action of Probiotics in Gut where (a) Adhesion of Microorganisms and then their Colonization (b) Indicates Activation and Enhancement of Immune system (c) Creat an Epithelial Barrier (d) Competition with Pathogens (e) Bacteriocin Production

non-motile, non-endospore forming, strictly anaerobic (few studies report that it can tolerate low levels of oxygen) and widely distributed in among intestinal microflora of various animals including humans [10]. In humans it is more abundant in intestinal mucosal layer of caecum of both healthy adults and infants [11].

Akkermansia muciniphila, is among the most frequent species among the meta genomes of healthy gut, rather is considered in amongst top 20 of commonly reported species. It is reported to colonize healthy gut during first year after birth and gradually reaches the level of healthy adults but later the number reduces with age. Introduction of this bacterial species to gut is due to its presence in human milk, therefore milk carries *Akkermansia muciniphila* from mothers to the feeding infants. Its presence in newborn gut is evident of this transmission. At this stage the acid tolerance capabilities along with the ability to utilize milk polysaccharides enables this species helps it to colonize the gastrointestinal tract [11].

Despite thepresence of few probiotic capabilities and reports of involvement of *Akkermansia muciniphila* in disease control , this bacterium is still not efficiently

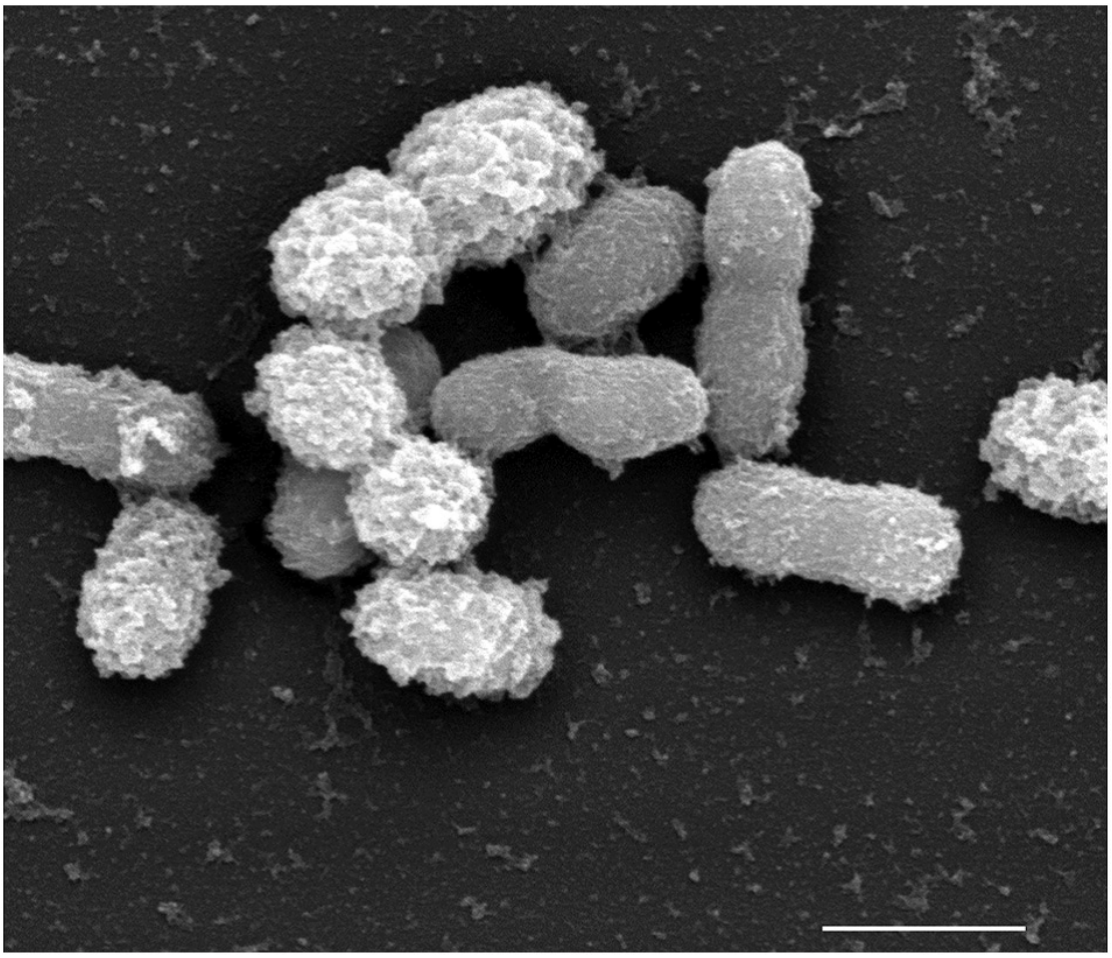

FIGURE 1.2: *Akkermansia muciniphila* ATCC BAA-835 Micrograph taken by the Scanning Electron Microscope [1].

used as probiotic strain.*Akkermansia muciniphila*, is been focus mainly to understand the mechanism by which this specie is related to disease (Figure 1.3). In most cases it is not considered as cause of diseases, but its contribution in onset of various diseases is debatable and marks questions on its safety for use in humans. This is the major reason for which this bacterium has not widely been used in foods and medicines but there have been suggestive evidences that this bacterium could be used safely in humans [12].

Although *Akkermansia muciniphila* is not pathogenic and never reported to be a primary cause of any diseases , but its property of adhesion is always questioned. It also possesses the capability to adhere with intestinal mucosal layer and degrade it, which further enhance the concerns regarding its safety. Although it is well reported that contrary to pathogenic bacteria *Akkermansia muciniphila* only

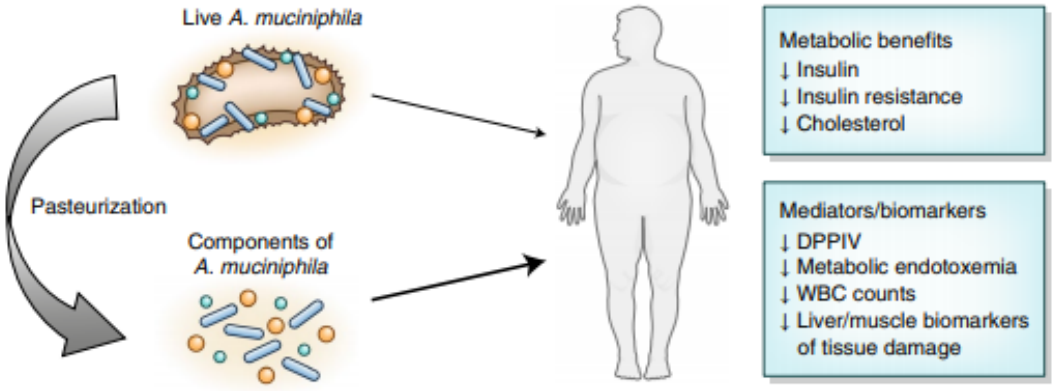

FIGURE 1.3: Mechanisms of Action, by which *Akkermansia muciniphila* Helps in Health Maintenance[2].

adhere and degrade outer mucosal area and never reach inner layers, as reaching inner layers require more pathogenic genes. In addition to this, degradation of mucosal layers is part of intestinal self-renewal balance and is a normal process [12]. Although Gram negative bacteria with lipopolysaccharides, *Akkermansia muciniphila* is not reported to be associated with endotoxemia rather presence of this bacteria is found to be associated with reduction of endotoxins level in mice with high fat diet. Like all other mucin degrading bacteria, *Akkermansia muciniphila* is also with the capability to regulate host immune system through various cytokines including necrosis factors such as TNF -alpha, INF-alpha, and interleukins such as IL-10 and IL-4 [10].

Similarly adherence ability of *Akkermansia muciniphila* with mucosal layer is an important characteristic of a potential probiotic. Intestinal mucosal linings are there to prevent microbial/pathogenic attacks on epithelial layer and this mucosa also provide nutrition to the adhered microbes. Microbes attached to the intestinal mucosa provide competition to potential pathogens and do not allow them to attack intestinal epithelium. *Akkermansia muciniphila* is reported to be a typical representative of this competition [11]. The frequency and distribution of *Akkermansia muciniphila* varies in different regions of intestine depending upon the nutrient availability.

As discussed earlier that *Akkermansia muciniphila* has not been reported to be

cause of any diseases but found to be associated with various diseases. This association is two facet, one the distribution and frequency of *Akkermansia muciniphila* increase in diseases condition, but on the other hand it is well reported that decrease in *Akkermansia muciniphila* number is strongly associated with metabolic diseases such as obesity, diabetes, hypertension, inflammatory bowel disease, autism and atopy [12]. The association is well evident from observational and animal model studies. Although the mechanisms by which this association works is still unclear but yet we can conclude that *Akkermansia muciniphila* is a key player in maintaining homeostasis and healthy physiology of human gut.

Obesity is emerged as a major threat to human health and focus has been to find effective remedies against its onset, control and weight reduction. As it is an aesthetic issue as well therefore a lot of investment is done on finding solutions for effective weight loss. Interestingly, *Akkermansia muciniphila* has found to be effective against obesity prevention. Similarly, decrease in number of *Akkermansia muciniphila* in children suffering with IgE-related atopic diseases, suggest an important role of this bacteria in immune modulation [13].

Akkermansia muciniphila, has a great potential as probiotic that can make good use of gastrointestinal mucin, but its safety is debated. Various studies are evident of the safety and suggest oral administration of *Akkermansia muciniphila* but more human trials are required. Computational biology and bioinformatics tools including comparative genomins and pangenome analysis can provide significant insight into genome of *Akkermansia muciniphila* and help us to understand population structure as well evolutionary history [14].

1.4 Aim and Objectives

Obesity and metabolic diseases have increases enormously in past few years either due to change in lifestyle or food intake. Lots of weight loss remedies are in use and probiotics are one of them. There is a requirement to identify probiotic strains that could help in weight loss or can prevent obesity. One of the bacterial specie

which shows a potential to be used against obesity as probiotic is *Akkermansia muciniphila*. This normal gut micro flora resides in human intestinal mucosa. Although *Akkermansia muciniphila* is not reported to be cause of any disease, this strain adhere to intestinal mucosa and degrade it. This question on its safety leads to a debate on its use as probiotics. This study is designed with an aim to explore genome of *Akkermansia muciniphila*, strains and get an insight into core, accessory and unique genes it posses to check if this bacteria or some of its starin could be used as probiotic. We have also tried to identify genetic differences among genomes of probiotic, non pathogenic-non probiotics, pathogenic strains. The study is designed with given objectives.

1. Selection of *Akkermansia muciniphila* strains
2. Determination of core and variable genome in *Akkermansia muciniphila* strains
3. Determination of genetic potential as probiotic

Chapter 2

Literature Review

This chapter covers the review of literature published in recent years with respect to probiotic potential. The role of *Akkermansia muciniphila* and its genomic potential as promising probiotic.

2.1 Probiotics and their Role in Gut Health

The word probiotics confer for the live microorganisms that are new word meaning for new life. when they are administered an adequate amount for humans and animals which give beneficial effects [3]. Alternatively, these probiotics have been defined as live microbial supplements that played an active role in the intestinal microbial balance by maintaining the human health [4].

Probiotic most commonly used to improve the health of both animals and humans bythe modulation of intestinal microbiota. According to the reference of beneficial gut microbiota, the well-known genera *Bifidobacterium* and *Lactobacillus* are available in the human gut both together are play active role against the infectious pathogens and also boost up the immune systems of humans [5]. There are many beneficial effects of probiotics directly relate to consumption quantity [6].

These probiotics microorganisms improve the intestinal health of humans by regulating the balance in gut microbiota, enhancing the bioavailability of nutrients, reduce the risk of infections and also enhance lactose tolerance [7]. The availability of gut microbiota is too much matter they can be found in both products dairy as well nondairy. The probiotics most commonly recommended as food supplements after the antibiotics therapy during serious illness, because they destroy the harmful microflora present in the digestive tract. Regular consumption of gut microbiota enhances the positive impact on the human body and established good relations in the population of beneficial microbes in the intestinal flora. The initial role of probiotics to protect against the GI infectious diseases [8]. There are any disturbance in the population of gut microbiota leads to serious GIT diseases and enhanced the infectious of pathogens so the probiotics therapy is recommended to maintain the balance of in beneficial gut microbiota [9].

The development of alternative method such as alternative therapies and adjuvants developments is based on the replacements of bacteria make them more resistant against the antibiotics and leads to adverse effects on the probiotics flora, which enhance the risk of infections [10]. In the last few years advancement of medical sciences also enhanced the knowledge about the gut intestinal microbiota, genetics, immunity and infectious host diseases. Such information provides a suitable way for developments of new appropriate probiotics strains with diseases-specific and could also provide information about the use of the probiotic and how they affect the specific pathological conditions. However, the developments of new probiotics undergo the clinical trial on animals before humans in order to maintain the authenticity, safety, suitability, efficacy and benefits of probiotics for human consumptions [11].

2.2 Significance of Probiotics

It is now a fact that the local community of gut microbiota in the host body is host-specific, location-specific and very diverse in composition and has a lot of

beneficial characteristics. It is still not clear which species of gut microbiota act as a key role in gut beneficial properties. Figure 2.1 show the role of probiotics in maintenance of health. These benefits are the reason for which the number of products containing probiotics especially dairy products are increasing [12].

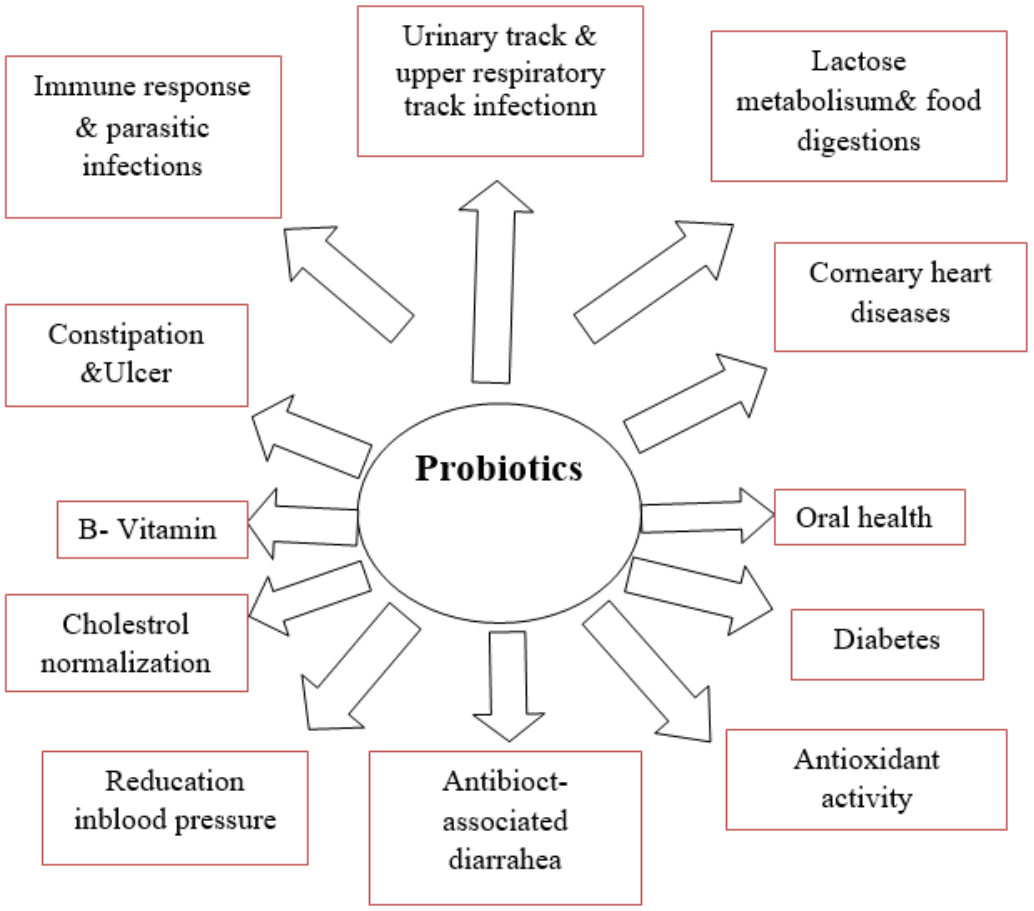

FIGURE 2.1: Role of Probiotics in Human Health [13].

2.3 Mechanisms of Probiotic Action

The functional role of probiotics is not directly related to the colonial population of the intestinal tract. For example, the some of the gut microbiota such as *Bifidobacterium longum* become a member of intestinal gut microflora, while other probiotics like a *Lactobacillus caseiiin* directly played effective role remodeling or influencing the existing community. There are following major role of probiotics in the host body mentioned in Table 2.1 [14].

TABLE 2.1: Mechanisms of Probiotic Actions [15].

Mode	**Process**	**Mechanism**	**Examples**
Barrier function	Reduced the apoptosis in epithelial cells	TNF-α production reduced	*Lactobacillus rhamnosusGG*
	Mucin production enhanced	MUC 2 expression enhanced	*Lactobacillus* .spp
Host cell antimicrobial peptides	hBD protein defensins	Defensins regulation level enhanced up	*E.coli* strains DSM 17252S2
	Cathelicidins	By production of butyrate	
Antimicrobial probiotics factors	Luminal PH lowering	SCFA's secretion	All most all bacteria are probiotics
	Production of bacteriocin	Probiotics gram-positive	
	Production of microcin	Probiotics gram-positive	
Adherence at epithelial	Probiotics compete with pathogens	Production of protein directly or indirectly that stoped the adherence	
Modulation of immune	pro-inflam matory molecules are blocked	By attenuating IL-8 secretions	*Salmonella tyhimurium* VSL#3 probiotics

TABLE 2.1: Mechanisms of Probiotic Actions [15].

Mode	Process	Mechanism	Examples
	Mucosal immunit enhanced	Enhanced the production of IgA	*L. casei*
Quorum sensing signaling interference	Communication between the pathogenic bacteria blocked	Secretes the molecules that blocked the quorum sensing	*L. acidophilus*

2.4 Probiotics as Barriers

Probiotics are competent in changing many components of epithelial intestinal function by regulating mucin production quantity and decreasing the apoptosis of intestinal cells. One of the most common examples of *Lactobacillus rhamnosusGG* present in dairy products as supplements can influence inhibiting tumor necrosis factor (TNF) in intestinal epithelial cells and can the ability to protect cytokine-induced apoptosis in epithelial intestinal cells [16]. *Lactobacillus* species have been playing an active role in the expression of mucin in the intestinal cells in vivo in host epithelial cells thus this mucin production as blocking agents against the pathogenic strain of E. coli in invasion and adherence [17, 18]. *Lactobacillus rhamnosusGG* has played an active role in intestinal cells programmed cell death and inflammation prevention [19]. And also act as an active partner in the regeneration of the mucosal wall and shown the mitogenic effects [20].

2.5 Production of Antimicrobial Substances

Gut microbiota beneficial probiotics induce the changes inside host epithelial cells and produced peptides that are directly released from the epithelial cells these

peptides interference the pathogenic interaction on the epithelial cells and stop the invasion of the pathogen at epithelial cells. Inside the curtain epithelial cells antimicrobial peptides cathelicidins and defensins (hBD protein) released from the cells and expressed the antimicrobial activity against the wide variety of bacteria, fungi and viruses [21].

There are certain probiotics such as *Lactobacilli* species and *E.coli* DSM 17252G2 strains the ability to have shown antimicrobial substances such as defensins [22]. The healthy individuals who have received the 3 weeks proper probiotics treatments who had increased the level of fecal hBD proteins for the 9 weeks [23]. The gut probiotics who have released the antimicrobial substances short fatty acid (SCFA), such as lactic acid and acetic acid, defensins, nitric oxide and bacteriocins hydrogen peroxides which reduced the ph of the lumen that makes them unsuitable environments for the bacterial growths [24].

SCFA causes the chemical changes in the outer membrane gram-negative to act as an inhibiting factor for the growth of pathogens [25]. Bacteriocins another antibacterial factor that easily permeable to the inner membrane of gram-negative bacteria, ultimately lead to disruption and pore formations [26]. Microcins target the inner membrane of gram-negative bacteria and the enzymes which are actively involved in the synthesis of DNA or RNA structure, or proteins synthesis enzymes [27].

2.6 Competition for an Attachment to Intestinal Cells

Probiotics are more competent about pathogenic bacteria and compete for pathogens for the adherence of epithelial cells and more than normal level attached to the mucus layer in a well specific strains manner. The inhibitory factor of *L. helvetics* R0052 outer surface proteins act as a resistance barrier against the adherence of *Escherichia coli* O157: H7 [28]. *S. boulardiisecretes* a substance thermo-labile that

acts as an antibacterial adherence factor that reduced the adherence of pathogens [29].

2.7 Immune Modulation

It has been reported *L. caseiact* as beneficial probiotics have been shown to supplement total and enhanced pathogenic specific secretory of level IgA at the specific infection site in mice and also stimulate the Bclass cell to switch the IgA [30]. There are no specific antibodies produce against the L. casei, so that indicating the immune system of the host does not produce any specific anti-body against the beneficial bacterium [31].

L. casei beneficial probiotics regulate the transcription of the number of different genes that code the pro-inflammatory factor such as chemokines, adherence molecules and cytokines molecules that induced the invasion of *S. flexneriin* intestinal cells. Thesefactors produced the anti-inflammatory result that stope the NF-kβ pathways, particularly through the stabilization of I-kBα [32].

2.8 Intervention with Quorum Sensing Signaling

Quorum sensing signaling is a well-mechanized system in between bacteria to communicate with each other and with the surrounding environment through chemical molecules that are called auto-inducers [33].

This quorum sensing mechanism facilitates the bacteria in colonization and regulation of all important traits of enteric microbes to causes the serious infection in their host body [34].

The probiotics strains such as *Lactobacillus acidophilusin* the gut of the human body secrets the molecule that targets the genes of *E. coli* O157 and stope the transcription and opposed the pathogenicity of bacteria in the human's body [35].

2.9 Role of Probiotics Against Gastrointestinal Diseases

The probiotics research is categorized on two main stages to evaluation of the infectious diseases and their prevention. First stage laboratory studies and second is a clinical trial to check the efficacy and safety Table 2.2.

TABLE 2.2: List of Different Strains of Probiotics Against Gastrointestinal Infectious Diseases [36].

Disease	**Probiotics strain**	**Comment**
Prevention of antibiotic-associated diarrhea (ADD)	*S. boulardii*	The number needed for the treatment of cases is 10.2 prevent.
	Lactobacillus rhamnose GG	Effect on the children and adults in RCT.
Prevention of infection *Clostridiumz difficle* infection (CDI)	*S. boulardii LGG*	Statistically, the result is not significant.
Resist recurrence of after-treatment of CDI	*S. boulardii*	Reduction of CDI recurrence infection
Eradication of *Helicobacter pylori*	*Lactobacillus rhamnose* GG *S. boulardii*	During treatment side effects improve the compliance.
Colitis ulcers	*E. coli* Nissle 1917	Effective role in maintenance of remission.
	VSL	Effective role in the induction and maintenance of colitis ulcer.

TABLE 2.2: List of Different Strains of Probiotics Against Gastrointestinal Infectious Diseases [36].

Disease	**Probiotics strain**	**Comment**
Crohn's diseases	*Lactobacillus rhamnose GG* *Lactobacillus johnsonii LA1*	No Role in Stimulation and prolong remission of CD.
Irritable bowel syndrome	*Bifidobacterium infantis*	Improve the enhancement of IBS syndrome
Acute pancreatitis	*Lactobacillus plantarum*	Incidence of infection enhanced by the PROTERIA trial
Necrotizing enterocolitis (NCE)	*Bifidobacterium spp*, &*Lactobacillus acidophilus*	Probiotics can reduce the NEC and mortality
Multiorgan dysfunction syndrome (MODS)	VSL	Enhanced the concentration of IgG and IgA but the mods are not affected \ By Probiotics.
Immune response and allergy	*Lactobacillus rhamnose GG*	When they are given to pregnant mother decreased the atopic dermatitis
Ventilator-associated pneumonia (VAP)	*Lactobacillus rhamnose GG*	Probiotics also played an effective role in the treatment of (VAP).

2.9.1 Antibiotics-Associated Diarrhea

The prevalence of antibiotic-associated diseases (AAD) ranges from 30% to 5% in the host. The risk of diseases increases by the amino penicillin therapies(Ampicillin or Amoxicillin) which is a combination of clindamycin, cephalosporin and clavulanic acid[37]. The alternative method to adopt to reduce the antibiotic associated-diseases (AAD) such as conjunction of probiotics with antibiotics have been studied on the adults and children. The major changes were observed after the conjunction of probiotics with antibiotics in the gut microbiota decreased the total number and diversity of bacteria such as Bifiobacteria and Bacteroides associated with amylolytic activity decreased and increase the number of facultative bacteria such as Clostridia, Fusobacteria and Eubacteria species [38].

The patients are treated with antibiotics curtains changes observed in the body such as decreased the production of short fatty acid chain and increased the proteolytic activity was noted [39]. Several clinical trials have been conducted using *Saccharomyces boulardii* how much they are effective for the prevention of AAD. After the clinical trial, it has been proved the *Saccharomyces boulardii* is acts as the most effective agent against the AAD [40-42]. Several years the trials were conducted on *Sacchromyces boulardiito* check the efficacy and the effectiveness of probiotics against the prevention of AAD. Randomized control trials on the Sacchromyces boulardii showed a 95% positive result agains the AAD prevention in the adult body [43].

2.9.2 Infections of *Clostridium difficile*

Clostridum difficile is a gram-positive bacterium, that is spore-forming which causes severe gastrointestinal infection with colitis and diarrhea. In the last few decades, *Clostridium difficile* infection CDI has been reported according to severity and incidence. The clinical result reported the CDI infection is asymptomatic mild diarrhea, pseudo membranous colitis. The infections of CDI is the most challenging aspects of diseases. According to recoded data 25% of patients of CDI that

have been treated with the metronidazole and vancomycin but after the 4 week the repeated symptoms of typically disease appears. Due to increasing numbers, frequent death rates and raising reappearance, there is a need for more effective prevention and treatment therapy against the CDI [44].

It proved Probiotics *S. boulardii* produces 54 KDa protease that acts as defensive against the *C. difficile* infection and degraded the A and B toxin which produces *C. difficle* infection and also degraded the colonial receptor site for *C. difficle*. The beneficial bacteria *S. boulardii* also enhanced the level of antibodies IgA level in the intestine that act as antitoxic secretory substance [45].

S. bulardii probiotic supplement that has been studied in the treatment and prevention of recurrence infection of *Clostrium difficile*. The study is based on the several randomized controlled trials of *Saccharomyces boulardii* or *Lactobacillus spp* combination of *C. difficle* toxin [42, 46-49]. Another randomized controlled trial was done on recurrent patients of CDI. Patients of CDI were given two doses in different concentration metronidazole (1g/d) and vancomycin (2g/d or 500mg/d) and *S. boulardii* (1g/d for 4 weeks).

The patients were treated with high doses of probiotics and vancomycin had significantly recurrence rates are reduced (16.7%) and that compared with the placebo and vancomycin (50%) [50]. The probiotics given in the low concentration dose with metronidazole or vancomycin did not show prominent effect against CDI.

S. boulradii only probiotic was shown effective protection against the recurrent infection of *Clostrium difficle* [51]. Probiotics are available in the market as in the form of capsule products such as *Sacchromyces boulardii* present in form of florastor capsules. *Lactobacillus spp* are also available in many other forms of different capsule product culturelle capsule, lactinex and fem-dophilus.

Align probiotics capsules, attune nutrition bars and adult formula CP-1 capsules are also present in the combination form of *Lactobacillus spp* and *Bifidobacterium spp*. There has been enhanced practice of using the probiotics combine with metronidazole and vancomycine for the prevention of recurrence CDI.

2.9.3 *Helicobacter Pylori* infections

The strong gastrointestinal infection bacteria *Helicobacter pylori*, morphological small curved spiral rod-shaped bacterium, this bacterium has strong relation with duodenal peptic ulceration. Helicobacterium pylori is the main infectious agent of causing gastric cancer and chronic gastritis as well as gastric malignancies. Recently therapy which is based on the eradication of this bacterium is a combination of proton pump inhibitors and antibiotics. In vitro study about probiotics showed maximum antimicrobial effect against the Helico bacterium pylori and resist the adherence of bacterium as well as probiotics enhanced the production of metabolites and antimicrobial molecules [52]. A randomized, double-blind, controlled trial was conducted on the 60, participants all were treated with Lactobacillus GG on day 1-14 and with antibiotic therapy on day 1-7 [52] [53]. Probiotics played a very effective role in the diagnosis of diseases and improved the symptoms such as taste disturbance, including nausea and diarrhea; however, eradication treatment did not significantly improve the epigastric pain. In another randomized, double-blind, trial was conducted on the infection patients of 85 H. pylori and these patients were treated with different amount of probiotics such as Lactobacillus acidophilus, Bifidobacterium lactis(group 3), *Saccharomyces boulardii* (group 2), Lactobacillus GG (group 1), on the days of 1-14, with H. pylori treatments on the day of 1-7 [54]. After the different trials of probiotic, it is proved that supplementation with S. boulardii in the treatment of H. pylori infection significantly reduced the adverse effects of therapy especially diarrhea and enhanced the eradication rate of disease [55].

2.9.4 Irritable Bowel Syndrome (IBS)

Intestinal bacteria that are called gut microbiota probiotics after the epidemiological, clinical and physiological studies have suggested the effective role against the pathogenesis of IBS. Many previous studies it is proved that gastroenteritis is one main reason for the IBS [56]. In the last two years, studies that continually

raises the issue of gastroenteritis are directly related to developing the risk of IBS [57]. Physiological studies of humans and animals intestine are directly related to the active function of gut microbiota and alteration in the composition of gut microbiota showed a strong effect on the physiological function of the intestine and IBS [58].

The IBS risk enhanced by the following reason such as elevated luminal gas production, dysbiosis, gastroenteritis and gastrointestinal beneficial gut microbiota and immune activation act as the therapeutic role in IBS [59]. During the serious methodological flaws, various RCTs can check the efficacy of probiotics in IBS patients [60]. Recently Benner and colleague they were reported after the 16 RCT evaluation probiotics in the treatment of IBS, only *Bifidobacterium infantis* probiotics which played an effective role in the improvement of symptoms in the IBS patients [61].

After the detailed study, it is proved that beneficial probiotics played a beneficial impact on the global symptoms than on the flatulence and abdominal pain [62, 63]. In the market *Bifidobacterium infants* available as in the form of Align capsules or present with other probiotics in the form of OWP probiotics capsules and VSL♯3 packets [64].

2.9.5 Ventilator-Associated Pneumonia

Ventilator-associate pneumonia (VAP), is a more adverse form of pneumonia diseases that lead to serious complications in respiration after 48 hours of endo tracheal intubation, the patients shift to intensive care units in the US [65]. The patients of Ventilator-associated pneumonia stayed to remain in the ICU till than the normal function of the lungs started [66].

The patients who suffered in serious infection of pneumonia and gone to VAP the chances of death these patients raises 2 to 10 folds higher as compared to those patients who are mechanically ventilated [67, 68]. The pathogens which may be associated with VAP they are more complex and they have formed biofilms

with aerodigestive tract bacteria and release the contaminated secretion micro-aspiration [69, 70]. Raising the rate of resistance against the antibiotics has promoted the alternative method adapted for the treatments to prevent [71].

In the clinical trial, Forestier et al, using *Lactobacillus caseir hamnosus (Lcr35)* which are played an effective role in VAP in all of the probiotics groups as compared to placebo group (2.9% vs 7.5%). They are reduced the colonization rate of Pseudomonas aeruginosa in gastric as well in the respiratory tract [72].

2.9.6 Allergy and Enhancement of Immune Response

Current study based on the mucosal immunology which build the relation between microbes and host at the early stages when the immune system and mucosal barrier both are still immature [73]. Probiotics act as beneficial potential agent that increase the innate immunity and changes the pathogens inflammation via regulating toll-like receptor signaling pathways [74].

The mode delivery has a great impact on the composition of gut microbiota and also beyond immediate neonatal periods. The infant born delivery also regulate the effective role in the composition of gut microbiota the vaginally born infant and infant born by the cesarean section both have major difference of culture gut microbiota up to 6 months of age [75].

2.10 Safety Concerns with Use of Bacteria as Probiotics

New various evolutionary pressures the DNA of microbes kept changing, these changes are referred as genome plasticity [5]. This phenomenon arises from continuous changes including mutations especially point mutations and conversions, genetic rearrangements as a result of inversions and translocations, indels even

insertions from other bacterial/ viral species such as conjugation plasmids, transposons, bacteriophage and many others. The genomic modifications result in adaptations and behavioral changes in bacterial species based on environmental pressures a specie encounter [18].

Pathogenicity islands and resistance islands are the regions of microbial genomes which possess genes encoding virulence factors and antibiotic resistance genes respectively. These genes are also present in bacterial species classified as potential probiotics.

These genes and their adaption are usually result of genome plasticity. As potential probiotic strain, bacterial species should not have any virulence gene and its should not possess ant antibiotic resistance gene while it should just have natural resistance mechanisms [21].

Probiotics are required to have specific characteristics properties which are encoded in their genes. Mining of bacterial genome for this characteristic is important but it is equally important to check that bacterial species do not posses any unwanted character. Genomic instability in probiotics could be detected by comparative genomic analysis [22]. Genomic stability of probiotic strain is always required to be assured and measures are required to be in place to avoid any mutation or variation [23].

2.11 Acid and Bile Tolerance

As probiotic bacteria are usually introduced through oral rout, therefore their ability to survive in harsh gastrointestinal environment is very important. These conditions involve very low pH, this strongly acidic condition do not allow most of bacteria to survive, similarly survival in presence of bile salts is another stress probiotic have to face. In order to be a potential probiotic bacterium should have gees responsible for tolerance against low pH, heat, cold, oxidative stress and osmosis. All commercially available strains including *L. helveticus MTCC5463*, cheese

starter DPC4571 and DPC5463 posses various genes for acid and bile tolerance [15,76].

2.12 Competitive Exclusion of Pathogens

In order to be a good probiotic candidate, a bacterium should have few genes which provide it with competitive advantage over pathogens, one of the mechanisms by which probiotics impart health benefits. In order to have competitive advantage bacteria should posses genes for bacteriocins or antimicrobial substances, betterment in the state of epithelial barrier, variations and activation of immune system and adhesion to epithelial wall. The potential is more enhanced if bacterium possess capabilities to produce compounds for coaggregation, aggregation and adhesion as well as biosynthesis pathways activation [12].

2.13 Adhesion

Adhesion to intestinal epithelium is the most important property after survival in gastrointestinal tract. Adhesion provides probiotic a potential and competitive advantage over pathogens. Host and probiotic bacterial interaction is dependent on adhesion related proteins. These proteins identify specific receptors in host epithelial cells and binds to them. The binding then activates innate responses including colonization. Adhesion process is mediated by fimbriae or pili present on bacterial surfaces. MSCRAMMs (Microbial Surface Components Recognizing Adhesive Matrix Molecules) also mediate in adhesion [21].

Potential probiotic characterization always involved the hunting of adhesion mediator genes such as *L. rhamnosus* possess spaCBA operon with sortase dependent pili and three secreted genes [22]. *B. coagulans HS243* posses eleven adhesion related genes including fibronectin binding proteins, enolase, flagellar hook associated proteins [77]. Detection of genomic islands using bioinformatics and computational pipelines reveals various potential genes [78].

2.14 Bacteriocin Production

The antagonism of probiotics against E.coli is well known and it was also known from very start that probiotics do this antagonism through certain antimicrobial compounds. Antimicrobial properties of dairy probiotics such as cheese, yogurt and fermented milk products, have long been known, but the concept of bacteriocin production is a bit new [24].

In 1993, first classification of bacteriocins was proposed [25] and after that various attempts have been made to reclassify them [18]. Bacteriocins are divided into four major classes where Class I comprise thermostable compounds also referred as lantibiotics, which which are produced mainly bt gram positive bacteria [79]. Class II includes bacteriocins slightly heavier than class I i.e 10KDa molecular weight and this class is further divided into various subclasses [25].

2.15 Immuno Modulation

Microbes or bacteria whenever enter mammalian body they trigger immune response. As bacteria especially pathogens and microflora have co evolved with mammalians including humans, they impart certain benefits to each other. Gut microbiota for example provides resistance against various diseases. The development of human immune system and its efficiency against diversified pathogens is an outcome of this evolution.

Probiotics have also shown a lot of potential in controlling diseases not only related to digestive tract or gastrointestinal tract but the spectrum goes ahead to neurodegenerative diseases and even cancer. Antibiotic resistance, drug side effects and lack of effective medicines,have shifted the focus on use of probiotic.

Gut microbiota especially in reference to probiotics, activate immune systems and make is stronger the evidence of which is provided by germ free animals who are more prone to develop not only diseases but also deficiencies [26][80].

2.16 *Akkermansia muciniphila* and its Potential as Probiotic

As discussed in Introduction section of this thesis, *Akkermansia muciniphila* and its potential as probiotic is debatable. Although *Akkermansia muciniphila* doesn't cause any disease, but some of its properties make its safety questionable. Table 2.3 Summarizes the Association of Frequency Distribution of *Akkermansia muciniphila* in human gut and the diseases state [76].

TABLE 2.3: Correlation between *Akkermansia muciniphila* and Disease in Humans [78].

Sr. No.	**Disease state**	**Analysis method**	**Obs. & Assoc.**	**Ref**
1.	Type 2 diabetes	Metagenome	Frequency of *A. muciniphila* is less abundant in Diabetics	[31]
2.	Overweight and obese adults	Metagenomic analysis and real time PCR	Frequency of *A. muciniphila* is less abundant in obese patients	[32]
3.	Children with atopic diseases	Pyrosequencing	Frequency of *A. muciniphila* is less abundant in patients as a result decreased efficiency of immune system	[32]

TABLE 2.3: Correlation between *Akkermansia muciniphila* and Disease in Humans [78].

Sr. No.	**Disease state**	**Analysis method**	**Obs. & Assoc.**	**Ref**
4.	Outstanding athletes	16 r RNA sequencing	Frequency of *A. muciniphila* is more abundant in athletes and individuals with low BMI	[34]
5.	Overweight and obese adults	16S r RNA sequencing	No association found	[33]
6.	Autistic children	Real time PCR	Frequency of *A. muciniphila* is less abundant in autistic patients	[34]
7.	Appendicitis, IBD and other diseases	FISH (Fluorescence in situ hybridization)	Frequency of *A. muciniphila* is less abundant in appendicitis patients	[32]
8.	Overweight lactating women	Real time PCR	Frequency of *A. muciniphila* is more abundant in lactating overweight mothers	[33]

Chapter 3

Material and Methods

Probiotic potential of any bacterial strain usually involves its isolation and characterization based on its capabilities of lysozyme tolerance, acid tolerance, antimicrobial activities, resistance to antibiotics, aggregation ability, antioxidant production, and hydrophobicity.

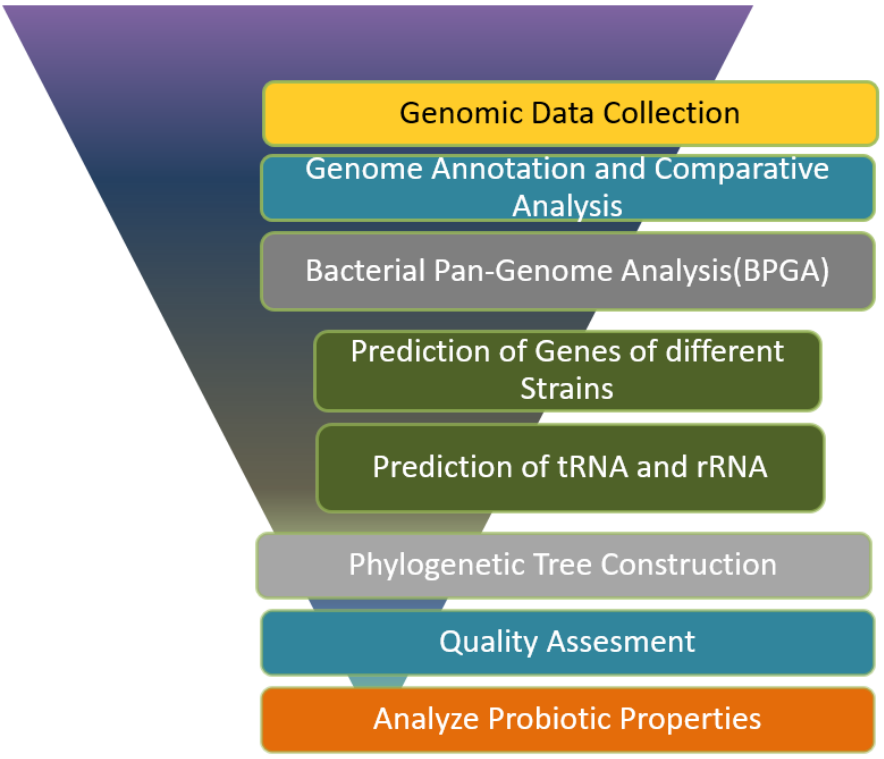

FIGURE 3.1: Summarizes the Methodological Steps used to Analyze Probiotic Potentials of *Akkermansia muciniphila*.

In this project we have used an insilico approach to determine the safety and then probiotic potential of *Akkermansia muciniphila*. The methodology used to analyze this bacterial strain is presented in Figure3.1.

3.1 Genomic Data Collection

First step of methodology was to collect genomic data for analysis, this step comprises three sections i.e. selection of strains based on literature, retrieval of bacterial sequences from databases and selection of reference genome for further analysis [81]

3.1.1 Selection of Strains

Selection of strains was done based on literature survey, the source of literature was NCBI and PubMed. There are 138 genome sequences of *Akkermansia muciniphila* available at NCBI database, but most of them are either exist as scaffold or are incomplete. The ambiguity in source of isolation was another parameter considered, the strains with complete genome but ambiguous source of isolation were not considered. Total nineteen strains were selected based on complete genomic information as well as information regarding source of isolation. Annotated genomic sequences of these strains were retrieved from NCBI databases.

3.1.2 Retreival of Bacterial Sequences

Genomic data is retrieved from NCBI (National center for biological information) database (https://www.ncbi.nlm.nih.gov/genome). Both nucleotide and protein sequences data is retrieved to be used for further analysis. A total of nineteen strains are selected to be used in this study. All the genomes are annotated through RAST [82]. Further, Genome sizes, G + C content, average number of genes, coding DNA sequences (CDS) and other general features are compared to see the

variations between strains. The RAST annotation facilitates to determine the features, assigned to subsystems and help to check the presence in all organisms.

3.1.3 Selection of Reference Strain

The availability of nearly complete *Akkermansia muciniphila* genomes are useful to define the core, accessory and unique genomic features for all the strains. The comparison of strain ATCC BAA-835 with other strains of probiotic, and pathogenic strains, facilitates to find core genes, accessory and unique genes. The genome of *Akkermansia muciniphila* ATCC BAA-835 is used as reference strain [83].

3.1.4 Quality Assessment

Strains of *Akkermansia muciniphila* were selected from NCBI and quality assessment check by CheckM and Patric Databases (https://www.patricbrc.org/). All strains were human and complete genome and quality was good.

3.2 Bacterial Pan-Genome Analysis

Pangenome is total or entire gene set of a particular genus or population under study, using all the available genomes of that genus or population. In order to identify strain-specific genomic features in a genome and determine the genomic diversity among the *Akkermansia muciniphila* strains, the computational pipeline BPGA tool was used [84]. The fundamental purpose of pangenome profile analysis to determine the frequency distribution of the selected strains which in this case are 19*Akkermansia muciniphila* starins. The full GenBank files of all selected genomes were downloaded from NCBI to be used as input for the BPGA analysis. BPGA further processed these files for orthologous cluster analysis and generated an input file containing a total of 4933 annotated genes. In BPGA core, Accessory

and unique gene families were identified using pangenome sequence extraction module. Homologous gene families which were unique to a particular strain were extracted using exclusive gene family analysis module. The pan-genome functional analysis module was used to find the Clusters of Orthologous Groups of proteins (COGs) and KEGG pathway distribution. Evolutionary analysis done by BPGA was based on concatenated core gene alignment using a binary pan-matrix file that depicts the presence or absence of the genes

3.2.1 Comparative Analysis of Orthologous Genes

Comparative analysis of core genes to detect the presence of single copy genes and multiple copy genes in all selected nineteen strains was performed using OrthoMCL(https://orthomcl.org/orthomcl/). A particular Cluster of Orthologous genes is a group of genes which have evolved together and are evolutionary counterparts or orthologues. Within the Clusters of Orthologous Gene (COG), clusters including DNA replication, transcription and translation, metabolism, growth and stress response, there is a long list of functional categories of these core genes. These categories were analyzed using webMGA server. (http://weizhonglab.ucsd.edu/webMGA). This analysis was done to provide insights into the diversity of genes within a particular COG category, i.e. similarity between core genes annotation. These accessory genes analysis was also helpful to understand the subsystems as well as the abundance of these genes among various groups. In this way we can identify important unique genes and their characteristic role within a particular strain.

3.2.2 Phylogenetic Analysis

Evolutionary trees are constructed based on similarities and differences among gene sequences. Although these trees are predictions not a definitive fact but they provide information related to the evolution from common ancestor. These

trees could be constructed using methods which involve construction using whole-genome methods or concatenated single gene sequence. BPGA tool was used to construct Phylogenetic Tree [84].

3.3 Antibiotic Resistance (Resistome) Determinants

Antibiotic resistance is one of the characteristics which is favorably required to be present in a probiotic. Intrinsic resistance to antibiotics provides capabilities to probiotic strains to regain their abundance in gut after the use of antibiotics against pathogens. On the other hand, the resistance against antibiotics in bacterial species is a global concern. Screening of probiotic bacteria for antibiotic resistance genes ensures their safety to be used as probiotics so that they cannot transfer these resistance genes to other bacteria through horizonal gene transfer mechanisms. Comprehensive Antibiotic Resistance Database CARD, (https://card.mcmaster.ca/analyze/rgi), is a database used for screening of antibiotic resistance determinants. The database was used to identify that either a particular strains harbors gene for resistance against various drugs as well compare and evaluate the differences [85].

3.4 Virulence Factors

The capability of a particular bacterial strain to cause disease is refereed as its pathogenicity, while the severity of damage or disease it will cause is its virulence. Molecules which enhance the capability and severity of diseases causing abilities in a bacterial strain are called virulence factors. These factors include the molecules/proteins which enable bacteria to adhere and colonies, evade host immune system and many more. A potential probiotic strain should not possess these genes. Virulence Factor of Bacterial Pathogens Database VFDB(http://www.mgc.ac.cn/VFs/) is used to confirm the presence of putative virulence genes [86].

3.4.1 Genomic Islands Determination

Horizontal gene transfer results in formation of clusters of genes referred as Genomic Islands. Horizontal gene transfer could be outcome of any mechanisms including transposons, bacteriophages or plasmids. As these clusters were first studied in pathogenic bacteria therefore were referred as Pathogenicity Islands. Now a days they are usually referred with reference to property they impart such as Metabolic Islands, symbiosis Islands, Antibiotic resistance islands and so on. As these Islands are acquired by Horizontal gene Transfer therefore their presence may variate among closely related strains of same or different species. IslandViewer4 (https://www.pathogenomics.sfu.ca/islandviewer/)is a tool which utilizes three prediction algorithms including SIGI-HMM, IslandPath-DIMOB, and IslandPick. Bacteriocin Production and Bioactive Islands were determined using online data base BAGEL [87].

Chapter 4

Result and Analysis

Akkermansia muciniphila, is a member of normal gut microbiota, which is reported to have positive impacts on health in obese patients. These positive impacts make it a potential probiotic to be used against obesity. Although the bacterium is not directly involved in causing any diseases but certain properties it possesses creates a debate on its safe use. The thesis is designed as an attempt to analyze various genetic properties of *Akkermansia muciniphila* to evaluate its potential to be used as probiotic.

4.1 Genomic Data Collection

Akkermansia muciniphila, is a common inhabitant of mammalian gut, and is reported to have 138 different strains isolated from various sources. Some of these strains are sequences and whole genome annotated sequences are available.

4.1.1 Selection Of Strains

For this project first step was to select an inclusion and exclusion criteria for selection of bacterial strains, all the strains with complete genomic sequence available along with a known source of isolation were selected. Nineteen bacterial strains

isolated from humans which had complete whole genome annotated sequences were selected. Table 4.1 summarizes the details of selected strains of *Akkermansia muciniphila* selected for further analysis. Selected strains were verified using literature analysis and their genomic properties were analyzed. The whole genome sequences of all 19 strains were downloaded from NCBI database [88].

TABLE 4.1: List of Selected Strains After Literature Review, Strains were Selected Based on Availability of Complete Genome and Information Related to Source of Isolation.

S. No	**Genome Name**	**Genome ID**	**Source of isolation**	**Genome Status**
1	Refrence Strain *A. muciniphila* ATCCBAA-835	349741.6	Human Feces	Complete
2	*A. muciniphila* CBA5201	239935.2076	Human Feces	complete
3	A. muciniphila DSM 22959	239935.2131	feces	complete
4	*A. muciniphila* JCM 30893	239935.2189	Human feces	complete
5	*A. muciniphila* AMDK-7	239935.264	Korean Adult Feces	Complete
6	*A. muciniphila* AMDK-8	239935.265	Korean Adult Feces	Complete
7	*A. muciniphila* AMDK-10	239935.255	Korean Adult Feces	Complete
8	*A. muciniphila* AMDK-11	239935.256	Korean Adult Feces	Complete
9	*A. muciniphila* AMDK-12	239935.257	Korean Adult Feces	Complete
10	*A. muciniphila* AMDK-13	239935.266	Korean Adult Feces	Complete

11	*A. muciniphila* AMDK-14	239935.258	Korean Adult Feces	Complete	
12	*A. muciniphila* AMDK-15	239935.267	Korean Adult Feces	Complete	
13	*A. muciniphila* AMDK-16	239935.262	Korean Adult Feces	Complete	
14	*A. muciniphila* AMDK-17	239935.263	Korean Adult Feces	Complete	
15	*A. muciniphila* AMDK-18	239935.259	Korean Adult Feces	Complete	
16	*A. muciniphila* AMDK-19	239935.26	Korean Adult Feces	Complete	
17	*A. muciniphila* AMDK-20	239935.268	Korean Adult Feces	Complete	
18	*A. muciniphila* AMDK-21	239935.269	Korean Adult Feces	Complete	
19	*A. muciniphila* AMDK-22	239935.261	Korean Adult Feces	Complete	

S. No.	**Genome length**	**GC content**	**GenBank Accession**	**No of proteins**	**No of RNAs**
1	2664102	55.82311	CP001071	2498	62
2	2819944	55.323246	CP033388	2336	62
3	2819944	55.762352	CP042830	2109	62
4	2819944	55.637306	AP021898, AP021899	2252	62
5	2819944	55.29845	CP025823	2212	62
6	2819944	55.392254	CP025824	2150	62
7	2819944	55.24988	CP025825	2260	62
8	2819944	55.25507	CP025826	2218	62
9	2819944	55.255096	CP025827	2191	62
10	2819944	55.251026	CP025828	2156	62

11	2819944	55.25402	CP025829	2198	62
12	2819944	55.30021	CP025830	2352	62
13	2819944	55.30107	CP025831	2203	62
14	2819944	55.301273	CP025832	2208	62
15	2819944	55.30088	CP025833	2208	62
16	2819944	55.31754	CP025834	2195	62
17	2819944	55.317074	CP025835	2135	62
18	2819944	55.315376	CP025836	2112	62
19	2819944	55.315674	CP025837	2210	62

4.1.2 Selection of Reference Genome

A nucleotide sequence assembly used as representative example of genes present in a particular bacterial species is referred s reference genome. These reference genomes act as guide for annonation and assembly of new genomes.

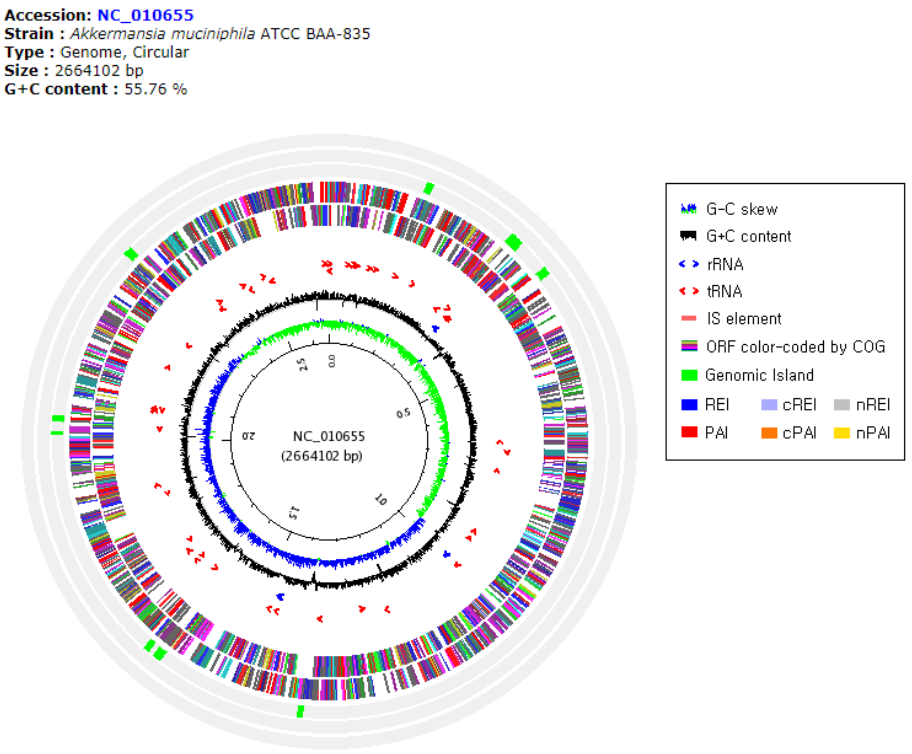

FIGURE 4.1: Circular Genomic View of *Akkermansia muciniphila* Strain ATCC BAA-835 along with Major Genomic Regions.

NCBI hosts a database for reference sequences. In case of *Akkermansia muciniphila*, strain ATCC BAA-835is often used as reference strain. A circular genome of this strain comprises 2664102 bp of nucleotides. Average GC content of the strain is 55.8% (Table 4.1) with total 88.8% of coding genome making predicted protein coding genes number of 2176 genes. Out of these protein coding genes 65% (1408) genes are assigned with a functional role while 35% (768) genes are hypothetical genes and 1.7 % (38) are classified as pseudogenes. Figure 4.1 shows the circular genome diagram as well as few genomic features.

4.1.3 Gene Prediction and Annotation

Gene prediction and annotation was performed by using Rapid Annotations using Subsystems Technology (RAST), online freely available tool. This online tool provides accession number that is a unique identification of every sequence. By submitting sequence in FASTA format, it provides size, GC%, No. of contigs, No. of Coding Sequences of strains. Table 4.2 summarizes the findings of Genome Annotation using RAST server Organism Genes Genes of known or predicted molecular function Protein-coding Genes tRNA genes rRNA genes Pseudo- genes Genes of unknown molecular function

TABLE 4.2: Summarizes the Findings of Genome Annotation Using RAST Server

S.No	**Organism**	**Genes**	**Genes of predicted molecular function**	**Protein coding Genes**
1	*A. muciniphila* ATCC BAA-835	2310	919	2246
2	*A. muciniphila* CBA5201	2373	543	2308
3	*A. muciniphila* DSM22959	2379	703	2315
4	*A. muciniphila* JCM30893	2379	524	2220
5	*A. muciniphila* AMDK-7	2316	559	2251

6	*A. muciniphila* AMDK-8	2327	700	2263
7	*A. muciniphila* AMDK-10	2282	670	2218
8	*A. muciniphila* AMDK-11	2257	703	2193
9	*A. muciniphila* AMDK-12	2257	706	2193
10	*A. muciniphila* AMDK-13	2269	659	2205
11	*A. muciniphila* AMDK-14	2265	693	2201
12	*A. muciniphila* AMDK-15	2277	536	2213
13	*A. muciniphila* AMDK-16	2258	704	2194
14	*A. muciniphila* AMDK-17	2271	538	2207
15	*A. muciniphila* AMDK-18	2277	539	2213
16	*A. muciniphila* AMDK-19	2204	695	2140
17	*A. muciniphila* AMDK-20	2214	684	2150
18	*A. muciniphila* AMDK-21	2222	676	2158
19	*A. muciniphila* AMDK-22	2194	680	2130

S.No	**Organism**	**tRNA genes**	**rRNA genes**	**Pseudo genes**	**Genes of unknown molecular function**
1	*A. muciniphila* ATCC BAA-835	52	9	9	1391
2	*A. muciniphila* CBA5201	53	9	29	1830
3	*A. muciniphila* DSM22959	52	9	37	1676
4	*A. muciniphila* JCM30893	52	9	30	1732
5	*A. muciniphila* AMDK-7	53	9	43	1757
6	*A. muciniphila* AMDK-8	52	9	43	1627
7	*A. muciniphila* AMDK-10	52	9	113	1612
8	*A. muciniphila* AMDK-11	52	9	38	1554
9	*A. muciniphila* AMDK-12	52	9	34	1551
10	*A. muciniphila* AMDK-13	52	9	107	1610
11	*A. muciniphila* AMDK-14	52	9	51	1572
12	*A. muciniphila* AMDK-15	52	9	61	1741

13	*A. muciniphila* AMDK-16	52	9	56	1554
14	*A. muciniphila* AMDK-17	52	9	53	1733
15	*A. muciniphila* AMDK-18	52	9	55	1738
16	*A. muciniphila* AMDK-19	52	9	44	1509
17	*A. muciniphila* AMDK-20	52	9	55	1530
18	*A. muciniphila* AMDK-21	52	9	66	1546
19	*A. muciniphila* AMDK-22	52	9	75	1514

As for the project the selection criteria were to select the annotated complete genomes, but annotation was performed again with RAST server to attach significant functional information with genomes. RAST, stands for Rapid Annotation using Subsystem Technology, is standard software pipeline established in 2008 for annotation of bacterial and archaeal species. Annotation was performed to identify the gene coding region and out of this region, the focus has been on the genes which have a known function. For this project, the pseudogenes as well as hypothetical genes were not taken into account. Similarly, genes encoding tRNA and rRNA were also not considered further. Convenience, consistency and speed of analysis are three major features of RAST for which this server was used to identify the protein coded regions in all the selected nineteen genomes.

4.2 Pan Genome Analysis

With the advent of next generation high throughput technologies in genomics, the focus has analysis shas shifted from one isolate to a complete/ entire picture from all reported strains. Therefore, we can say that we have shifted from genome to genomics. Pan genomics is one of the emerging Omics techniques where instead of analyzing genes present in single genome, the concept is to analyze the entire set of genes from bacterial population under study. In pangenome analysis, genes present in all the bacterial strains is referred as ‘core genome’, gene set present in few bacterial species is referred as ‘dispensable genome’ and gene set present in

only one bacterial species is referred as 'Unique genome'. 'Accessory genome' also referred as 'variable genome' is the set of genes which is not present in all strains, comprising dispensable and unique genome sets.

For the selected nineteen strains of *Akkermansia muciniphila* BPGA i.e. Bacterial Pan Genome Analysis Algorithm was used. BPGA is a step by step bioinformatics pipeline which analysis pangenome using core genome modules. The process starts with the input of data file, the input files could be in three formats GenBank file, NCBI FASTA format File, Protein Sequence file. The file format could be chosen as per analysis requirements. The tool can also process tab delimited binary files retrieved from another tool for pan genome analysis.

The strategy used for this thesis was all against all comparisons in all nineteen selected genomes instead of using sequence comparison against reference genome. To identify core and variable genomes all selected nineteen strains, whole genome sequences in NCBI format were used. Whole genome sequences were downloaded from NCBI Genome Database and BPGA file was generated as initial input file preparation process. This BPGA generated file comprised 4933 sequences of from all selected strains and this file with annotated sequences was later used as an input file for clustering. BPGA has three different tools options for clustering including Ortho MCL, CD-HIT and USEARCH. BPGA in default uses USEARCH but for this project we used all three available options in parallel to validate clustering.

All nineteen selected strains of *Akkermanisa muciniphila* were analyses and to fined core and pan genome, it was found that all selected strains share 1035 genes making core genome. The gene accumulation curve shown in Figure 4.2 indicate that as we add a new strain in analysis there is decrease in number of core genome. Similarly, the pan genome depicts an increase as number of genes with addition of each strain. This increasing trend in pan genome indicates that *Akkermanisa muciniphila* is an open genome where each strain has quite a high number of unique genes, we can conclude from this trend that with each strain number of unique genes are added to pangenome. In all strains under study, total of 1489 unique

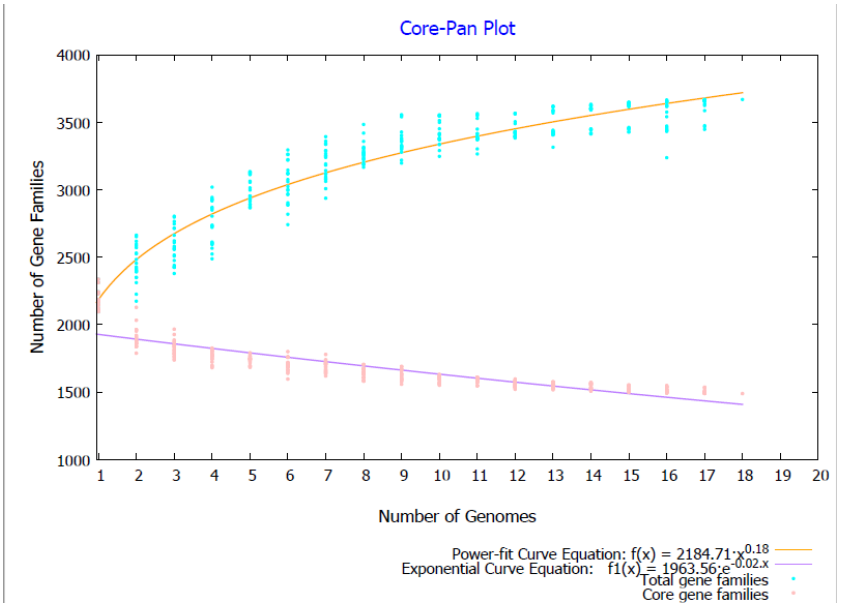

FIGURE 4.2: Gene Accumulation Curve of All Selected Strains, the Curve in Purple Indicates Core Genome and the Trend Indicates that the Number of Core Genes Decrease with Addition of New Strain. The curve in Orange Color Indicates that Pangenome which Show the Trend of Increase in Pangenome with Addition of Each Strain.

protein coding genes were found which is one third of average protein coding genes number in all selected strains.

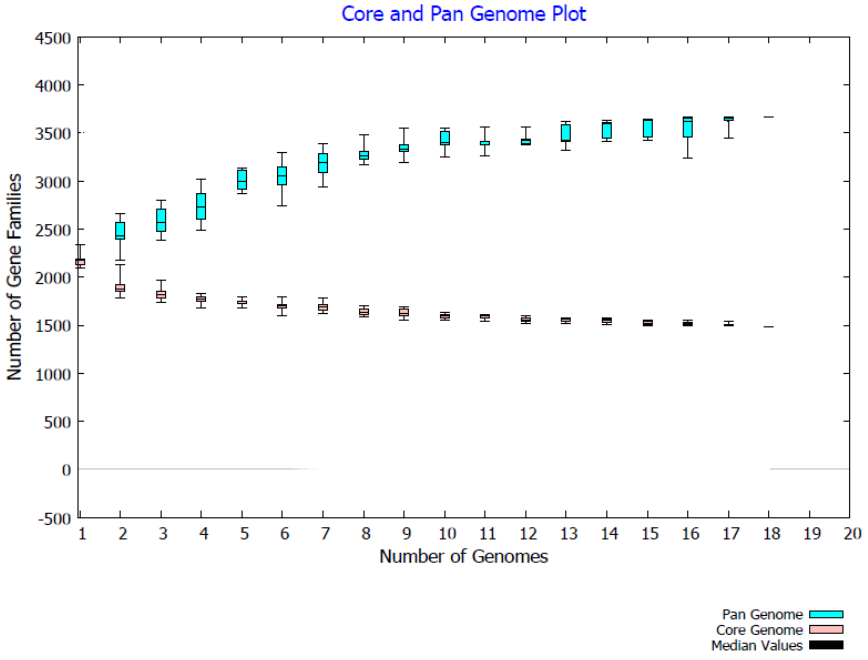

FIGURE 4.3: The Pan Genome Profile Trends Obtained using BPGA, Pangenome is Indicated by Cyan Color and Show an Increasing Trend, While Pink Color Show Core Genome Depicted a Decrease with Addition of Each Strain.

The number of unique genes which is quite high in case of *Akkermansia muciniphila* strains under study. Out of 1489 genes 1322 genes were strain specific genes and the remaining were additional accessory genes present in few strains as well. The high number of variable genomes indicates that horizontal gene transfer is quite frequent phenomenon in *Akkermansia muciniphila.* the reason for this could be the variation in environments from where strains were isolated. Although in inclusion criteria, strains were selected to be from humans, but as the gut microbiome depends a lot on the type of food habits an individual possesses. Similarly, the health status and the use of antibiotics also imparts a stress and makes a bacterium to uptake new genes. As gut is quite a crowded area therefore, we can predict that the genes acquired not only from other strains but also from other species based on the stress and available neighborhood. Figure 4.3, show core and accessory genome of all 19 selected strains and validate the curves obtained in Figure 4.2. Table 4.3 indicates observations against each strain.

TABLE 4.3: Information about Core, Variable and Unique genes.

S no.	**organisms**	**No. of core genes**	**No. of accessory genes**	**No. of unique genes**
1	*A. muciniphila* _ATCC_BAA_835	1489	615	6
2	*A. muciniphila* CBA5201	1489	601	222
3	*A. muciniphila* DCM22959	1489	604	3
4	*A. muciniphila* JCM30893	1489	653	194
5	*A. muciniphila* AMDK-7	1489	538	202
6	*A. muciniphila* AMDK- 8	1489	671	83
7	*A. muciniphila* AMDK- 10	1489	609	45
8	*A. muciniphila* AMDK-11	1489	687	0
9	*A. muciniphila* AMDK-12	1489	692	4
10	*A. muciniphila* AMDK-13	1489	624	34
11	*A. muciniphila* AMDK- 14	1489	666	9

12	*A. muciniphila* AMDK-15	1489	671	14
13	*A. muciniphila* AMDK-16	1489	683	7
14	*A. muciniphila* AMDK-17	1489	681	10
15	*A. muciniphila* AMDK-18	1489	679	10
16	*A. muciniphila* AMDK-19	1489	680	10
17	*A. muciniphila* AMDK- 20	1489	627	6
18	*A. muciniphila* AMDK-21	1489	605	15
19	*A. muciniphila* AMDK-22	1489	596	12

S no.	**Organisms**	**No. ofexcl. absent genes**	**Accessory Gene %**	**Unique genes %**
1	*A. muciniphila* ATCC_BAA_835	3	41.30289	0.402955003
2	*A. muciniphila* CBA5201	24	40.36266	14.90933512
3	*A. muciniphila* DCM22959	1	40.56414	0.201477502
4	*A. muciniphila* JCM30893	11	43.85494	13.02887844
5	*A. muciniphila* AMDK-7	16	36.13163	13.56615178
6	*A. muciniphila* AMDK- 8	3	45.0638	5.57421088
7	*A. muciniphila* AMDK- 10	47	40.89993	3.022162525
8	*A. muciniphila* AMDK-11	2	46.13835	0
9	*A. muciniphila* AMDK-12	1	46.47414	0.268636669
10	*A. muciniphila* AMDK-13	43	41.90732	2.283411686
11	*A. muciniphila* AMDK- 14	7	44.72801	0.604432505
12	*A. muciniphila* AMDK-15	6	45.0638	0.940228341
13	*A. muciniphila* AMDK-16	6	45.86971	0.470114171
14	*A. muciniphila* AMDK-17	3	45.73539	0.671591672
15	*A. muciniphila* AMDK-18	3	45.60107	0.671591672
16	*A. muciniphila* AMDK-19	3	45.61107	0.671591572
17	*A. muciniphila* AMDK- 20	7	42.1088	0.402955003
18	*A. muciniphila* AMDK-21	15	40.6313	1.007387508
19	*A. muciniphila* AMDK-22	14	40.02686	0.805910007

For further validation of pangenome analysis, Roary was used with the parameter of 90% BLAST p percentage identidity cuttoff. This tool clustered the genes further into hard core and soft core genes. Similarly Accessory genes are categorised as shell and cloud genomes. In case of *Akkermansia muciniphila* under study more than 99% of genes were classified as hard core while 95-99 % could be easily categorized as soft core. Shell genes were 15-95 % while cloud genes were less than 15 %. Figure 4.4 and Table 4.4 summarizes the results of Roary analysis [89].

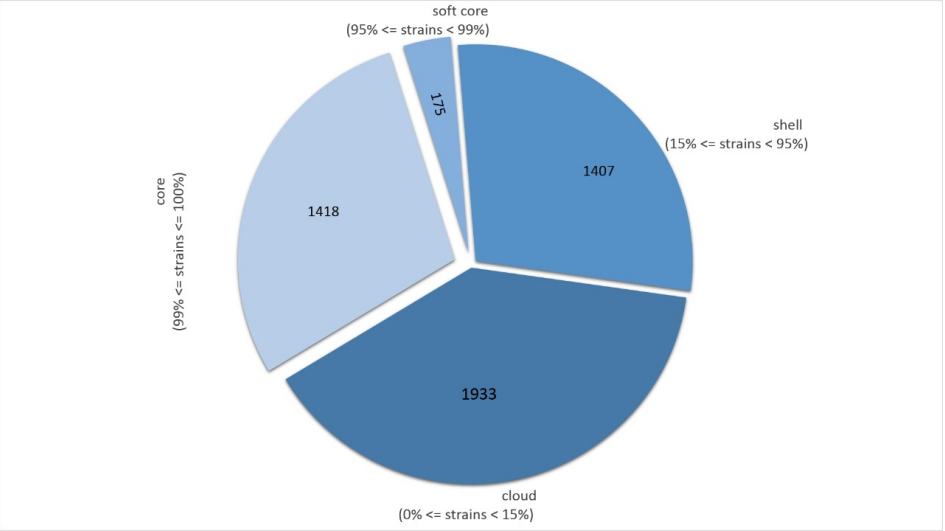

FIGURE 4.4: Information About Core Accessory Unique Genes from Roary

TABLE 4.4: Results from Roary for Pangenome Analysis

Core genes	(99% < strains < 100%)	1418
Soft core genes	(95% < strains <99%)	175
Shell genes	(15% < strains <95%)	1407
Cloud genes	(0% < strains <15%)	1933
Total genes	(0% < strains <= 100%)	4933

4.2.1 Exclusive Gene Family Analysis

In order to find genes which are exclusively present in a particular strain or are unique genes, a special feature of BPGA referred as 'Exclusive Gene Family Analysis' is used. Figure 4.5 summarizes the frequency of singletons or unique genome

of each strain [90].

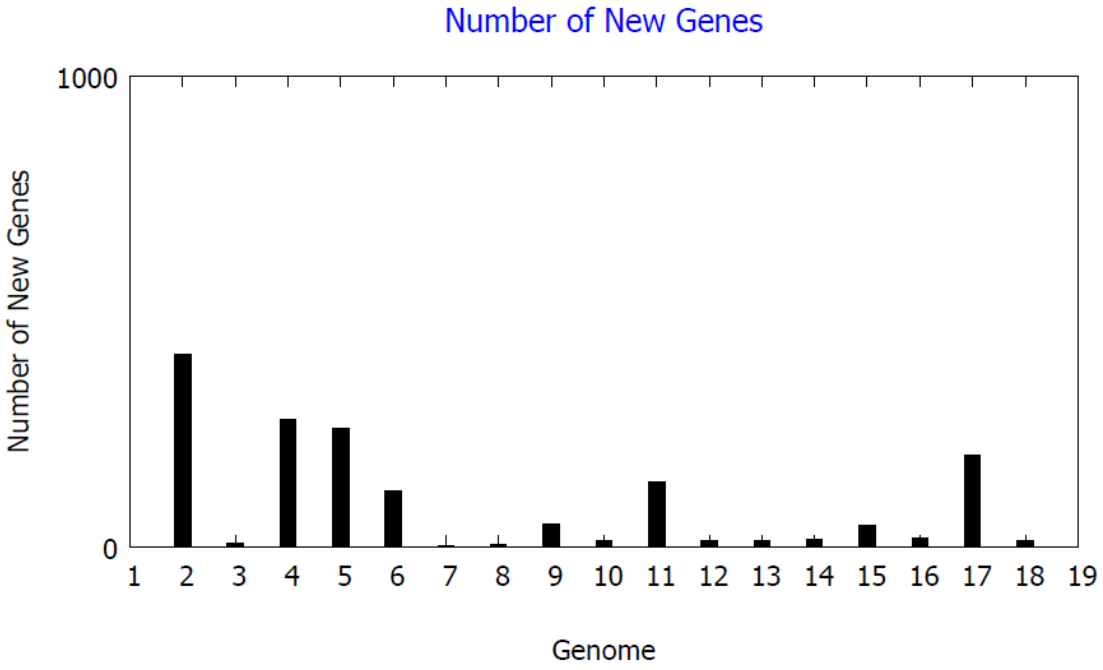

FIGURE 4.5: Number of New Genes or Unique Genes Added to Pangenome with Addition of Each New Strain in Pan Genomic Analysis

4.2.2 Sequence Extraction

For further analysis protein sequences and genome sequences were required therefore protein sequences of all core, unique and accessory genes were extracted as FASTA files using a special module of BPGA referred as 'Pan Genome Sequence Extraction'.

4.2.3 Phylogenetic Analysis

Phylogenetic analysis provides as insight into the diversity of strains under study. To construct phylogenetic tree BPGA can help to construct three different types of trees based on insilico Multi Locus Sequence Tags (MLST), or based on concatenated core gene alignment, or based on pan-matrix. The phylogenetic tree of all 19 strains was constructed using USEARCH clustering module. Figure 4.6 show phylogenetic tree constructed using BPGA.

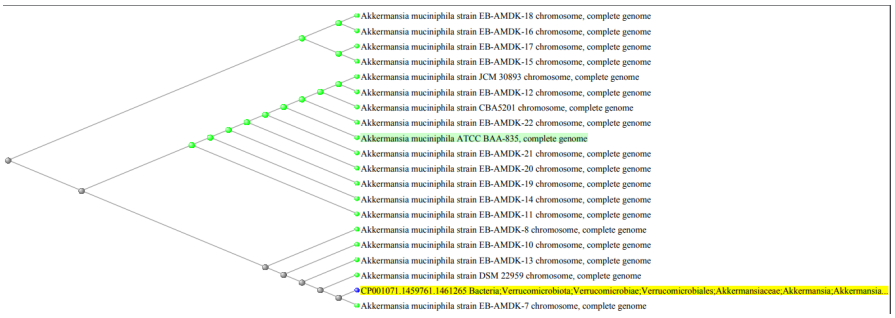

FIGURE 4.6: Phylogenetic Tree of All Selected Strains

4.2.4 Clusters of Orthologous Genes

Cluster of orthologous genes (COG) is a collection of genes from various organisms with common ancestors. COG analysis is one of the most important analysis after pan genome analysis to comprehend what role core and especially unique genes play in a particular organism. For COG analysis of all selected strains 'Pan Genome Functional Analysis' module of BPGA was used. The module uses COG function and KEGG pathway mapping for the given protein sequences (retrieved from BPGA earlier section 4.2.2) representing core and accessory genome. Figure 4.7 represents COG and distribution of core, accessory and unique genomes.

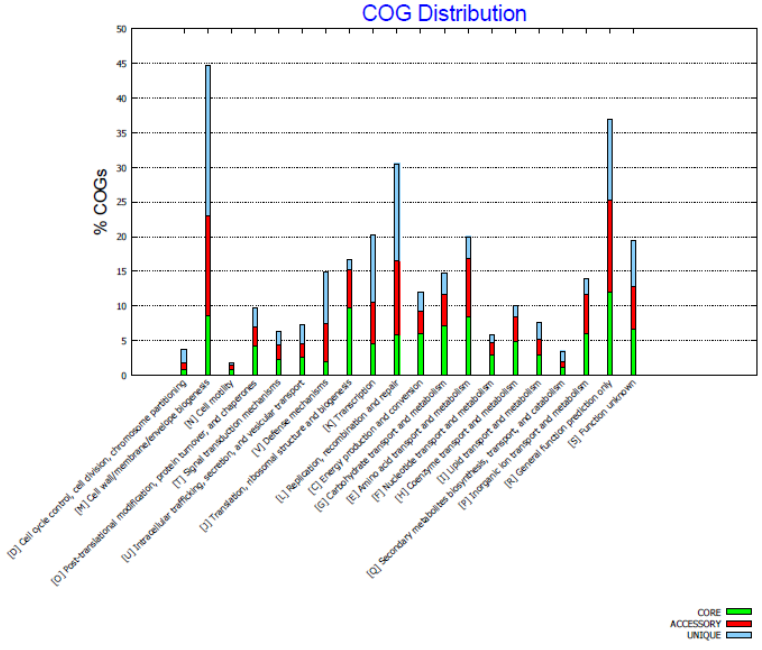

FIGURE 4.7: Clusters of Orthologous Groups (COG) and Distribution of the Core Genes, Accessory Genes and Unique Genes in *Akkermansia muciniphila* Strains.

The distribution pattern of COG depicts that most of core genes are involved in 'translation, ribosomal structure and biogenesis', 'cell wall, membrane, envelope biogenesis' and then 'carbohydrate, transport and metabolism'. On the other hand, most of the genes from accessory genome are related with 'Transcription, Replication, Recombination and Repair' then another cluster of 'Cell wall, membrane and envelope biogenesis'.

For pathway mapping, BPGA could map 1218 gene clusters out total 2790 gene clusters making 43.7% with KEGG pathways. BPGA used USEARCH, CD-HIT, and OrthoMCL tools for clustering and indicated that most of the pathways were related with metabolism. These core and accessory gene distribution among various clusters was validated by comparing our results with Clusters of Orthologous Groups Database.

In this database a difficulty was faced related to available data. Most of the genes of *Akkermansia muciniphila* were not available indicating that less data is available for this bacterium. As depicted in Figure 4.8 most of core genes are associated with metabolism, cell wall biogeneiss and process of transcription and translation indicating that these processes are conserved in all strains

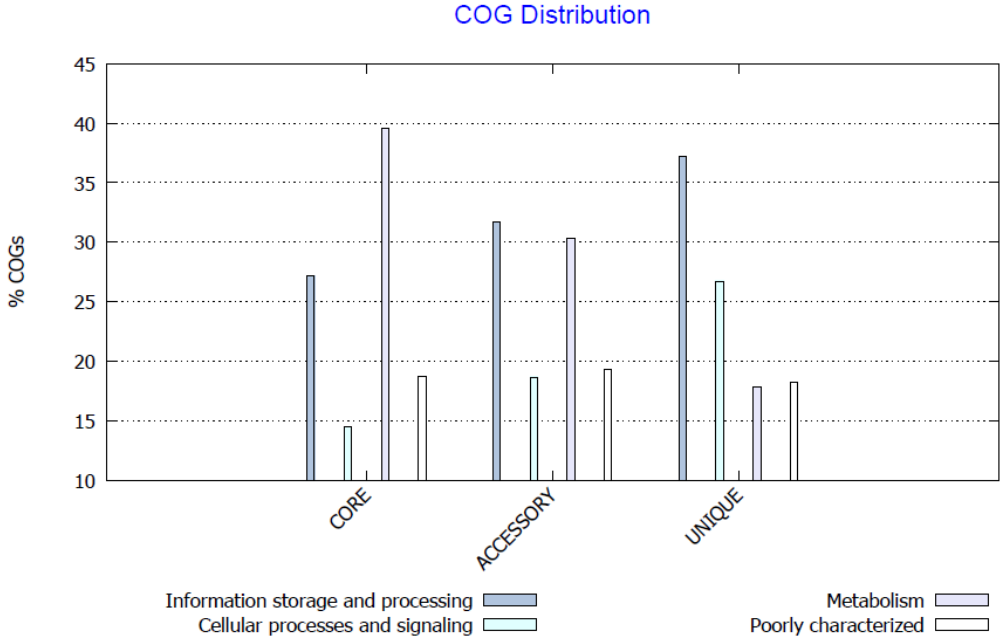

FIGURE 4.8: Distribution of the Core genes, Accessory Genes and Unique Genes Involve in Different Processes in *Akkermansia muciniphila* Strains.

4.2.5 Determination of Genetic Islands

Most of the adaptive characters of a bacterium are located in close proximity of each other in prokaryotic genome. This is often an indication of horizontal gene transfer, as well as it provides significance in expression of these adaptive traits. Genomic islands are usually characterized based on the adaptive advantage they provide. In this project antibiotic resistance Islands and pathogenicity islands were analyzed Results of which are summarized in Table 4.5. Presence of Bacteriocin production gene and pathogenicity islands were also predicted [91].

Table 4.5: Antibiotic Resistance Genes through CARD Analysis

Organism	**Category**	**Gene**	**Drug Class**
ATCC BAA-835	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone
AMDK-8	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone
AMDK-7	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone
AMDK-10	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone
AMDK-11	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone
AMDK-12	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone
AMDK-13	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone
AMDK-14	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone
AMDK-15	23 S rRNA methyletransferase	ErmB	Marcolide, Lincosamide, Streptogramin

AMDK-16	Antibiotic Efflux, Resist. in Cell Division 23 S rRNA methyletransferase	adef ErmB	Tetracyclin, Fluoriquinolone Marcolide, Lincosamide, Streptogramin
AMDK-17	Antibiotic Efflux, Resist. in Cell Division 23 S rRNA methyletransferase	adef ErmB	Tetracyclin, Fluoriquinolone Marcolide, Lincosamide, Streptogramin
AMDK-18	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone
AMDK-19	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone
AMDK-20	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone
AMDK-21	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone
AMDK-22	No data available		
AMDK-30893	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone
CBA-521	Antibiotic Efflux, Resist. in Cell Division	adef	Tetracyclin, Fluoriquinolone

Organism	**Resist. mechanism**	**Identity of matching region**	**Identity of Reference Region**
ATCC BAA-835	Antibiotic efflux	41.27	99.72
AMDK-8	Antibiotic efflux	41.27	99.72
AMDK-7	Antibiotic efflux	41.36	99.72
AMDK-10	Antibiotic efflux	41.42	99.72
AMDK-11	Antibiotic efflux	41.45	99.72

AMDK-12	Antibiotic efflux	41.45	99.72
AMDK-13	Antibiotic efflux	41.45	
AMDK-14	Antibiotic efflux	98.37	98.79
AMDK-15	Antibiotic Target Alteration	98.37	98.79
AMDK-16	Antibiotic efflux Antibiotic	41.52	114.26
	Target Alteration	98.37	98.79
AMDK-17	Antibiotic efflux Antibiotic	41.55	99.72
	Target Alteration	98.37	98.79
AMDK-18	Antibiotic efflux	41.55	99.72
AMDK-19	Antibiotic efflux	41.55	99.72
AMDK-20	Antibiotic efflux	41.36	99.72
AMDK-21	Antibiotic efflux	41.36	99.72
AMDK-22			
AMDK-30893	Antibiotic efflux	41.36	99.72
CBA-521	Antibiotic efflux	41.27	99.72

Table 4.5 is compiled from the results of CARD database which clearly indicates that all the selected strains of *Akkermansia muciniphila* possess some antibiotic resistance against first grade antibiotics but they do not show any indication of antibiotic resistance against all antibiotics or in other words, Multi Drug Resistance. Hence in this regard they are safe to be used as probiotics. Probiotic bacteria are part of normal gut microflora and intrinsic resistance to certain very commonly exposed antibiotics, problem arises if these strains develop resistance against most of the antibiotics and become resistant microbe or pathogen. In case of *Akkermansia muciniphila*, it is observed that all the selected strains are found resistance to commonly used antibiotics (which in one way is essential to keep the gut microbial composition), the genes encoding the resistance against these antibodies is present on mobile elements or plasmid, that indicates that bacteria especially all commensal bacteria in a common environmental niche share these

antibiotic resistance. Most of the accessory genes, which are acquired by the microbial strain to carry on life activities in a better way, in other words adaptive advantages, are acquired through horizontal gene transfer and they reside and move as a block referred as genomic island. As these are adaptive genes therefore it is always essential to analyses that what type of genes it possesses. Based on the type of genes, genetic islands are classified as pathogenicity islands, symbiosis islands, metabolic islands, resistance islands and fitness islands. In vase of *Akkermansia muciniphila*, pathogenicity islands were searched and no pathogenic gene was found all of the selected strains but in order to ensure not only the safety but the bacteriocin production capacity, all strains were analyzed through Island Viewer and BAGEL. Details of each starin are as follows

4.2.6 *Akkermansia muciniphila* ATCC BAA-835

Figure 4.9 Indicates the presence of Genomic islands in *Akkermansia muciniphila* ATCC BAA-835.

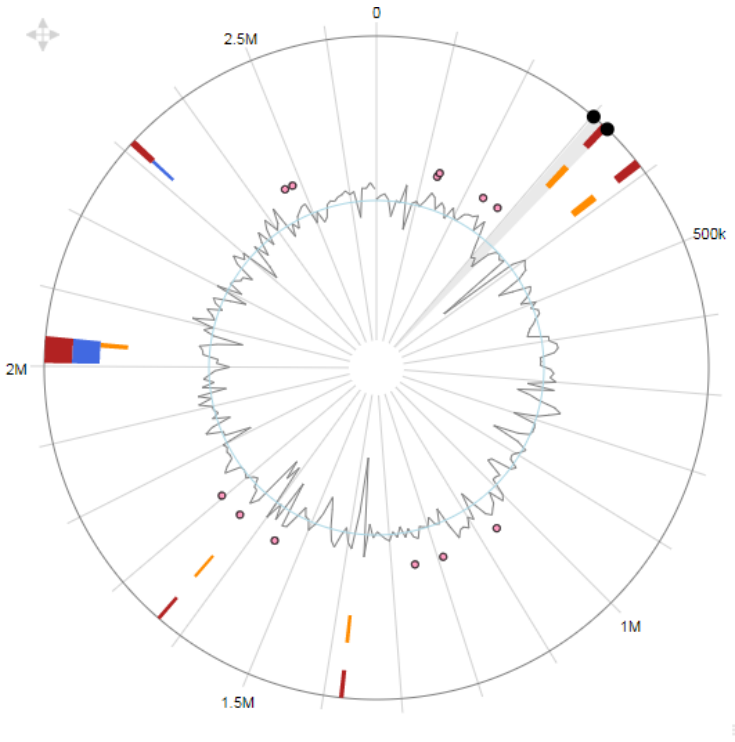

FIGURE 4.9: Genetic Islands in *Akkermansia muciniphila* ATCC BAA-835 as Predicted by Island Viewer.

This strain is selected as reference strain for further analysis. Table 4.6 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.10 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila* strains are human pathogens were predicted by Pathogenecity island viewer and no pathogenecity islands found in *Akkermansia muciniphila*. (http://bagel4.molgenrug.nl).

TABLE 4.6: Islands of *Akkermansia muciniphila* ATCC-BAA 835 Strain

Island start	**Island end**	**Length**	**Method**	**Gene name**	**Gene ID**	**Product**
378438	390250	11812	Predicted by at least one method	WP_012419377.1		ABC transporter ATP-binding protein
315805	326127	10322	Predicted by at least one method	WP_031930123.1		hypothetical protein
1374126	1379931	5805	Predicted by at least one method	WP_123038903.1		DUF2778 domain -containing protein
1632716	1637515	4799	Predicted by at least one method	WP_012420401.1		type II toxin-antitoxin system HicA family toxin
2004200	2038697	34497	Predicted by at least one method	WP_012420692.1	speA	biosynthetic arginine decarboxylase

TABLE 4.6: Islands of *Akkermansia muciniphila* ATCC-BAA 835 Strain

Island start	Island end	Length	Method	Gene name	Gene ID	Product
2004200	2038697	34497	Predicted by at least one method	WP_042448227.1		bifunctional adenosylcobinamide kinase/adenosylcobinamide-phosphate guanylyltransferase
2310076	2319099	9023	Predicted by at least one method	WP_012420939.1		iron-containing alcohol dehydrogenase

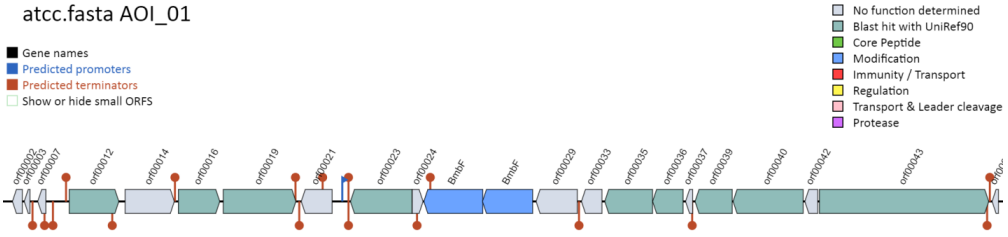

FIGURE 4.10: Bacteriocin of *Akkermansia muciniphila* ATCC- BAA 835 from BAGEL4

4.2.7 *Akkermansia muciniphila* CBA5201

Figure 4.11 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* CBA5201. Table 4.7 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.12 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila*

strains are human pathogens were predicted by Pathogenecity island viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

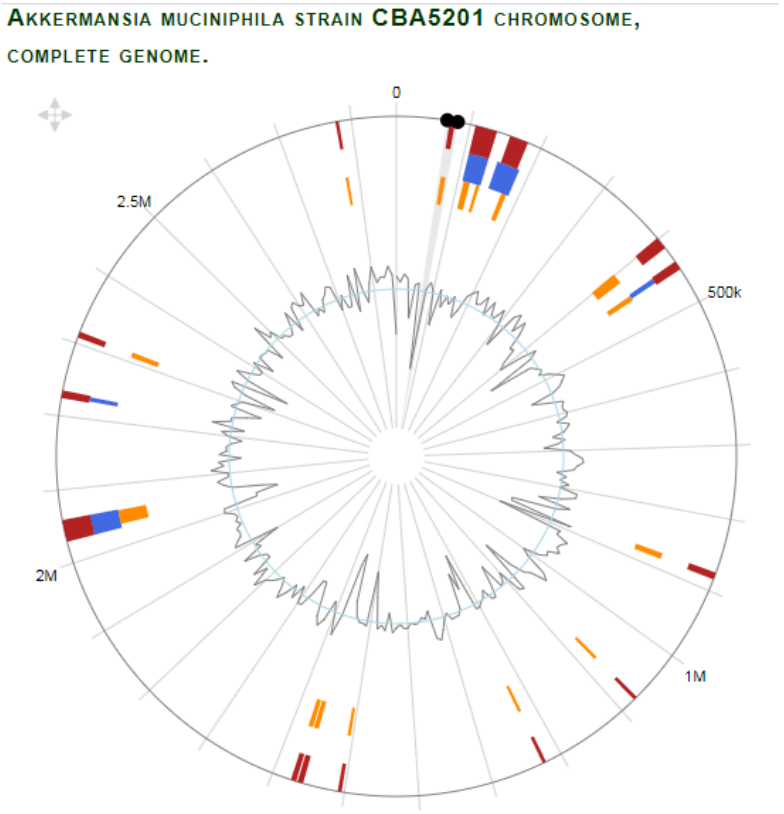

FIGURE 4.11: Genetic Islands in *Akkermansia muciniphila* CBA5201 as Predicted by Island Viewer.

TABLE 4.7: Islands of *Akkermansia muciniphila* CBA5201 Strain

Island start	Island end	Length	Gene name	Gene ID	Product
72057	79198	7141	WP_164917341.1		sugar-binding protein
107004	136880	29876	WP_102733850.1	tcdA	tRNA cyclic N6-threonylcarbamoyladenosine(37) synthase TcdA
130388	135551	5163	WP_102737849.1		VWA domain-containing protein

TABLE 4.7: Islands of *Akkermansia muciniphila* CBA5201 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
130388	135551	5163	WP_128804659.1		hypothetical protein
155490	179714	24224	WP_102743592.1		hypothetical protein
173058	181024	7966	WP_128804674.1		hypothetical protein
397828	415000	17172	WP_087393711.1		type II toxin-antitoxin system HicA family toxin
437222	448213	10991	WP_102737454.1		addiction module toxin, HicA family
437222	448213	10991	WP_102737591.1		YHYH domain-containing protein
441165	449809	8644	WP_164917364.1		hypothetical protein
441165	449809	8644	WP_102737449.1		hypothetical protein
441165	449809	8644	WP_102737450.1		hypothetical protein
441165	449809	8644	WP_102737452.1		hypothetical protein
441165	449809	8644	WP_128804703.1		hypothetical protein
441165	449809	8644	WP_102737454.1		addiction module toxin, HicA family
441165	449809	8644	WP_102737591.1		YHYH domain-containing protein

TABLE 4.7: Islands of *Akkermansia muciniphila* CBA5201 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
441165	449809	8644	WP_128804704.1		hypothetical protein
441165	449809	8644	WP_128220221.1		hypothetical protein
441165	449809	8644	WP_102737456.1		tyrosine-type recombinase/ integrase
441165	449809	8644	WP_161994078.1	ybeY	rRNA maturation RNase YbeY
441165	449809	8644	WP_102733694.1		HDIG domain-containing protein
874094	883636	9542	WP_046436576.1		hypothetical protein
1072668	1077975	5307	WP_128804751.1		restriction endonuclease subunit S
1509113	5186		WP_128804778.1		AKKM5201_RS06460
1503927	1509113	5186	WP_164917419.1		RHS repeat-associated core domain-containing protein
1556770	1564486	7716	WP_094140227.1		hypothetical protein
2029628	2059455	29827	WP_128804824.1		hypothetical protein

TABLE 4.7: Islands of *Akkermansia muciniphila* CBA5201 Strain

Island start	Island end	Length	Gene name	Gene ID	Product
2035747	2058180	22433	WP_046437171.1		hypothetical protein
2223865	2234230	10365	WP_081863700.1		substrate-binding domain-containing protein
2307928	2316277	8349	WP_102737558.1		master DNA invertase Mpi family serine-type recombinase
2778256	2782733	4477	WP_128804861.1		hypothetical protein

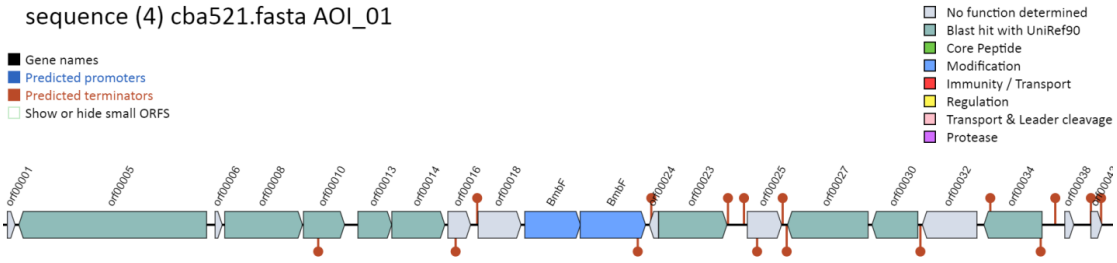

FIGURE 4.12: Bacteriocin of *Akkermansia muciniphila* CBA5201 from BAGEL4

4.2.8 *Akkermansia muciniphila* DSM 22959

Figure 4.13 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila*DSM 22959 .Table 4.8 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.14 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila* strains are

human pathogens were predicted by Pathogenecity island viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

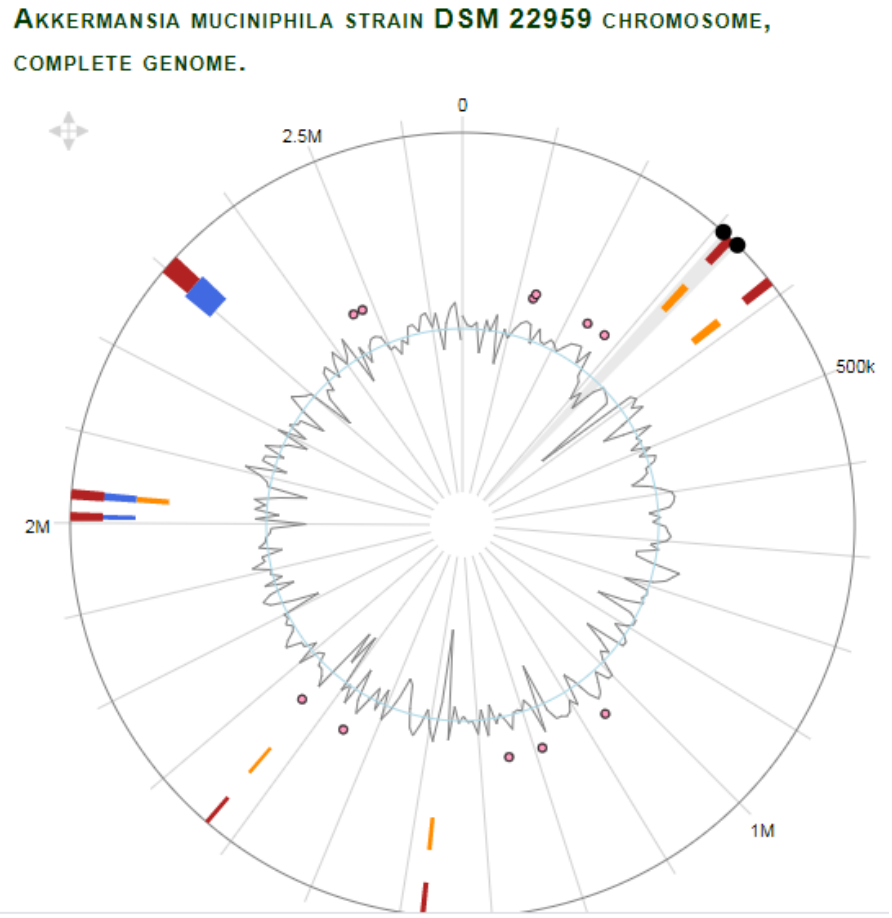

FIGURE 4.13: Genetic Islands in *Akkermansia muciniphila* DSM 22959 as Predicted by Island Viewer.

TABLE 4.8: Islands of *Akkermansia muciniphila* DSM22959 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
313861	324183	10322	WP_012419064.1	dnaA	chromosomal replication initiator protein DnaA
376494	387478	10984	WP_012419064.1	dnaA	chromosomal replication initiator protein DnaA

TABLE 4.8: Islands of *Akkermansia muciniphila* DSM22959 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
1372163	1377968	5805	WP_012419064.1	dnaA	chromosomal replication initiator protein DnaA
1630753	1635552	4799	WP_012419064.1	dnaA	chromosomal replication initiator protein DnaA
2002236	2011743	9507	WP_012419064.1	dnaA	chromosomal replication initiator protein DnaA
2002236	2011743	9507	WP_012420692.1	speA	chromosomal replication initiator protein DnaA
2025587	2036733	11146	WP_012420713.1	cobS	adenosyl cobinamide-GDP ribazolet ransferase
2025587	2036733	11146	WP_012420713.1	cobS	adenosyl cobinamide-GDP ribazolet ransferase
2027085	2034483	7398	WP_012419064.1	dnaA	chromosomal replication initiator protein DnaA

TABLE 4.8: Islands of *Akkermansia muciniphila* DSM22959 Strain

Island start	Island end	Length	Gene name	Gene ID	Product
2294595	2317121	22526	WP_012419064.1	dnaA	chromosomal replication initiator protein DnaA
2294595	2317121	22526	WP_012419064.1	dnaA	chromosomal replication initiator protein DnaA
2294595	2317121	22526	WP_012420923.1	murA	UDP-N-acetylglucosamine 1-carboxyvinyl transferase
2294595	2317121	22526	WP_012420925.1	aroC	chorismate synthase

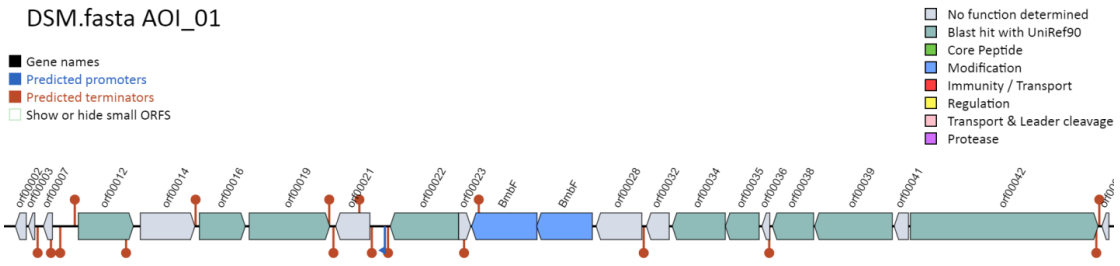

FIGURE 4.14: Bacteriocin of *Akkermansia muciniphila* DSM22959 from BAGEL

4.2.9 *Akkermansia muciniphila* JCM 30893

Figure 4.15 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* JCM 30893 Table 4.9 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.16 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila*

strains are human pathogens were predicted by Pathogenecity island viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

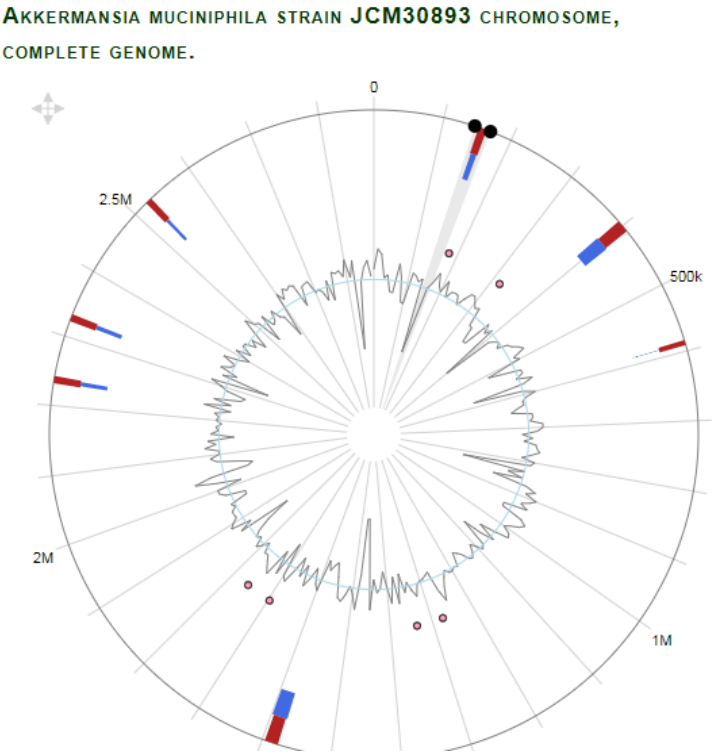

FIGURE 4.15: Genetic islands in *Akkermansia muciniphila*JCM 30893 as Predicted by Island Viewer.

TABLE 4.9: Islands of *Akkermansia muciniphila* JCM30893 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
150826	162569	11743	WP_102741725.1	thiS	sulfur carrier protein ThiS
391359	409573	18214	WP_162610405.1		HDIG domain-containing protein
584220	591615	7395	WP_065529667.1		competence/damage-inducible protein A
1576884	1596071	19187	WP_155844540.1		hypothetical protein
2231406	2241815	10409	WP_031931181.1	cobS	adenosylcobinamide-GDP ribazoletransferase

TABLE 4.9: Islands of *Akkermansia muciniphila* JCM30893 Strain

Island start	Island end	Length	Gene name	Gene ID	Product
2322132	2332660	10528	WP_094136305.1	infC	translation initiation factor IF-3
2522010	2531686	9676	WP_102743890.1	crcB	fluoride efflux transporter CrcB

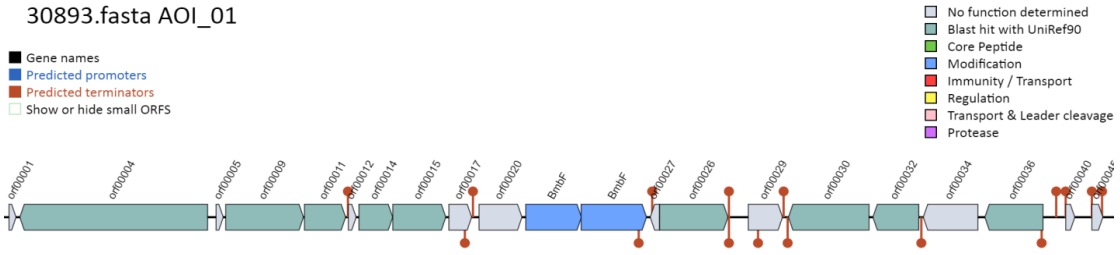

FIGURE 4.16: Bacteriocin of *Akkermansia muciniphila* JCM 30893 from BAGEL4

4.2.10 *Akkermansia muciniphila* AMDK-7

Figure 4.18 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-7. Table 4.10 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.17 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila*strains are human pathogens were predicted by Pathogenecityisland viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

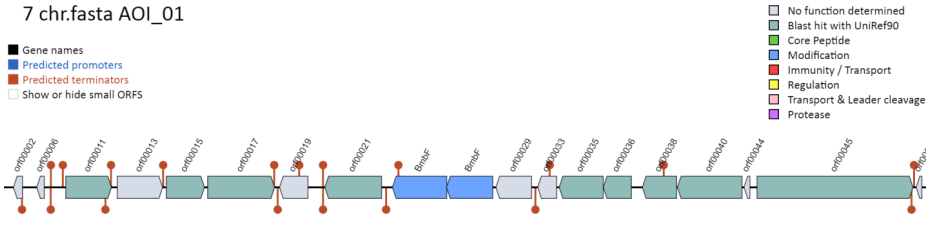

FIGURE 4.17: Bacteriocin of *Akkermansia muciniphila* AMBK-7 from BAGEL4

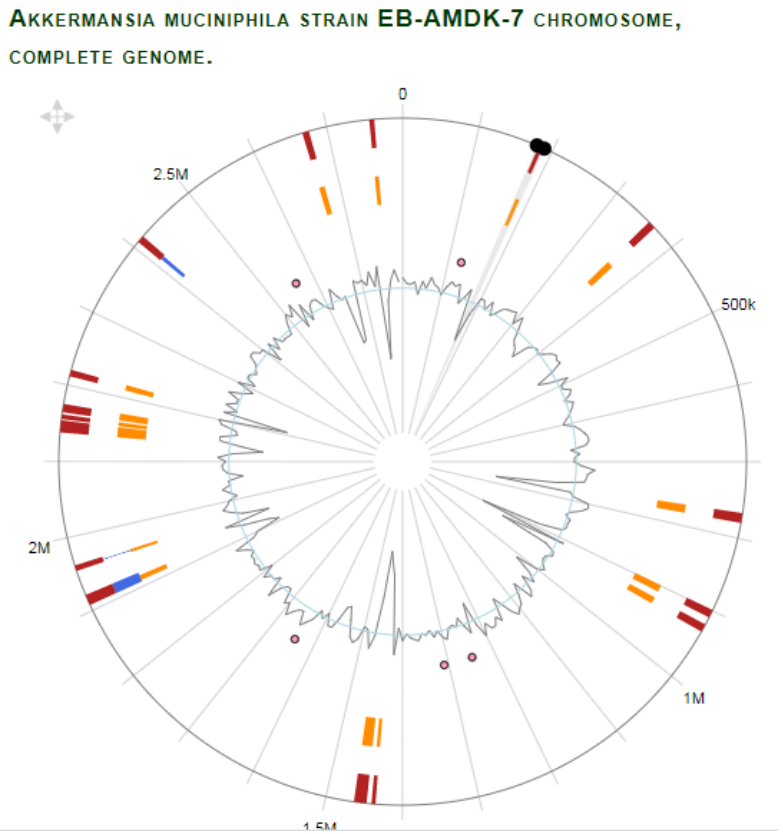

FIGURE 4.18: Genetic Islands in *Akkermansia muciniphila* AMDK-7 as Predicted by Island Viewer.

TABLE 4.10: Islands of *Akkermansia muciniphila* AMDK-7 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
181403	186601	5198	WP_12825 1927.1		DUF3800 domain-containing protein
355998	366318	10320	WP_01241 9320.1		transcriptional regulator
766899	779893	12994	WP_12825 2167.1		hypothetical protein
898913	909636	10723	WP_12825 2219.1		hypothetical protein
919872	930500	10628	WP_12825 2235.1		hypothetical protein

TABLE 4.10: Islands of *Akkermansia muciniphila* AMDK-7 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
1434911	1440184	5273	WP_128252423.1		RHS repeat-associated core domain-containing protein
1446538	1463048	16510	WP_081429137.1	kdpF	K(+)-transporting ATPase subunit F
1907827	1922076	14249	WP_128252627.1	grpE	nucleotide exchange factor GrpE
1910017	1917694	7677	WP_128252626.1		hypothetical protein
1955317	1962789	7472	WP_128252649.1	tsaB	tRNA (adenosine (37)-N6)-threonylcarbamoyltransferase complex dimerization subunit type 1 TsaB
2137570	2153386	15816	WP_128252712.1		hypothetical protein
2155024	2161149	6125	WP_128252713.1		hypothetical protein
2211592	2220164	8572	WP_128252731.1		DEAD/DEAH box helicase family protein
2408658	2418323	9665	WP_065529196.1	crcB	fluoride efflux transporter CrcB

TABLE 4.10: Islands of *Akkermansia muciniphila* AMDK-7 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
2667482	2676323	8841			acyltransferase family protein
2756084	2762533	6449	WP_094137864.1		hypothetical protein

4.2.11 *Akkermansia muciniphila* AMDK-8

Figure 4.19 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-8. Table 4.11 Summarizes the Details of Genomic Islands and Genes present in Respective Island. Figure 4.20 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila* strains are human pathogens were predicted by Pathogenecity island viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

TABLE 4.11: Islands of *Akkermansia muciniphila* AMDK-8 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
39377	56879	17502	WP_102732764.1		hypothetical protein
104973	113039	8066	WP_128153682.1		hypothetical protein
128901	136480	7579	WP_046436036.1		hypothetical protein
280462	287107	6645	WP_128220153.1		hypothetical protein

TABLE 4.11: Islands of *Akkermansia muciniphila* AMDK-8 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
371830	384399	12569	WP_102738840.1		trypsin-like peptidase domain-containing protein
405911	420594	14683	WP_102741300.1		site-specific integrase
415191	421375	6184	WP_102733694.1		HDIG domain-containing protein
897377	904708	7331	WP_094140696.1		IS1595 family transposase
1113796	1118641	4845	WP_128220348.1		type I restriction endonuclease subunit R
1501124	1506979	5855	WP_081429137.1	kdpF	K(+)-transporting ATPase subunit F
1621582	1636409	14827	WP_094135541.1		2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase
2102810	2109501	6691	WP_102735124.1		hypothetical protein
2686461	2690938	4477	WP_128157668.1		hypothetical protein
2771676	2788330	16654	WP_128220805.1		Txe/YoeB family addiction module toxin
2784103	2790038	5935	WP_128220750.1		hypothetical protein

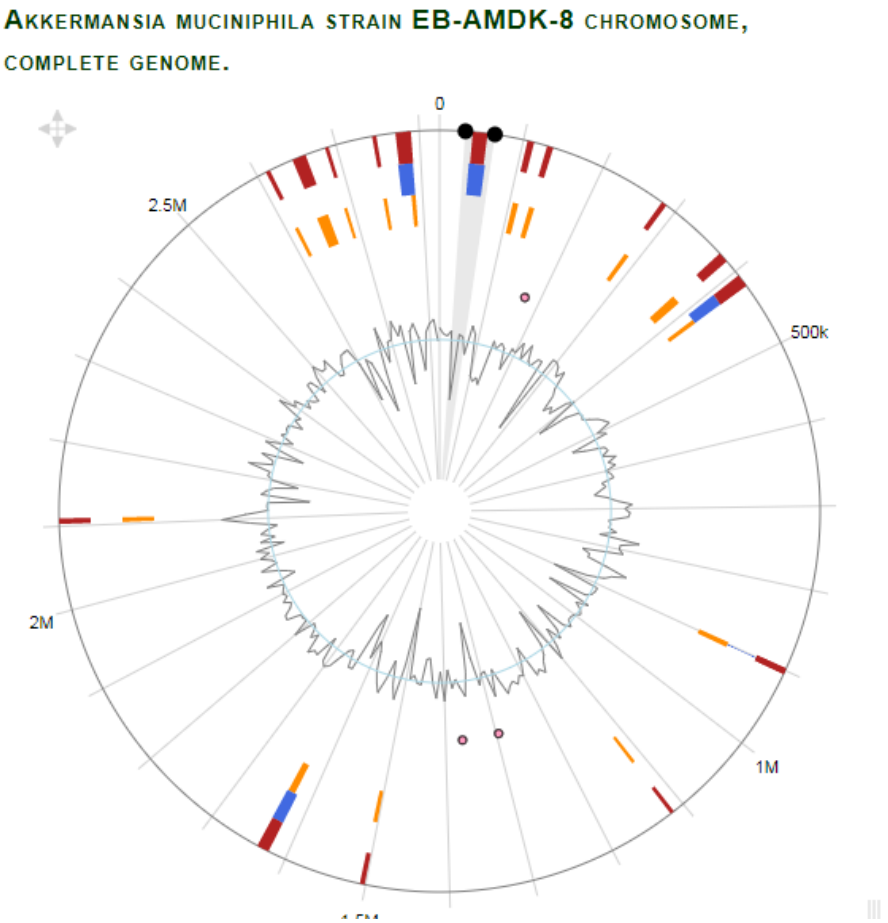

FIGURE 4.19: Genetic Islands in *Akkermansia muciniphila* AMDK-8.as Predicted by Island Viewer.

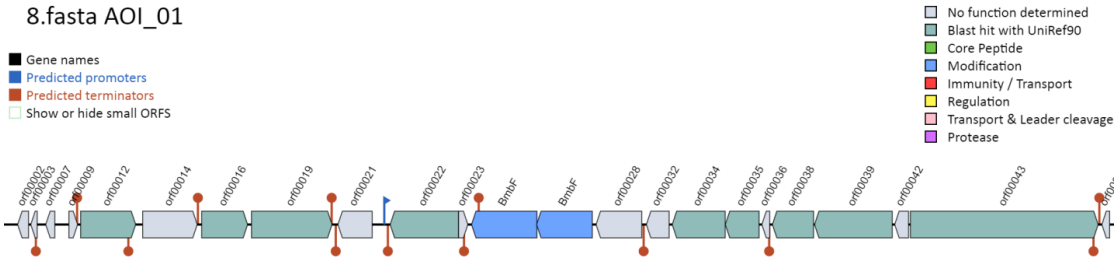

FIGURE 4.20: Bacteriocin of *Akkermansia muciniphila* AMBK-8 from BAGEL4

4.2.12 *Akkermansia muciniphila* AMDK-10

Figure 4.21 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-10. Table 4.12 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.22 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila*

strains are human pathogens were predicted by Pathogenecityisland viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

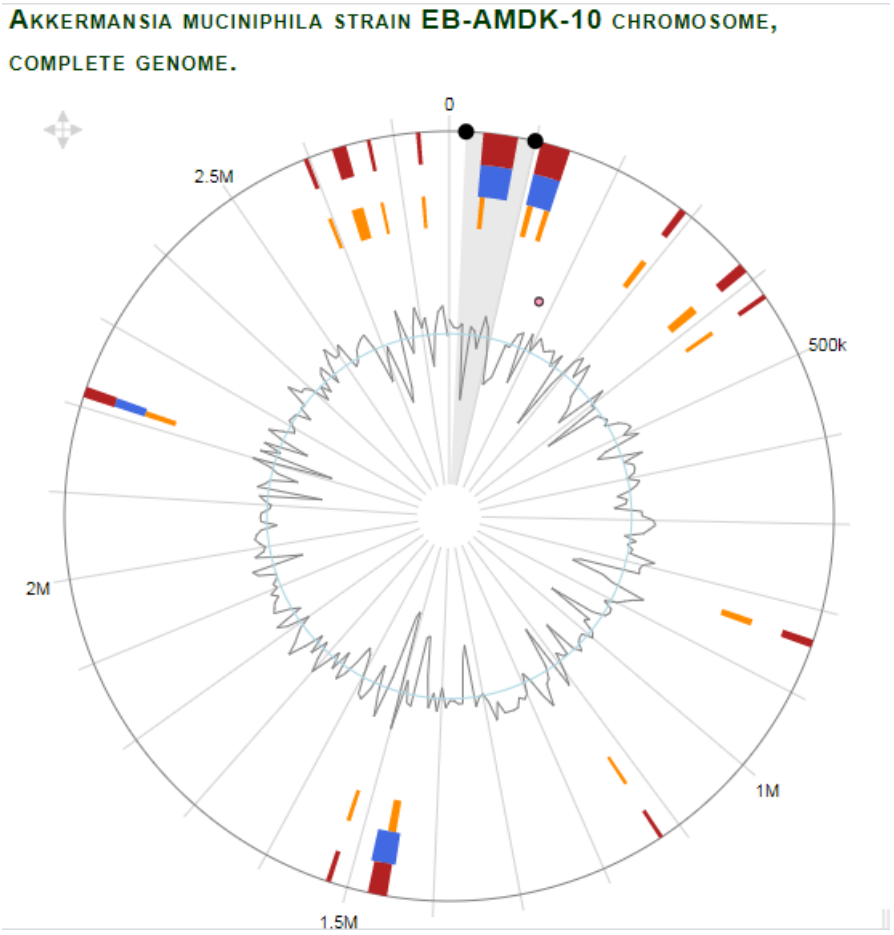

FIGURE 4.21: Genetic Islands in *Akkermansia muciniphila*AMDK-10 as Predicted by Island Viewer.

TABLE 4.12: Islands of *Akkermansia muciniphila* AMDK-10 Strain

Island start	Island end	Length	Gene name	Gene ID	Product
39306	79281	39975	WP_022197161.1	glsA	glutaminase A
42251	49038	6787	WP_128153368.1		PepSY domain-containing protein
106304	140674	34370	WP_128153382.1		hypothetical protein

TABLE 4.12: Islands of *Akkermansia muciniphila* AMDK-10 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
108575	116639	8064	WP_128153682.1		hypothetical protein
132786	140031	7245	WP_046436036.1		hypothetical protein
283478	291928	8450	WP_046434858.1		hypothetical protein
376710	388739	12029	WP_128190976.1		hypothetical protein
420636	426433	5797	WP_022198322.1	ybeY	rRNA maturation RNase YbeY
420636	426433	5797	WP_102733694.1		HDIG domain-containing protein
1122546	1127355	4809	WP_123044031.1		type I restriction endonuclease subunit R
1453447	1475779	22332	WP_123044062.1	ilvD	dihydroxy-acid dehydratase
1455174	1466546	11372	WP_102733457.1		hypothetical protein
1455174	1466546	11372	WP_102733459.1		hypothetical protein
1455174	1466546	11372	WP_128153481.1		hypothetical protein
1455174	1466546	11372	WP_128153482.1		hypothetical protein

TABLE 4.12: Islands of *Akkermansia muciniphila* AMDK-10 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
1455174	1466546	11372	WP_123044 063.1		hypothetical protein
1455174	1466546	11372	WP_128153 483.1		hypothetical protein
1455174	1466546	11372	WP_128153 484.1		hypothetical protein
1455174	1466546	11372	WP_123044 064.1		hypothetical protein
1455174	1466546	11372	WP_10273 4350.1		IS1595 family transposase
1455174	1466546	11372	WP_10273 4349.1		outer membrane beta-barrel protein
1455174	1466546	11372	WP_10273 4348.1		hypothetical protein
1455174	1466546	11372	WP_102734 347.1		RHS repeat-associated core domain-containing protein
1455174	1466546	11372	WP_10273 4346.1		hypothetical protein
1455174	1466546	11372	WP_128153 485.1		hypothetical protein

TABLE 4.12: Islands of *Akkermansia muciniphila* AMDK-10 Strain

Island start	Island end	Length	Gene name	Gene ID	Product
1519694	1525504	5810	WP_123038903.1		DUF2778 domain-containing protein
2211213	2223743	12530	WP_123044222.1	coaD	pantetheine-phosphate adenylyl transferase
2212243	2219999	7756	WP_128153571.1		hypothetical protein
2593118	2598330	5212	WP_094136088.1	cas2	CRISPR-associated endonuclease Cas2
2668514	2672991	4477	WP_128153561.1		hypothetical protein
2725796	2730878	5082	WP_046437351.1	lepB	signal peptidase I

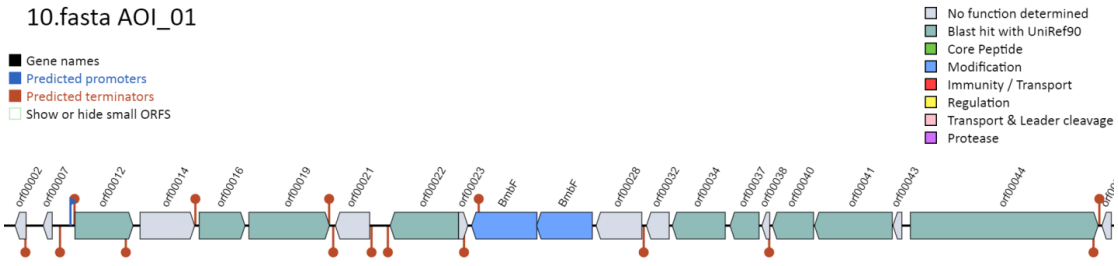

FIGURE 4.22: BAGEL Results

4.2.13 *Akkermansia muciniphila* AMDK-11

Figure 4.23 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-11. Table 4.13 Summarizes the Details of Genomic Islands and Genes

present in Respective Island. Figure 4.24 Indicates the Bacteriocin Producing genes predicted by BAGEL. The probabilities that *Akkermansia muciniphila*strains are human pathogens were predicted by Pathogenecityisland viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

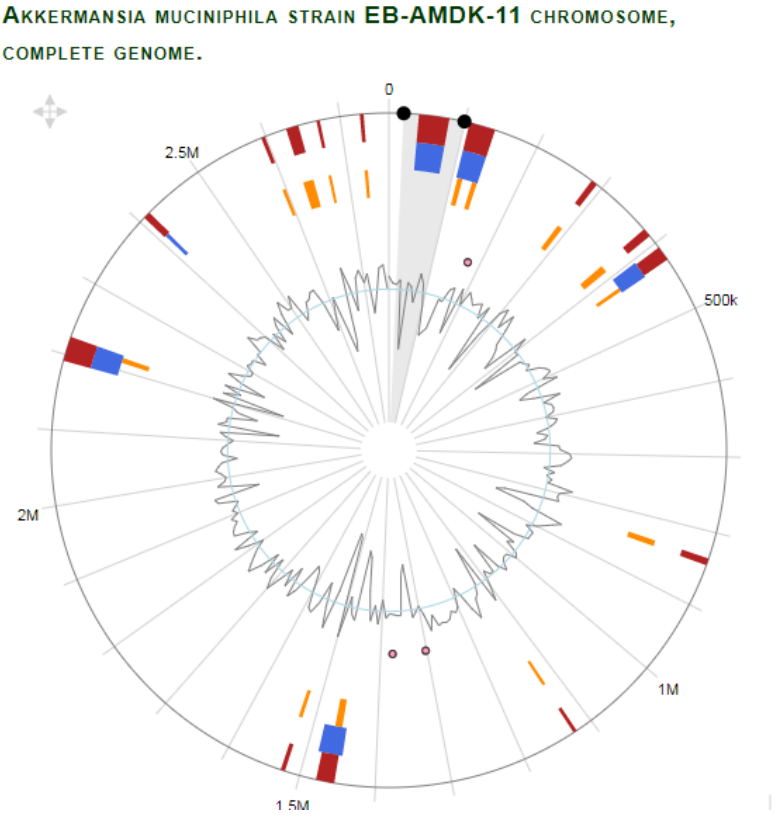

FIGURE 4.23: Genetic Islands in *Akkermansia muciniphila*AMDK-11 as Predicted by Island Viewer.

TABLE 4.13: Islands of *Akkermansia muciniphila* AMDK-11 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
39309	79291	39982	WP_10273 2764.1		hypothetical protein
106311	140689	34378	WP_12825 1664.1		tyrosine-type recombinase/ integrase
108588	116653	8065	WP_10273 3853.1		SEL1-like repeat protein

TABLE 4.13: Islands of *Akkermansia muciniphila* AMDK-11 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
132468	140046	7578	WP_094139775.1		HipA domain-containing protein
283506	291956	8450	WP_123043932.1		RHS repeat-associated core domain-containing protein
376742	388355	11613	WP_012419320.1		transcriptional regulator
407271	426477	19206	WP_102733694.1		HDIG domain-containing protein
420676	426477	5801	WP_102732238.1		AAA family ATPase
1122685	1127494	4809	WP_123044031.1		type I restriction endonuclease subunit R
1453693	1478015	24322	WP_102734347.1		RHS repeat-associated core domain-containing protein
1455420	1466777	11357	WP_102734347.1		RHS repeat-associated core domain-containing protein

TABLE 4.13: Islands of *Akkermansia muciniphila* AMDK-11 Strain

Island start	Island end	Length	Gene name	Gene ID	Product
1519943	1525753	5810	WP_08142 9137.1	kdpF	K(+)-trans porting ATPase subunit F
2191442	2224114	32672	WP_102733 478.1	xseA	exodeoxy ribonuclease VII large subunit
2593548	2598761	5213	WP_10273 4057.1	cas1c	type I-C CRISPR-associated endonuc lease Cas1
2628065	2644673	16608	WP_0221 97099.1		Glycosyl transferase

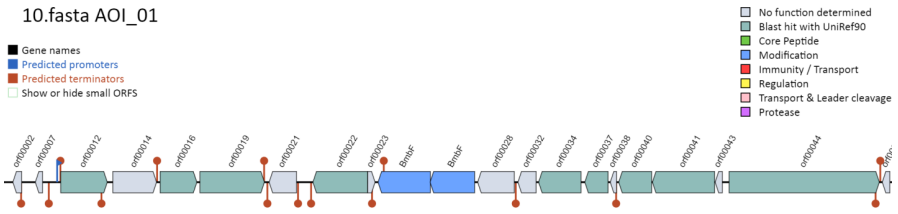

FIGURE 4.24: Bacteriocin of *Akkermansia muciniphila* AMBK-11 from BAGEL4

4.2.14 *Akkermansia muciniphila* AMDK-12

Figure 4.25 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-12. Table 4.14 Summarizes the Details of Genomic Islands and Genes

Present in Respective Island. Figure 4.26 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila*strains are human pathogens were predicted by Pathogenecityisland viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

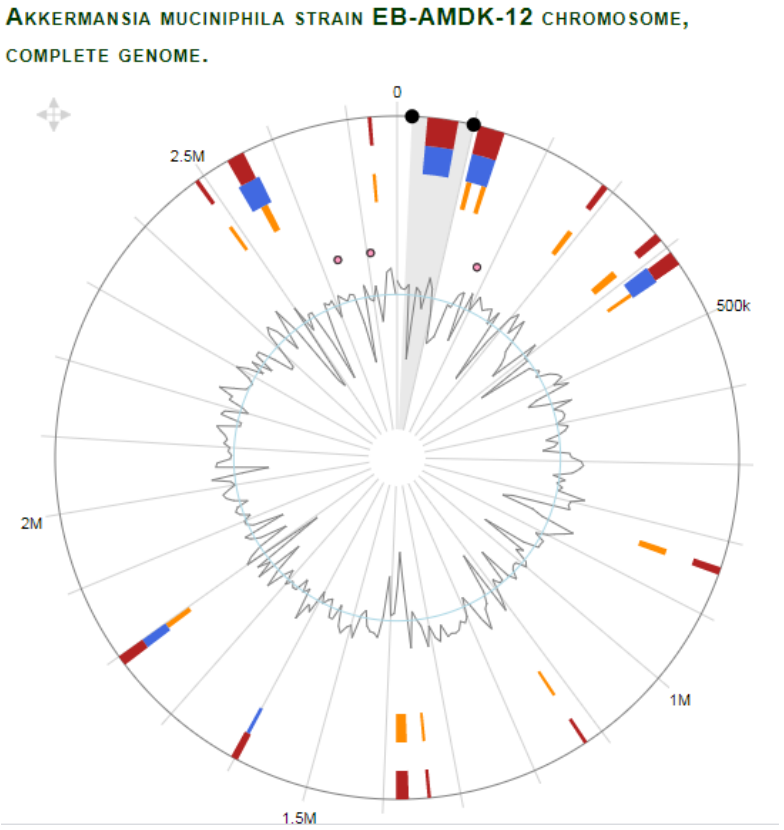

FIGURE 4.25: Genetic Islands in *Akkermansia muciniphila*AMDK-12 as Predicted by Island Viewer.

TABLE 4.14: Islands of *Akkermansia muciniphila* AMDK-12 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
39309	79351	40042	WP_102732764.1		hypothetical protein
39309	79351	40042	WP_123044205.1		RHS repeat-associated core domain-containing protein

TABLE 4.14: Islands of *Akkermansia muciniphila* AMDK-12 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
132468	140046	7578	WP_046436036.1		hypothetical protein
283508	291958	8450	WP_123043937.1		hypothetical protein
376744	388357	11613	WP_094140769.1		hypothetical protein
407273	426479	19206	WP_128153412.1		hypothetical protein
420678	426479	5801	WP_102733694.1		hypothetical protein
834954	844973	10019	WP_128153446.1		hypothetical protein
1122691	1127500	4809	WP_123044031.1		hypothetical protein
1338363	1342840	4477	WP_128153561.1		hypothetical protein
1367126	1383734	16608	WP_102734081.1		hypothetical protein
1595641	1605183	9542	WP_012420931.1		hypothetical protein
1786737	1800195	13458	WP_123044222.1	coaD	pantetheine-phosphate adenylyl transferase
1790914	1799165	8251	WP_128153532.1		restriction endonuclease subunit S

TABLE 4.14: Islands of *Akkermansia muciniphila* AMDK-12 Strain

Island start	Island end	Length	Gene name	Gene ID	Product
2486025	2491835	5810	WP_123038903.1		DUF2778 domain-containing protein
2535753	2559006	23253	WP_102734350.1		IS1595 family transposase
2545001	2556358	11357	WP_128153482.1		hypothetical protein
2726244	2731326	5082	WP_046437351.1	lepB	signal peptidase I

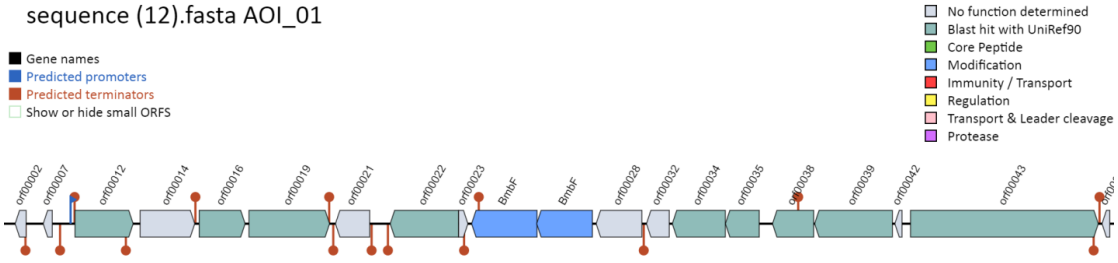

FIGURE 4.26: Bacteriocin of *Akkermansia muciniphila* AMBK-12 from BAGEL4

4.2.15 *Akkermansia muciniphila* AMDK-13

Figure 4.27 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-13 . Table 4.15 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.28 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila* strains are human pathogens were predicted by Pathogenecity island viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

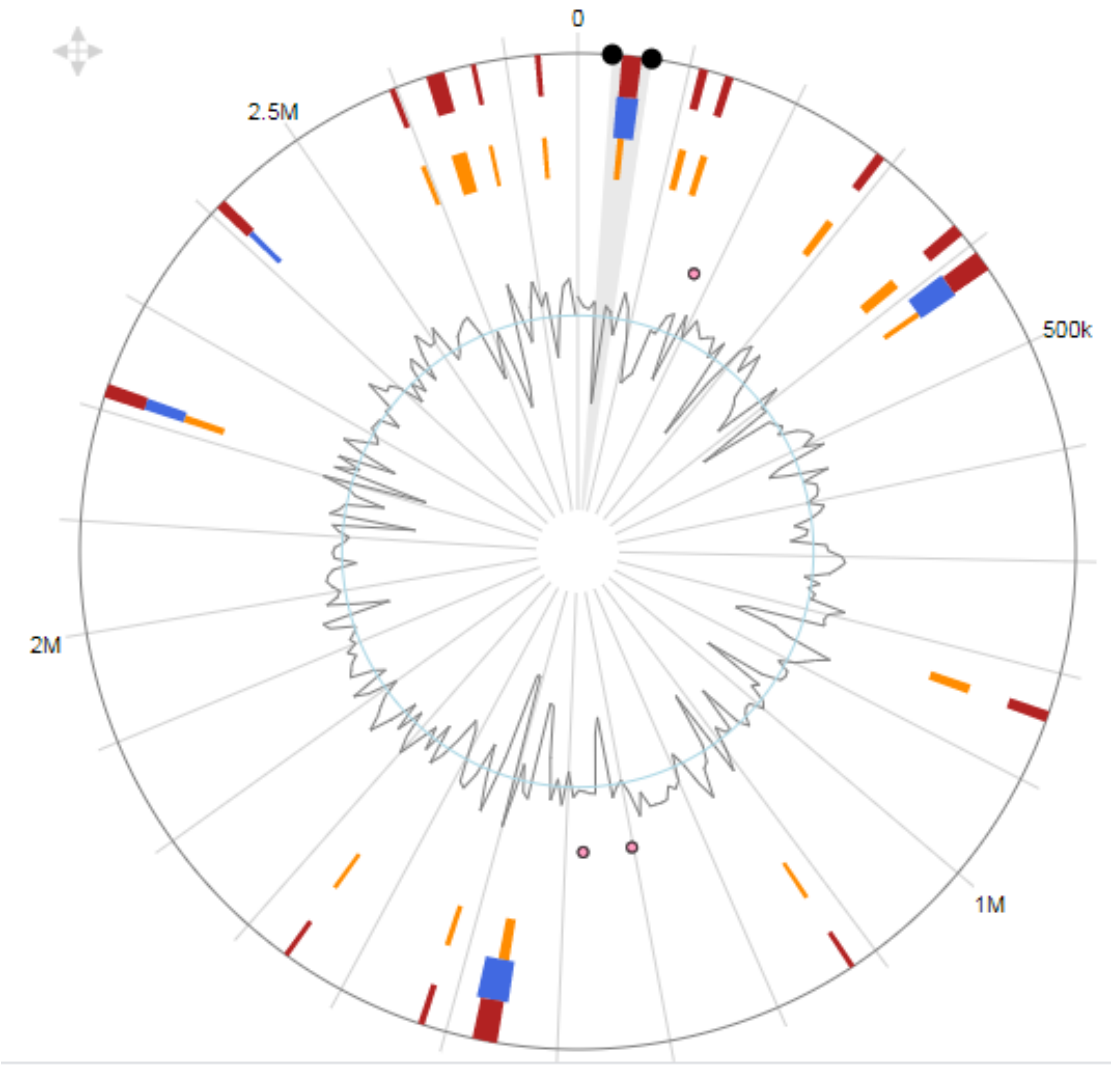

FIGURE 4.27: Genetic Islands in *Akkermansia muciniphila* AMDK-13 as Predicted by Island Viewer.

TABLE 4.15: Genetic islands in *Akkermansia muciniphila* AMDK-13

Island start	Island end	Length	Gene name	Gene ID	Product
39319	56816	17497	WP_022197161.1	glsA	glutaminase A
42264	49050	6786	WP_102732766.1		IS3 family transposase

TABLE 4.15: Genetic islands in *Akkermansia muciniphila* AMDK-13

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
108585	116624	8039	WP_128153378.1		hypothetical protein
132463	140040	7577	WP_094139775.1		HipA domain-containing protein
283486	291936	8450	WP_123043937.1		hypothetical protein
376710	388323	11613	WP_094140769.1		hypothetical protein
407723	426422	18699	WP_161994078.1	ybeY	rRNA maturation RNase YbeY
420642	426422	5780	WP_161994078.1	ybeY	rRNA maturation RNase YbeY
834793	844877	10084	WP_128153446.1		hypothetical protein
1122574	1127383	4809	WP_123044029.1		restriction endonuclease subunit S
1453540	1475868	22328	WP_123044062.1	ilvD	dihydroxy-acid dehydratase
1455267	1466621	11354	WP_102734345.1		hypothetical protein
1519852	1525595	5743	WP_081429137.1	kdpF	K(+)-transporting ATPase subunit F

TABLE 4.15: Genetic islands in *Akkermansia muciniphila* AMDK-13

Island start	Island end	Length	Gene name	Gene ID	Product
1654823	1659553	4730	WP_1809 71958.1		hypothetical protein
2211342	2223872	12530	WP_1230 44222.1	coaD	pantetheine-phosphate adenylyltransferase
2212372	2220128	7756	WP_12815 3571.1		restriction endonuclease subunit S
2407470	2417117	9647	WP_022197 217.1	crcB	fluoride efflux transporter CrcB
2593213	2598426	5213	WP_09413 6088.1	cas2	CRISPR-associated endonuclease Cas2
2627727	2644333	16606	WP_1027 34081.1		glycosyl transferase family 2 protein
2668618	2673095	4477	WP_09413 9728.1		hypothetical protein
2725915	2730997	5082	WP_04643 7351.1	lepB	signal peptidase I

4.2.16 *Akkermansia muciniphila* AMDK-14

Figure 4.29 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-14. Table 4.16 Summarizes the Details of Genomic Islands and Genes

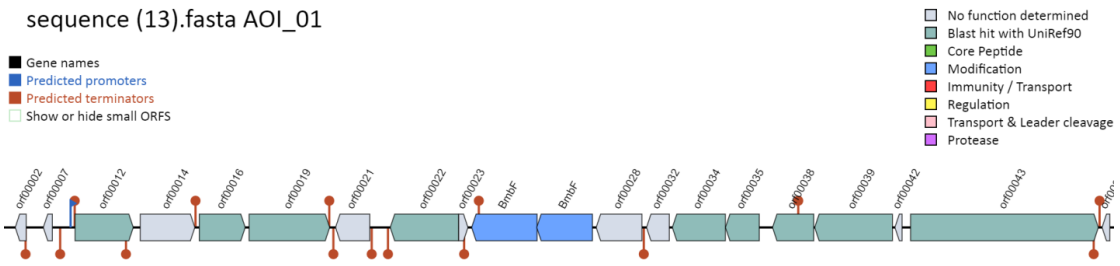

FIGURE 4.28: Bacteriocin of *Akkermansia muciniphila* AMBK-13 from BAGEL4

Present in Respective Island. Figure 4.30 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila*strains are human pathogens were predicted by Pathogenecityisland viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

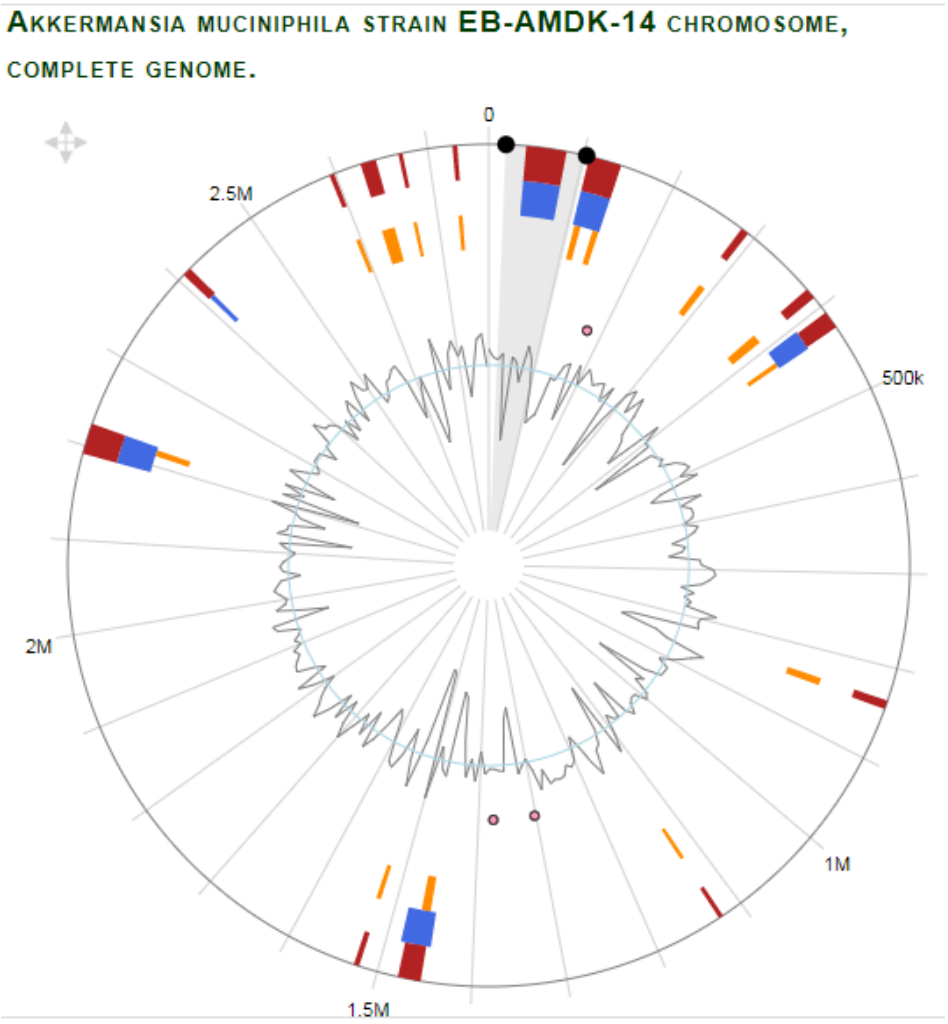

FIGURE 4.29: Genetic Islands in *Akkermansia muciniphila* AMDK-14 as Predicted by Island Viewer.

TABLE 4.16: Islands of *Akkermansia muciniphila* AMDK-14 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
39308	81891	42583	WP_102732764.1		hypothetical protein
108585	116624	8039	WP_128153378.1		hypothetical protein
132463	140041	7578	WP_046436036.1		hypothetical protein
283488	291950	8462	WP_102734797.1		hypothetical protein
376735	388348	11613	WP_094140769.1		hypothetical protein
407749	426469	18720	WP_102733694.1		HDIG domain-containing protein
420668	426469	5801	WP_102732238.1		AAA family ATPase
834926	844207	9281	WP_123043999.1		hypothetical protein
1453637	1477959	24322	WP_128153485.1		hypothetical protein
1455364	1466736	11372	WP_128153485.1		hypothetical protein
1519887	1525696	5809	WP_123038903.1		DUF2778 domain-containing protein
2191369	2224040	32671	WP_123044222.1	coaD	pantetheine-phosphate adenylyl transferase
2212540	2220296	7756	WP_128153571.1		hypothetical protein
2407635	2417283	9648	WP_012420932.1		hypothetical protein
2593428	2598641	5213	WP_102734060.1	cas5c	type I-C CRISPR -associated protein Cas5
2627944	2644552	16608	WP_102734081.1		glycosyltransferase family 2 protein
2668837	2673314	4477	WP_128157668.1		hypothetical protein
2726135	2731217	5082	WP_046437351.1	lepB	signal peptidase I

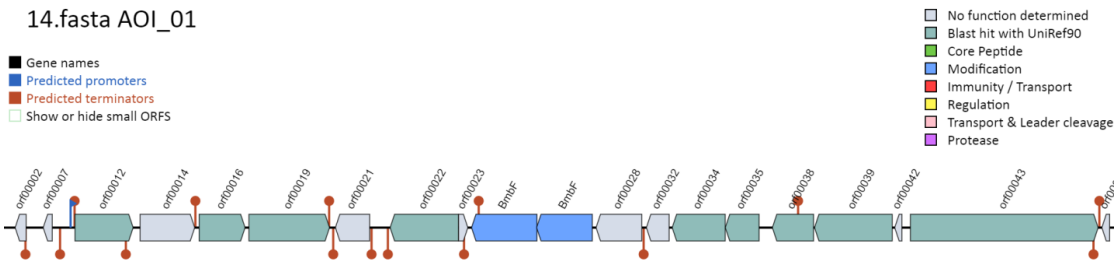

FIGURE 4.30: Bacteriocin of *Akkermansia muciniphila* AMBK-14 from BAGEL4

4.2.17 *Akkermansia muciniphila* AMDK-15

Figure 4.31 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-15.Table 4.17 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.32 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila* strains are human pathogens were predicted by Pathogenecityisland viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

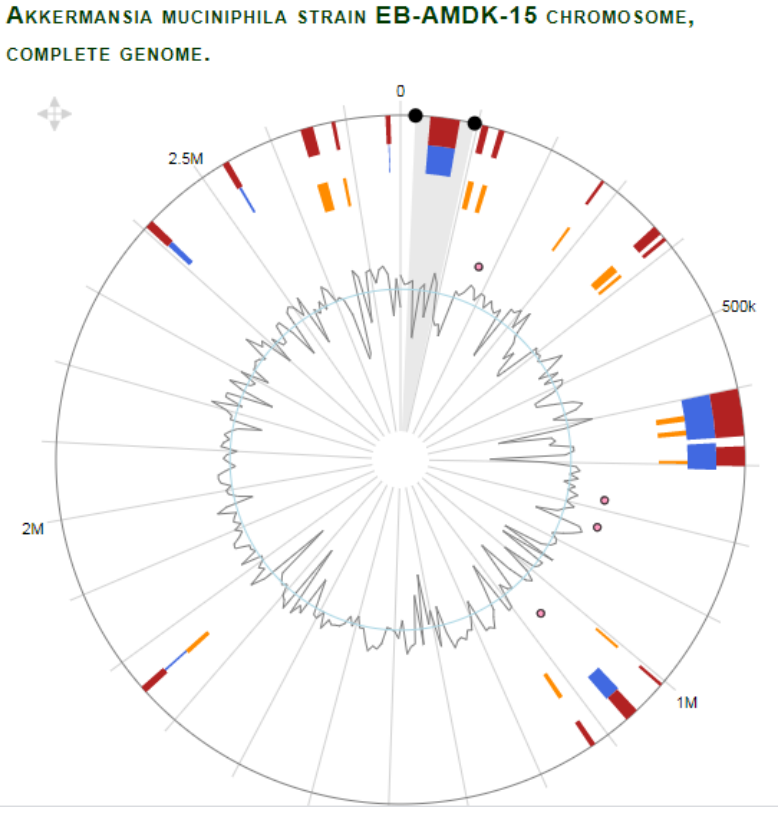

FIGURE 4.31: Genetic Islands in *Akkermansia muciniphila* AMDK-15 as Predicted by Island Viewer.

TABLE 4.17: Islands of *Akkermansia muciniphila* AMDK-15 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
38405	76617	38212	WP_128192695.1		hypothetical protein
105416	114344	8928	WP_102732793.1		hypothetical protein
128034	135611	7577	WP_102732799.1		HipA domain-containing protein
275145	279399	4254	WP_102732994.1		hypothetical protein
364859	377709	12850	WP_128154364.1		hypothetical protein
383419	388446	5027	WP_022198860.1		hypothetical protein
600713	662294	61581	WP_102733050.1		hypothetical protein
624858	633757	8899	WP_081429137.1	kdpF	K(+)-transporting ATPase subunit F
647468	655329	7861	WP_102732964.1		VWA domain-containing protein
675204	700876	25672	WP_128154479.1		hypothetical protein
694083	699127	5044	WP_128154497.1		hypothetical protein

TABLE 4.17: Islands of *Akkermansia muciniphila* AMDK-15 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
1004560	1008762	4202	WP_012419831.1		ParB-like nuclease domain-containing protein
1051542	1069509	17967	WP_022198166.1	dnaN	DNA polymerase III subunit beta
1119931	1127011	7080	WP_128154681.1		type I restriction endonuclease subunit R
1750974	1757973	6999	WP_022198314.1		DUF2971 domain-containing protein
1751755	1760185	8430	WP_022198321.1		site-specific integrase
2404170	2415226	11056	WP_102732654.1		carbohydrate kinase
2530887	2539578	8691	WP_128155289.1		potassium channel family protein
2640085	2656692	16607	WP_022197102.1		glycosyl transferase family 2 protein
2750864	2757136	6272	WP_102732736.1		recombinase family protein

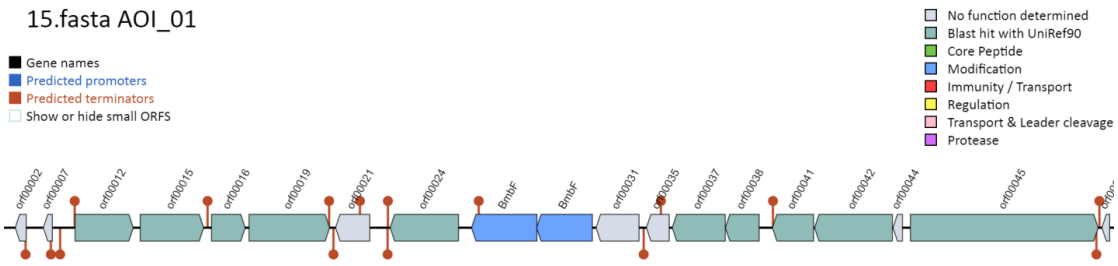

FIGURE 4.32: Bacteriocin of *Akkermansia muciniphila* AMBK-15 from BAGEL4.

4.2.18 *Akkermansia muciniphila* AMDK-16

Figure 4.33 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-16. Table 4.18 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.34 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila* strains are human pathogens were predicted by Pathogenecity island viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

TABLE 4.18: Islands of *Akkermansia muciniphila* AMDK-16 Strain

Island start	Island end	Length	Gene name	Gene ID	Product
38375	54441	16066	WP_145963244.1		hypothetical protein
103148	136572	33424	WP_128154229.1		hypothetical protein
105425	114353	8928	WP_102732793.1		hypothetical protein
128043	135620	7577	WP_128154226.1		hypothetical protein
275152	279815	4663	WP_102732991.1		hypothetical protein
363899	377697	13798	WP_145963247.1		hypothetical protein

TABLE 4.18: Islands of *Akkermansia muciniphila* AMDK-16 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
383407	388396	4989	WP_022198856.1		tape measure protein
600673	662255	61582	WP_123044217.1		hypothetical protein
624819	633718	8899	WP_081429137.1	kdpF	K(+)-transporting ATPase subunit F
647429	655290	7861	WP_102732964.1		VWA domain-containing protein
675166	700838	25672	WP_022196927.1	ilvD	dihydroxy-acid dehydratase
694045	699089	5044	WP_128154497.1		hypothetical protein
1004479	1008681	4202	WP_012419831.1		ParB-like nuclease domain-containing protein
1051444	1069411	17967	WP_022198166.1	dnaN	DNA polymerase III subunit beta
1119835	1126915	7080	WP_128154681.1		type I restriction endonuclease subunit R
1750915	1757914	6999	WP_022198323.1		HDIG domain-containing protein
2404128	2415184	11056	WP_102732654.1		carbohydrate kinase
2530844	2539535	8691	WP_128155289.1		potassium channel family protein
2640043	2656651	16608	WP_022197102.1		glycosyltransferase family 2 protein
2680926	2685402	4476	WP_102735771.1		hypothetical protein
2750830	2757603	6773	WP_102732737.1		hypothetical protein

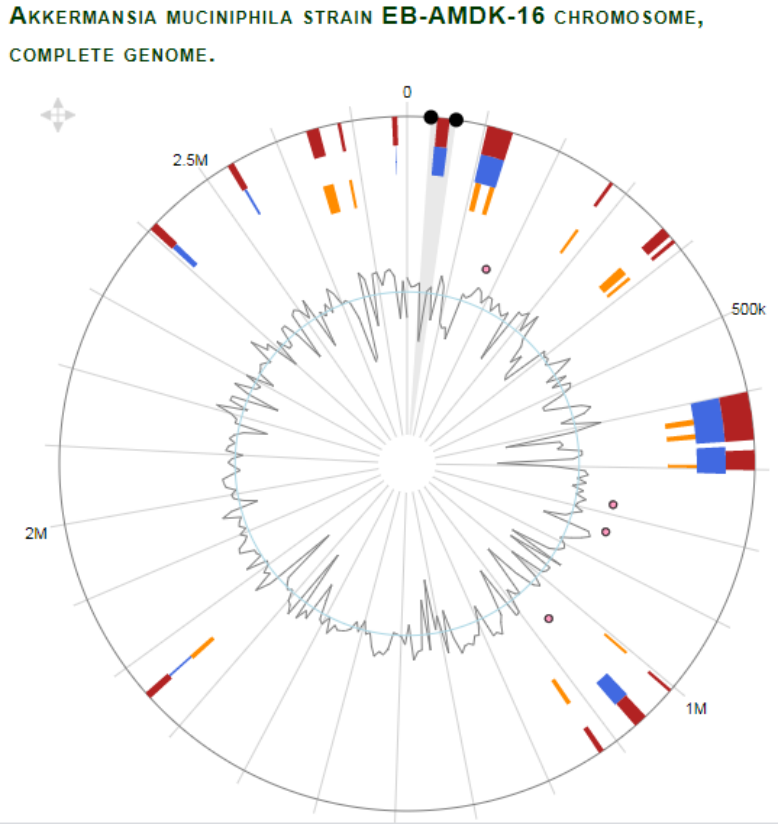

FIGURE 4.33: Genetic Islands in *Akkermansia muciniphila* AMDK-16 as Predicted by Island Viewer.

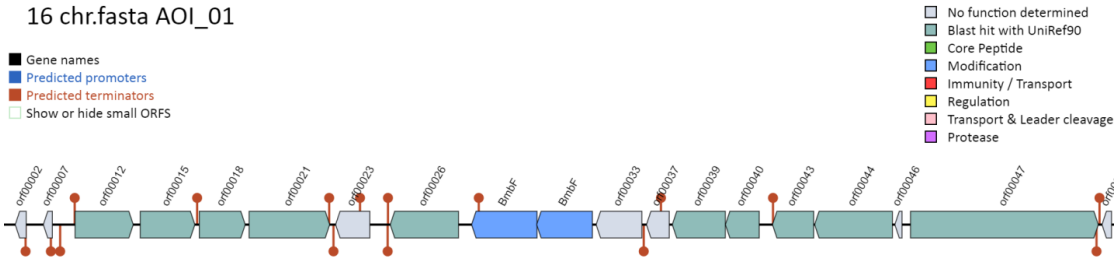

FIGURE 4.34: Bacteriocin of *Akkermansia muciniphila* AMBK-16 from BAGEL4

4.2.19 *Akkermansia muciniphila* AMDK-17

Figure 4.35 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-17. Table 4.19 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.36 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila*

strains are human pathogens were predicted by Pathogenecityisland viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

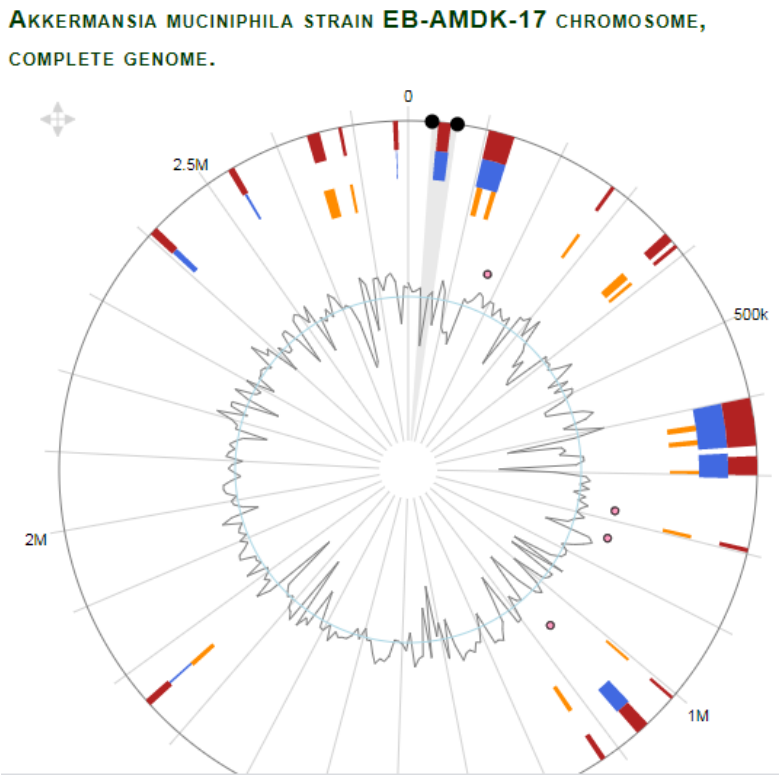

FIGURE 4.35: Genetic Islands in *Akkermansia muciniphila* AMDK-17 as Predicted by Island Viewer.

TABLE 4.19: Genetic Islands in *Akkermansia muciniphila* AMDK-17

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
315805	326127	10322	WP_031930123.1		hypothetical protein
378438	390250	11812	WP_012419377.1		ABC transporter ATP-binding protein
1374126	1379931	5805	WP_123038903.1		DUF2778 domain-containing protein
1632716	1637515	4799	WP_012420401.1		type II toxin-antitoxin system HicA family toxin

TABLE 4.19: Genetic Islands in *Akkermansia muciniphila* AMDK-17

Island start	Island end	Length	Gene name	Gene ID	Product
2004200	2038697	34497	WP_012420692.1	speA	biosynthetic arginine decarboxylase
2004200	2038697	34497	WP_042448227.1		bifunctional adenosylcobinamide kinase/adenosyl cobinamide-phosphate guanylyltransferase
2310076	2319099	9023	WP_012420939.1		iron-containing alcohol dehydrogenase

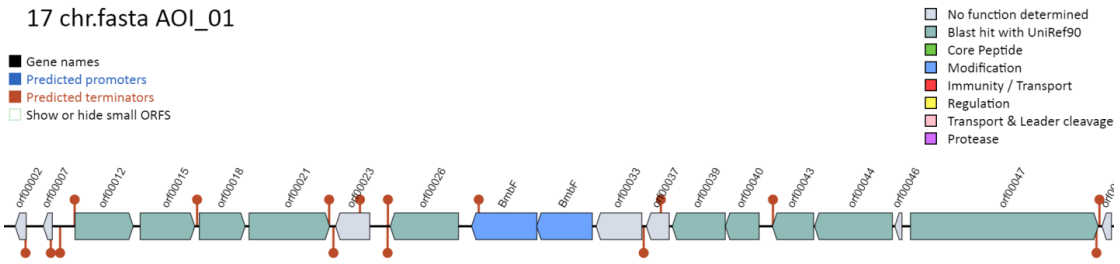

FIGURE 4.36: Bacteriocin of *Akkermansia muciniphila* AMBK-17 from BAGEL4

4.2.20 *Akkermansia muciniphila* AMDK-18

Figure 4.37 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-18. Table 4.20 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.38 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila*

strains are human pathogens were predicted by Pathogenecityisland viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

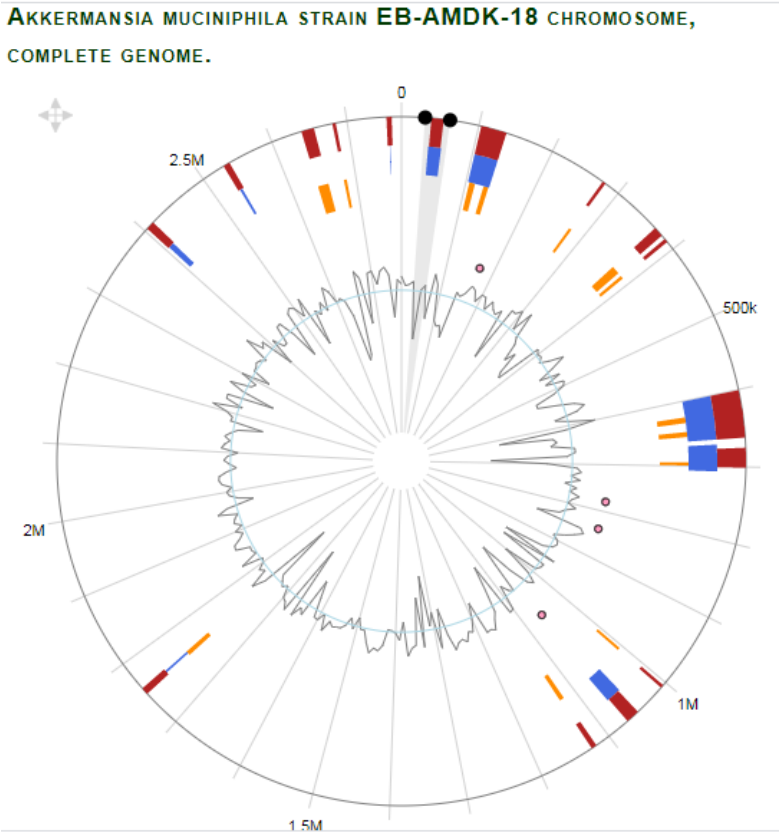

FIGURE 4.37: Genetic Islands in *Akkermansia muciniphila* AMDK-18 as Predicted by Island Viewer.

TABLE 4.20: Islands of *Akkermansia muciniphila* AMDK-18 Strain

Island start	Island end	Length	Gene name	Gene ID	Product
38412	54442	16030	WP_102732764.1		hypothetical protein
103147	136570	33423	WP_128154229.1		hypothetical protein
105424	114352	8928	WP_102732793.1		hypothetical protein
128042	135618	7576	WP_128154226.1		hypothetical protein
275152	279406	4254	WP_102732991.1		hypothetical protein
364869	377740	12871	WP_094140769.1		hypothetical protein
383429	388420	4991	WP_102731926.1		hypothetical protein

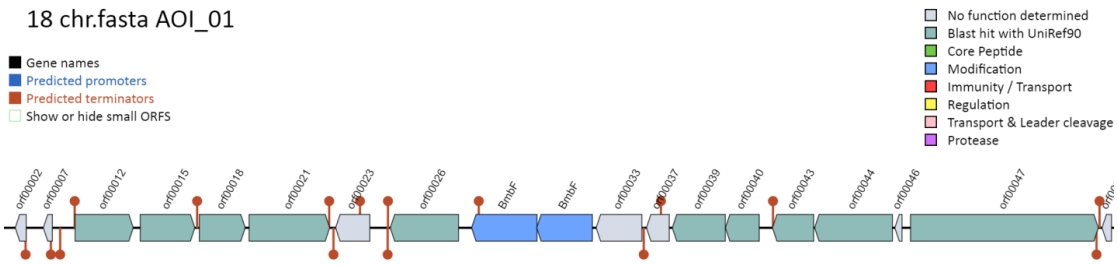

FIGURE 4.38: Bacteriocin of *Akkermansia muciniphila* AMBK-18 from BAGEL4

4.2.21 *Akkermansia muciniphila* AMDK-19

Figure 4.39 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-19. Table 4.21 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.40 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila* strains are human pathogens were predicted by Pathogenecityisland viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

TABLE 4.21: Islands of *Akkermansia muciniphila* AMDK-19 Strain

Island start	Island end	Length	Gene name	Gene ID	Product
101105	137054	35949	WP_046436080.1		tyrosine-type recombinase/ integrase
103376	111415	8039	WP_128153378.1		hypothetical protein DUF3320
116342	125106	8764	WP_102734753.1		domain-containing protein
128834	136411	7577	WP_102734759.1		hypothetical protein
367295	378718	11423	WP_022198322.1	ybeY	rRNA maturation RNase YbeY

TABLE 4.21: Islands of *Akkermansia muciniphila* AMDK-19 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
808554	832656	24102	WP_102732483.1	mnmA	tRNA 2-thiouridine (34) synthase MnmA
810154	831135	20981	WP_102734931.1		trimeric intracellular cation channel family protein
874762	886808	12046	WP_094140708.1		helix-turn-helix transcriptional regulator
1093969	1099465	5496	WP_128153805.1		type I restriction endonuclease subunit R
1480095	1485950	5855	WP_123038903.1		DUF2778 domain-containing protein
2553401	2558614	5213	WP_102734060.1	cas5c	type I-C CRISPR-\associated protein Cas5
2587918	2604526	16608	WP_022197102.1		glycosyl transferase family 2 protein
2686155	2691237	5082	WP_046437351.1	lepB	signal peptidase I

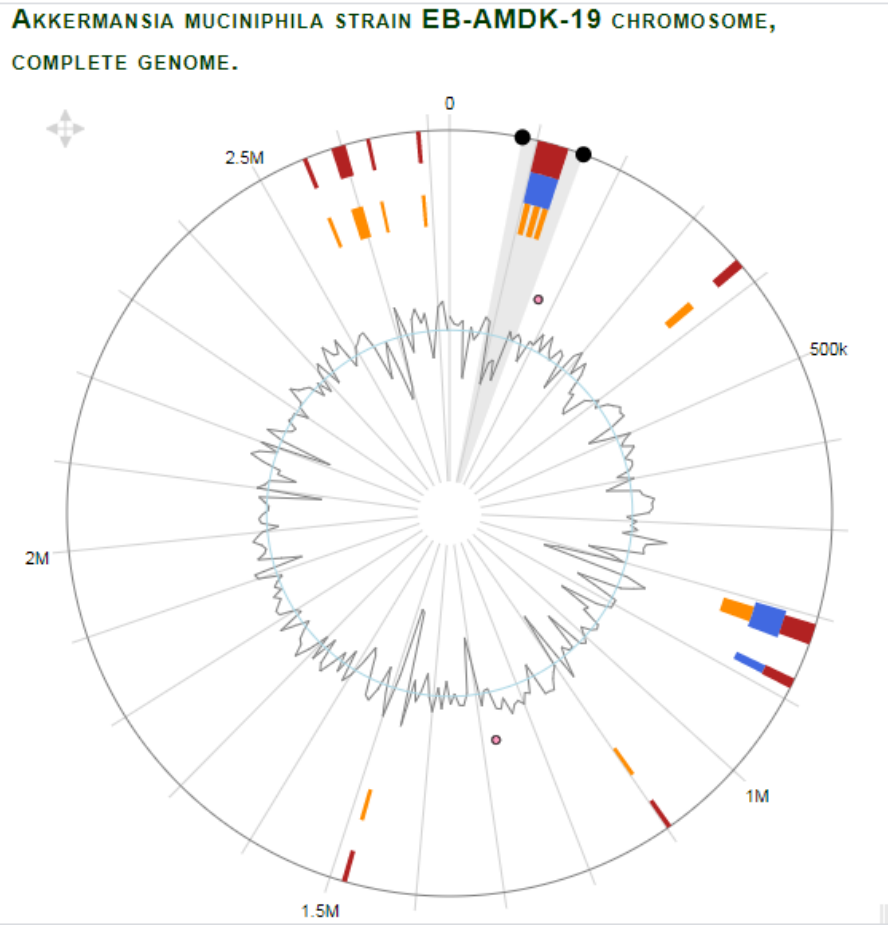

FIGURE 4.39: Genetic Islands in *Akkermansia muciniphila* AMDK-19 as Predicted by Island Viewer.

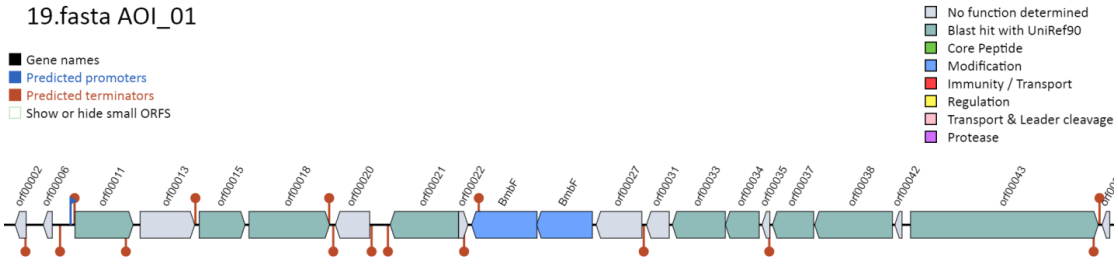

FIGURE 4.40: Bacteriocin of *Akkermansia muciniphila* AMBK-19 from BAGEL4

4.2.22 *Akkermansia muciniphila* AMDK-20

Figure 4.41 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-20. Table 4.22 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.42 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila*

strains are human pathogens were predicted by Pathogenecityisland viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

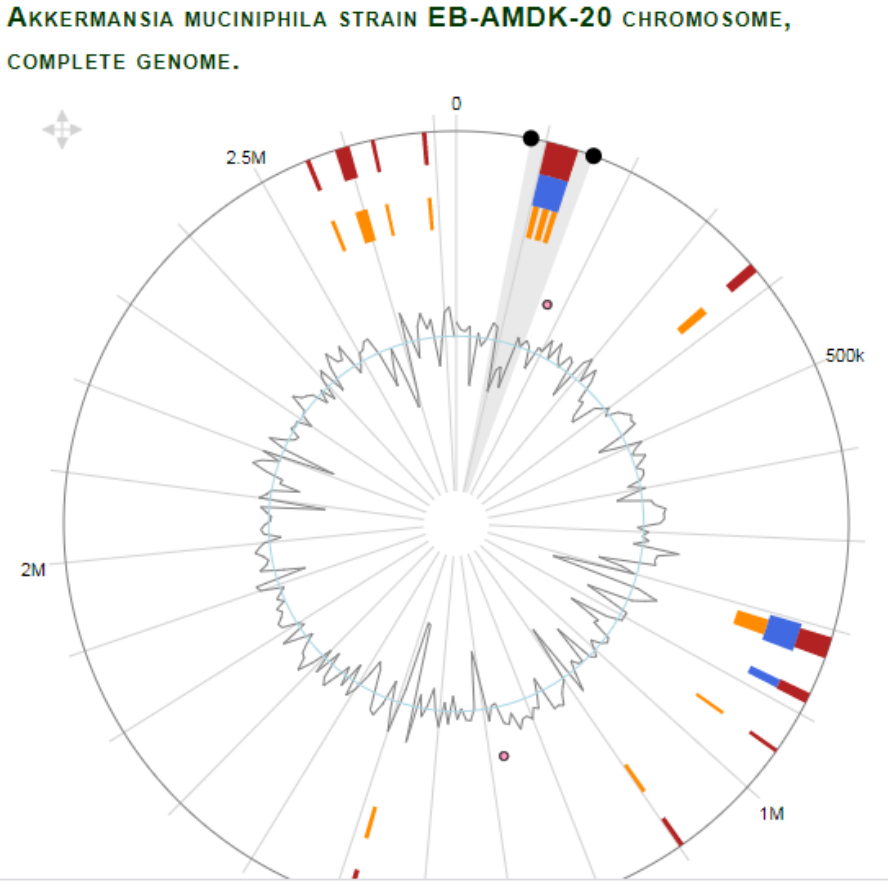

FIGURE 4.41: Genetic Islands in *Akkermansia muciniphila* AMDK-20. as predicted by Island Viewer.

TABLE 4.22: Islands of *Akkermansia muciniphila* AMDK-20 Strain

Island start	Island end	Length	Gene name	Gene ID	Product
101102	137049	35947	WP_046436080.1		tyrosine-type recombinase /integrase
103373	111436	8063	WP_046436078.1		hypothetical protein
116338	125102	8764	WP_128153685.1		hypothetical protein

TABLE 4.22: Islands of *Akkermansia muciniphila* AMDK-20 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
129162	136406	7244	WP_046436046.1		phosphatidy linositol kinase
367288	378711	11423	WP_102733694.1		HDIG domain-containing protein
810138	831119	20981	WP_102734931.1		trimeric intracellular cation channel\ family protein
874745	886791	12046	WP_128153779.1		PEP-CTERM sorting domain-containing protein
946145	950694	4549		guaA	glutamine-hydrolyzing GMP synthase
946145	950694	4549	WP_128153785.1		hypothetical protein
1093926	1099422	5496	WP_128153803.1		restriction endonuclease subunit S
1480067	1485921	5854	WP_081429137.1	kdpF	K(+)-transporting ATPase subunit F
2553348	2558560	5212	WP_094136088.1	cas2	CRISPR-associated endonuclease Cas2
2686094	2691176	5082	WP_094137864.1		hypothetical protein

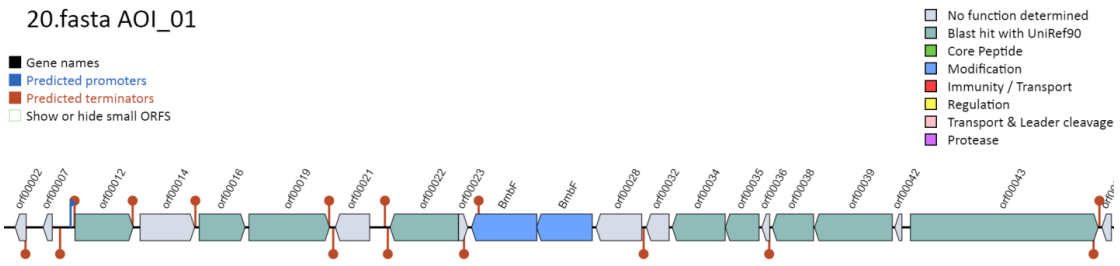

FIGURE 4.42: Bacteriocin of *Akkermansia muciniphila* AMBK-20 from BAGEL4

4.2.23 *Akkermansia muciniphila* AMDK-21

Figure 4.43 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-21.Table 4.23 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.44 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila* strains are human pathogens were predicted by Pathogenecity island viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

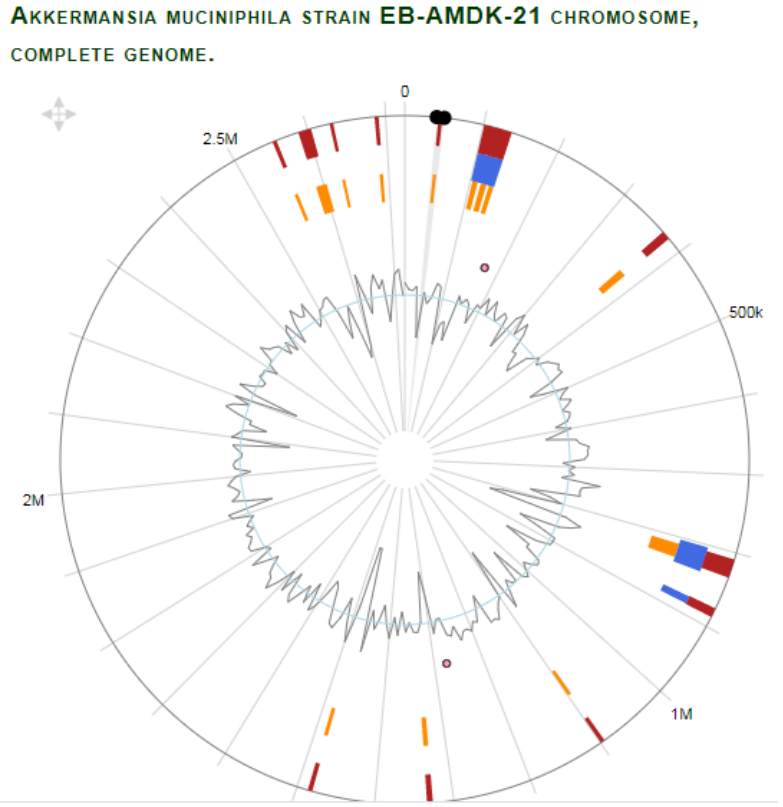

FIGURE 4.43: Genetic islands in *Akkermansia muciniphila* AMDK-21 as predicted by Island Viewer.

TABLE 4.23: Genetic islands in *Akkermansia muciniphila* AMDK-21

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
42719	47476	4757	WP_102734724.1		hypothetical protein
101072	137019	35947	WP_046436036.1		hypothetical protein
103343	111404	8061	WP_128153682.1		hypothetical protein
116307	125071	8764	WP_128153685.1		hypothetical protein
129131	136376	7245	WP_046436036.1		hypothetical protein
367245	378667	11422	WP_102733694.1		HDIG domain-containing protein
808484	832586	24102	WP_102733745.1		MBL fold metallo-hydrolase
810084	831065	20981	WP_102734931.1		trimeric intracellular cation channel family protein
874687	886733	12046	WP_094140708.1		helix-turn-helix transcriptional regulator
1093891	1099387	5496	WP_128153805.1		type I restriction endonuclease subunit R
1326520	1334008	7488	WP_022197112.1		hypothetical protein

TABLE 4.23: Genetic islands in *Akkermansia muciniphila* AMDK-21

Island start	Island end	Length	Gene name	Gene ID	Product
1480037	1485891	5854	WP_081429137.1	kdpF	K(+)-transporting ATPase subunit F
2553316	2558529	5213	WP_102734692.1		CRISPR-associated helicase/ endonuclease Cas3
2587831	2604438	16607	WP_022197102.1		glycosyltransferase family 2 protein
2628722	2633199	4477	WP_094139728.1		hypothetical protein

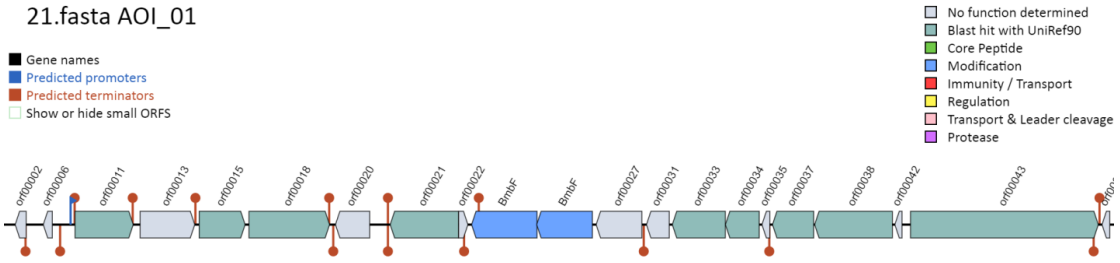

FIGURE 4.44: Genetic Islands in *Akkermansia muciniphila* ATCC BAA-835 as Predicted by Island Viewer.

4.2.24 *Akkermansia muciniphila* AMDK-22

Figure 4.45 Indicates the Presence of Genomic Islands in *Akkermansia muciniphila* AMDK-22. Table 4.24 Summarizes the Details of Genomic Islands and Genes Present in Respective Island. Figure 4.46 Indicates the Bacteriocin Producing Genes Predicted by BAGEL. The probabilities that *Akkermansia muciniphila*

strains are human pathogens were predicted by Pathogenecityisland viewer and no pathogenecity islands found in *Akkermansia muciniphila*.

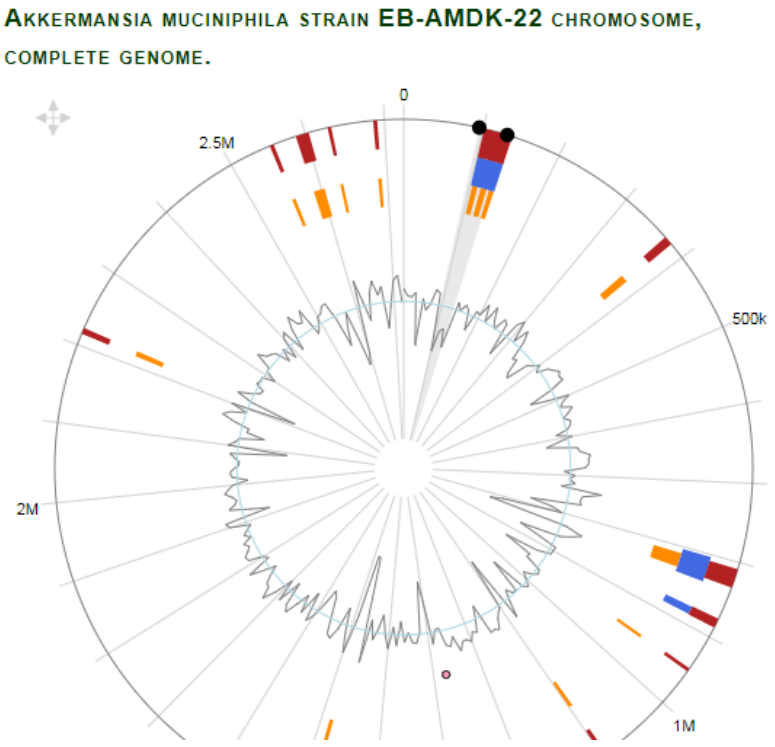

FIGURE 4.45: Genetic Islands in *Akkermania muciniphila* AMDK-22 as Predicted by Island Viewer.

TABLE 4.24: Islands of *Akkermansia muciniphila* AMDK-22 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
101101	137050	35949	WP_046436080.1		tyrosine-type recombinase/ integrase
103372	111435	8063	WP_128153682.1		hypothetical protein
129162	136407	7245	WP_046436036.1		hypothetical protein
367284	378706	11422	WP_094140769.1		hypothetical protein

TABLE 4.24: Islands of *Akkermansia muciniphila* AMDK-22 Strain

Island start	**Island end**	**Length**	**Gene name**	**Gene ID**	**Product**
808516	832618	24102	WP_102734924.1		hypothetical protein
810116	831097	20981	WP_102734931.1		trimeric intracellular cation channel family protein
874722	886768	12046	WP_128153779.1		PEP-CTERM sorting domain-containing protein
946120	950669	4549		guaA	glutamine-hydrolyzing GMP synthase
1093917	1099413	5496	WP_128153805.1		type I restriction endonuclease subunit R
1480056	1485911	5855	WP_081429137.1	kdpF	K(+)-transporting ATPase subunit F
2214709	2222414	7705	WP_102739544.1		AAA domain-containing protein
2553325	2558538	5213	WP_094136088.1	cas2	CRISPR-associated endonuclease Cas2

TABLE 4.24: Islands of *Akkermansia muciniphila* AMDK-22 Strain

Island start	Island end	Length	Gene name	Gene ID	Product
2587841	2604448	16607	WP_022197102.1		glycosylt ransferase family 2 p rotein
2686072	2691153	5081	WP_046437351.1	lepB	signal peptidase I

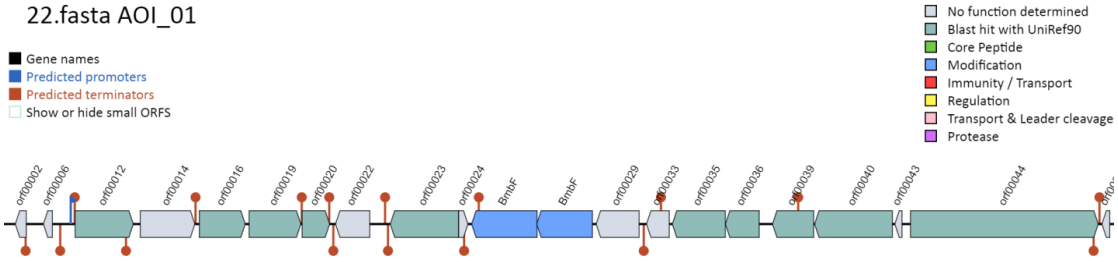

FIGURE 4.46: Bacteriocin of *Akkermansia muciniphila* AMBK-22 from BAGEL4

Chapter 5

Conclusions and Recommendations

Due to rising resistance against various antibiotics among pathogens, as well as the negative health impact of antibiotics of host health, a focus has been diverted from treatment to prevention. It is now preferred to boost health and immunity of host to fight against pathogen instead of using chemical entities to kill pathogens. Use if probiotics is one of these strategies, where health benefits or beneficialcapacities of normal gut microflora is used. It is quite a common practice to use probiotics against gut dysbiosis. But the potential use of probiotics against other diseases is yet to be explore and has great potential. Obesity is very prevalent disease and enjoys the status of a pandemic, Pakistan lies at position nine in list of obese countries. Various surgical procedure, therapeutic interventions are used to control obesity along with changes in food and exercise regimen. *Akkermansia muciniphila*is reported repeatedly to be associated with control of obesity. *Akkermansia muciniphila* is a normal gut microflora and is part of healthy gut microbiome. In principles this bacterial species has a great potential to be used as a probiotic against obesity. on the other hand, the questions have been raised against the safety of this bacterial species. This project was designed to check the probiotic potentials and safety of *Akkermansia muciniphila* as probiotic. In silico pipeline of Pangenome analysis was utilized to check the presence of virulent

genes in order to ensure safety. This analysis also ensured the genome plasticity and vulnerability of the bacterial species to possess virulent genes. To determine the evolution in the strains phylogenetic analysis was performed. Probiotic potentials were determined using BAGEL.

It was found by pangenome analysis as well COG and phylogenetic analysis that genome of all selected strains (selected based on availability of whole genome sequence and human origin) is stable and no frequent shuffling are observed in these genomes. Another feature for safety was resistome analysis and it was found that all selected strains of *Akkermansia muciniphila* just show intrinsic resistance against commonly used antibiotics, a required character for a potential probiotic to maintain healthy gut population. No multidrug resistance was found in any of the selected strains. For further validation of results, all the potential islands were analyzed using Island Viewer and analysis of genes revealed that no potential virulent genes are present.

Bacteriocin productions another important character, these small peptides are secreted by a bacterium to inhibit growth of closely related bacterial species. All the bacterial strains were found to possess bacteriocin production genes. Hence, we can conclude based on these observations that *Akkermansia muciniphila* specifically the 19 selected strains in this study are found to safe for use as probiotic against obesity. The major constrains or limiting factors in generalizing this opinion of safety is un availability of whole genome sequence of various strains of human origin, it is necessary to select few strains, perform sequence analysis to explore presence of virulent determinants or pathogenic genes. For future, it is strongly recommended to have validation by in vivo studies so that we can have a better idea about the safety and probiotic potential of *Akkermansia muciniphila* against obesity.

Bibliography

[1]. Derrien, M., C. Belzer, and W.M. de Vos, Akkermansia muciniphila and its role in regulating host functions. Microbial pathogenesis, 2017. 106: p'p. 171-181.

[2]. Anhê, F.F., J.D. Schertzer, and A. Marette, Bacteria to alleviate metabolic syndrome. Nature medicine, 2019. 25(7): p'p. 1031-1033.

[3]. WHO, F., Human vitamin and mineral requirements. Food and Agriculture Organization of the United Nations, Bangkok, Thailand. ISBN, 2002: p'p. 1014-9228.

[4]. "AFRC, R.F., Probiotics in man and animals. Journal of applied bacteriology", 1989. 66(5): p'p. 365-378;

[5]. Kalam-Salminen, L., Hoidon laatu lapsivuodeosastoilla Suomessa ja Virossa: näkökulmia asiakaskeskeiseen laatuun ja sen kehittämiseen. 2005: Turun yliopisto, p'p. 203-255.

[6]. Kleerebezem, M., et al., Complete genome sequence of Lactobacillus plantarum WCFS1. Proceedings of the National Academy of Sciences, 2003. 100(4): p'p. 1990-1995.

[7]. Nagpal, R., et al., Gut microbiota in health and disease: an overview focused on metabolic inflammation. Beneficial microbes, 2016. 7(2): p'p. 181-194.

[8]. D'emilio, C., et al., Apparatus and method for detecting weapons of mass destruction. 2006, Google Patents.

[9]. Ley, R.E., et al., Human gut microbes associated with obesity. nature, 2006. 444(7122): p'p. 1022-1023.

[10]. Wilkens, H., et al., Effect of inhaled iloprost plus oral sildenafil in patients with primary pulmonary hypertension. Circulation, 2001. 104(11): p'p. 1218-1222.

[11]. "Forestier, C., et al., Probiotic activities of Lactobacillus casei rhamnosus: in vitro adherence to intestinal cells and antimicrobial properties". Research in Microbiology, 2001. 152(2): p'p. 167-173;

[12]. "Lavasani, S., et al., A novel probiotic mixture exerts a therapeutic effect on experimental autoimmune encephalomyelitis mediated by IL-10 producing regulatory T cells". PloS one, 2010. 5(2): p'p. e9009-9012;

[13]. "Nagpal, R., et al., Probiotics, their health benefits and applications for developing healthier foods: a review. FEMS microbiology letters", 2012. 334(1): p'p. 1-15;

[14]. "Derrien, M., et al., Akkermansia muciniphila gen. nov., sp. nov., a human intestinal mucin-degrading bacterium". International journal of systematic and evolutionary microbiology, 2004. 54(5): p'p. 1469-1476;

[15]. Koskenniemi, K., et al., Proteomics and transcriptomics characterization of bile stress response in probiotic Lactobacillus rhamnosus GG. Molecular & Cellular Proteomics, 2011. 10(2).

[16]. Yan, F. and D.B. Polk, Probiotics as functional food in the treatment of diarrhea. Current Opinion in Clinical Nutrition & Metabolic Care, 2006. 9(6): p'p. 717-721.

[17]. Mack, D., et al., Extracellular MUC3 mucin secretion follows adherence of Lactobacillus strains to intestinal epithelial cells in vitro. Gut, 2003. 52(6): p'p. 827-833.

[18]. Mattar, A., et al., Probiotics up-regulate MUC-2 mucin gene expression in a Caco-2 cell-culture model. Pediatric surgery international, 2002. 18(7): p'p. 586-590.

[19]. "Gaudier, E., et al., The VSL♯ 3 probiotic mixture modifies microflora but does not heal chronic dextran-sodium sulfate–induced colitis or reinforce the mucus barrier in mice". The Journal of nutrition, 2005. 135(12): p'p. 2753-2761;

[20]. "Caballero-Franco, C., et al., The VSL♯ 3 probiotic formula induces mucin gene expression and secretion in colonic epithelial cells". American Journal of Physiology-Gastrointestinal and Liver Physiology, 2007. 292(1): p'p. G315-G322;

[21]. "Kelsall, B., Innate and adaptive mechanisms to control of pathological intestinal inflammation". The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland, 2008. 214(2): p'p. 242-259;

[22]. Möndel, M., et al., Probiotic E. coli treatment mediates antimicrobial human β-defensin synthesis and fecal excretion in humans. Mucosal immunology, 2009. 2(2): p'p. 166-172.

[23]. Schlee, M., et al., Probiotic lactobacilli and VSL♯ 3 induce enterocyte β-defensin 2. Clinical & Experimental Immunology, 2008. 151(3): p'p. 528-535.

[24]. Penner, R., R.N. Fedorak, and K.L. Madsen, Probiotics and nutraceuticals: non-medicinal treatments of gastrointestinal diseases. Current opinion in pharmacology, 2005. 5(6): p'p. 596-603.

[25]. Alakomi, H.-L., et al., Lactic acid permeabilizes gram-negative bacteria by disrupting the outer membrane. Appl. Environ. Microbiol., 2000. 66(5): p'p. 2001-2005.

[26]. "Liévin-Le Moal, V. and A.L. Servin, The front line of enteric host defense against unwelcome intrusion of harmful microorganisms: mucins, antimicrobial peptides, and microbiota". Clinical microbiology reviews, 2006. 19(2): p'p. 315-337;

[27]. Vassiliadis, G., et al., Isolation and characterization of two members of the siderophore-microcin family, microcins M and H47. Antimicrobial agents and chemotherapy, 2010. 54(1): p'p. 288-297.

[28]. "Johnson-Henry, K.C., et al., Surface-layer protein extracts from Lactobacillus helveticus inhibit enterohaemorrhagic Escherichia coli O157: H7 adhesion to epithelial cells. Cellular microbiology", 2007. 9(2): p'p. 356-367;

[29]. "Wu, X., et al., Saccharomyces boulardii ameliorates Citrobacter rodentium-induced colitis through actions on bacterial virulence factors. American Journal of Physiology-Gastrointestinal and Liver Physiology", 2008. 294(1): p'p. G295-G306;

[30]. Galdeano, C.M. and G. Perdigon, The probiotic bacterium Lactobacillus casei induces activation of the gut mucosal immune system through innate immunity. Clin. Vaccine Immunol., 2006. 13(2): p'p. 219-226.

[31]. Ogawa, M., et al., Protective Effect of Lactobacillus casei Strain Shirota on Shiga Toxin-Producing Escherichia coliO157: H7 Infection in Infant Rabbits. Infection and immunity, 2001. 69(2): p'p. 1101-1108.

[32]. "Tien, M.-T., et al., Anti-inflammatory effect of Lactobacillus casei on Shigella-infected human intestinal epithelial cells". The Journal of Immunology, 2006. 176(2): p'p. 1228-1237;

[33]. Miller, M.B. and B.L. Bassler, Quorum sensing in bacteria. Annual Reviews in Microbiology, 2001. 55(1): p'p. 165-199.

[34]. Kendall, M.M. and V. Sperandio, Quorum sensing by enteric pathogens. Current opinion in gastroenterology, 2007. 23(1): p'p. 10-15.

[35]. Medellin-Peña, M.J., et al., Probiotics affect virulence-related gene expression in Escherichia coli O157: H7. Appl. Environ. Microbiol., 2007. 73(13): p'p. 4259-4267.

[36]. Gogineni, V.K., L.E. Morrow, and M. Malesker, Probiotics: mechanisms of action and clinical applications. J Prob Health, 2013. 1(101): p'p. 2.

[37]. "Bartlett, J.G., Management of Clostridium difficile infection and other antibiotic-associated diarrhoeas. European journal of gastroenterology & hepatology", 1996. 8(11): p'p. 1054-1061;

[38]. Mueller, S., et al., Differences in fecal microbiota in different European study populations in relation to age, gender, and country: a cross-sectional study. Appl. Environ. Microbiol., 2006. 72(2): p'p. 1027-1033.

[39]. Guigoz, Y., J. Doré, and E.J. Schiffrin, The inflammatory status of old age can be nurtured from the intestinal environment. Current Opinion in Clinical Nutrition & Metabolic Care, 2008. 11(1): p'p. 13-20.

[40]. "Szajewska, H., M. Ruszczyński, and A. Radzikowski, Probiotics in the prevention of antibiotic-associated diarrhea in children: a meta-analysis of randomized controlled trials". The Journal of pediatrics, 2006. 149(3): p'p. 367-372. e1;

[41]. Adam, J., et al., Controlled double-blind clinical trials of Ultra-Levure: Multicentre study by 25 physicians in 388 cases. Gazette Medicale de France, 1977. 84: p'p. 2072-2078.

[42]. Kotowska, M., P. Albrecht, and H. Szajewska, Saccharomyces boulardii in the prevention of antibiotic-associated diarrhoea in children: a randomized double-blind placebo-controlled trial. Alimentary pharmacology & therapeutics, 2005. 21(5): p'p. 583-590.

[43]. "McFarland, L.V., Systematic review and meta-analysis of Saccharomyces boulardii in adult patients". World journal of gastroenterology: WJG, 2010. 16(18): p'p. 2202;

[44]. Castagliuolo, I., et al., Saccharomyces boulardii protease inhibits Clostridium difficile toxin A effects in the rat ileum. Infection and immunity, 1996. 64(12): p'p. 5225-5232.

[45]. Qamar, A., et al., Saccharomyces boulardii stimulates intestinal immunoglobulin a immune response to clostridium difficiletoxin a in mice. Infection and immunity, 2001. 69(4): p'p. 2762-2765.

[46]. "McFarland, L.V., et al., Prevention of b-lactam-associated diarrhea by Saccharomyces boulardii compared with placebo". American Journal of Gastroenterology, 1995. 90(3): p'p. 439-448;

[47]. Plummer, S., et al., Clostridium difficile pilot study: effects of probiotic supplementation on the incidence of C. difficile diarrhoea. International Microbiology, 2004. 7(1): p'p. 59-62.

[48]. Thomas, M.R., et al. Lack of effect of Lactobacillus GG on antibiotic-associated diarrhea: a randomized, placebo-controlled trial. in Mayo Clinic Proceedings. 2001. Elsevier.

[49]. Can, M., et al., Prophylactic Saccharomyces boulardii in the prevention of antibiotic-associated diarrhea: a prospective study. Medical science monitor, 2006. 12(4): p'p. PI19-PI22.

[50]. Surawicz, C.M., et al., The search for a better treatment for recurrent Clostridium difficile disease: use of high-dose vancomycin combined with Saccharomyces boulardii. Clinical infectious diseases, 2000. 31(4): p'p. 1012-1017.

[51]. Hickson, M., et al., Use of probiotic Lactobacillus preparation to prevent diarrhoea associated with antibiotics: randomised double blind placebo controlled trial. Bmj, 2007. 335(7610): p'p. 80.

[52]. Armuzzi, A., et al., The effect of oral administration of Lactobacillus GG on antibiotic-associated gastrointestinal side-effects during Helicobacter pylori

eradication therapy. Alimentary pharmacology & therapeutics, 2001. 15(2): p'p. 163-169.

[53]. Dao, M.C., et al., Akkermansia muciniphila and improved metabolic health during a dietary intervention in obesity: relationship with gut microbiome richness and ecology. Gut, 2016. 65(3): p'p. 426-436.

[54]. "Cremonini, F., et al., Effect of different probiotic preparations on anti-Helicobacter pylori therapy-related side effects: a parallel group, triple blind, placebo-controlled study". The American journal of gastroenterology, 2002. 97(11): p'p. 2744-2749;

[55]. Szajewska, H., A. Horvath, and A. Piwowarczyk, Meta-analysis: the effects of Saccharomyces boulardii supplementation on Helicobacter pylori eradication rates and side effects during treatment. Alimentary pharmacology & therapeutics, 2010. 32(9): p'p. 1069-1079.

[56]. Marshall, J.K., et al., Incidence and epidemiology of irritable bowel syndrome after a large waterborne outbreak of bacterial dysentery. Gastroenterology, 2006. 131(2): p'p. 445-450.

[57]. "Porter, C.K., et al., The incidence and gastrointestinal infectious risk of functional gastrointestinal disorders in a healthy US adult population". American Journal of Gastroenterology, 2011. 106(1): p'p. 130-138;

[58]. "Husebye, E., et al., Influence of microbial species on small intestinal myoelectric activity and transit in germ-free rats". American Journal of Physiology-Gastrointestinal and Liver Physiology, 2001. 280(3): p'p. G368-G380;

[59]. Shah, E.D., et al., Abnormal breath testing in IBS: a meta-analysis. Digestive diseases and sciences, 2010. 55(9): p'p. 2441-2449.

[60]. Moayyedi, P., et al., The efficacy of probiotics in the treatment of irritable bowel syndrome: a systematic review. Gut, 2010. 59(3): p'p. 325-332.

[61]. "Brenner, D.M., et al., The utility of probiotics in the treatment of irritable bowel syndrome: a systematic review. American Journal of Gastroenterology", 2009. 104(4): p'p. 1033-1049;

[62]. "McFarland, L.V. and S. Dublin, Meta-analysis of probiotics for the treatment of irritable bowel syndrome". World journal of gastroenterology: WJG, 2008. 14(17): p'p. 2650;

[63]. for Nursing, N.C.C., Literature review of prognostic, resource use and quality of data, in Irritable Bowel Syndrome in Adults: Diagnosis and Management of Irritable Bowel Syndrome in Primary Care [Internet]. 2008, Royal College of Nursing (UK).

[64]. Hoveyda, N., et al., A systematic review and meta-analysis: probiotics in the treatment of irritable bowel syndrome. BMC gastroenterology, 2009. 9(1): p'p. 15.

[65]. "Society, A.T. and I.D.S.o. America, Guidelines for the management of adults with hospital-acquired, ventilator-associated, and healthcare-associated pneumonia". American journal of respiratory and critical care medicine, 2005. 171(4): p'p. 388;

[66]. Safdar, N., et al., Clinical and economic consequences of ventilator-associated pneumonia: a systematic review. Critical care medicine, 2005. 33(10): p'p. 2184-2193.

[67]. Craven, D.E., F.G. De Rosa, and D. Thornton, Nosocomial pneumonia: emerging concepts in diagnosis, management, and prophylaxis. Current opinion in critical care, 2002. 8(5): p'p. 421-429.

[68]. "Chastre, J. and J.-Y. Fagon, Ventilator-associated pneumonia. American journal of respiratory and critical care medicine", 2002. 165(7): p'p. 867-903;

[69]. Kollef, M.H., What is ventilator-associated pneumonia and why is it important? Respiratory care, 2005. 50(6): p'p. 714-724.

[70]. Kollef, M.H., Prevention of hospital-associated pneumonia and ventilator-associated pneumonia. Critical care medicine, 2004. 32(6): p'p. 1396-1405.

[71]. Forestier, C., et al., Oral probiotic and prevention of Pseudomonas aeruginosa infections: a randomized, double-blind, placebo-controlled pilot study in intensive care unit patients. Critical care, 2008. 12(3): p'p. R69.

[72]. Knight, D.J., et al., Effect of synbiotic therapy on the incidence of ventilator associated pneumonia in critically ill patients: a randomised, double-blind, placebo-controlled trial. Intensive care medicine, 2009. 35(5): p'p. 854.

[73]. Rook, G. and L. Brunet, Microbes, immunoregulation, and the gut. Gut, 2005. 54(3): p'p. 317-320.

[74]. Vanderpool, C., F. Yan, and B.D. Polk, Mechanisms of probiotic action: implications for therapeutic applications in inflammatory bowel diseases. Inflammatory bowel diseases, 2008. 14(11): p'p. 1585-1596.

[75]. Grönlund, M., et al., Importance of intestinal colonisation in the maturation of humoral immunity in early infancy: a prospective follow up study of healthy infants aged 0–6 months. Archives of Disease in Childhood-Fetal and Neonatal Edition, 2000. 83(3): p'p. F186-F192.

[76]. Van Passel, M.W., et al., The genome of Akkermansia muciniphila, a dedicated intestinal mucin degrader, and its use in exploring intestinal metagenomes. PloS one, 2011. 6(3): p'p. e16876.

[77]. "Singh, M., et al., Efficacy of different estrus synchronization protocols in repeat breeder cows". INDIAN JOURNAL OF ANIMAL SCIENCES, 2019. 89(9): p'p. 958-960;

[78]. Collado, M.C., et al., Intestinal integrity and Akkermansia muciniphila, a mucin-degrading member of the intestinal microbiota present in infants, adults, and the elderly. Applied and environmental microbiology, 2007. 73(23): p'p. 7767-7770.

[79]. Karpiński, T.M. and A.K. Szkaradkiewicz, Microbiology of dental caries. J Biol Earth Sci, 2013. 3(1): p'p. M21-4.

[80]. Jan, G., Belzacq AS, Haouzi D, Rouault A, Metivier D, Kroemer G, and Brenner C. Propionibacteria induce apoptosis of colorectal carcinoma cells via short-chain fatty acids acting on mitochondria. Cell Death Differ, 2002. 9: p'p. 179-188.

[81]. Shao, Z., H. Zhao, and H. Zhao, DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. Nucleic acids research, 2009. 37(2): p'p. e16-e16.

[82]. Overbeek, R., et al., The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). 2014. 42(D1): p'p. D206-D214.

[83]. van Passel, M.W.J., et al., The genome of Akkermansia muciniphila, a dedicated intestinal mucin degrader, and its use in exploring intestinal metagenomes. PloS one, 2011. 6(3): p'p. e16876-e16876.

[84]. "Cheung, E.Y., M.E. Gahan, and D. McNevin, Predictive DNA analysis for biogeographical ancestry. Australian Journal of Forensic Sciences" 2018. 50(6): p'p. 651-658;

[85]. Jia, B., et al., CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic acids research, 2016: p'p. gkw1004.

[86]. Chen, L., et al., VFDB: a reference database for bacterial virulence factors. 2005. 33(suppl_1): p'p. D325-D328.

[87]. Ribbera, A., et al., Comparative genomic and functional analysis of Lactobacillus casei and Lactobacillus rhamnosus strains marketed as probiotics. Applied and environmental microbiology, 2013. 79(6): p'p. 1923-1933.

[88]. Sherry, S.T., et al., dbSNP: the NCBI database of genetic variation. Nucleic acids research, 2001. 29(1): p'p. 308-311.

[89]. Page, A.J., et al., Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics, 2015. 31(22): p'p. 3691-3693.

[90]. Chaudhari, N.M., V.K. Gupta, and C. Dutta, BPGA-an ultra-fast pan-genome analysis pipeline. Scientific reports, 2016. 6(1): p'p. 1-10.

[91]. Bertelli, C., et al., IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. Nucleic acids research, 2017. 45(W1): p'p. W30-W35.