Norwegian University
of Life Sciences

Master's Thesis 2020 60 ECTS
Faculty of Chemistry, Biotechnology and Food Science

Investigating prevalence and geographical distribution of *Mycoplasma* sp. in the gut of Atlantic Salmon (*Salmo salar* L.)

Mari Raudstein
MSc Biotechnology

ACKNOWLEDGEMENTS

This master project was performed at the Faculty of Chemistry, Biotechnology and Food Science, at the Norwegian University of Life Sciences (NMBU), with Professor Knut Rudi as primary supervisor and associate professor Lars-Gustav Snipen as secondary supervisor.

To begin with, I want to express my gratitude to my supervisor Knut Rudi for giving me the opportunity to include fish, one of my main interests, in my thesis. Professor Rudi has helped me execute this project to my best ability by giving me new ideas and solutions to problems that would occur, as well as answering any questions I would have. Further, my secondary supervisor associate professor Snipen has helped me in acquiring and interpreting shotgun data for this thesis. A big thank you to both of you.

I would also like to thank laboratory engineer Inga Leena Angell for all the guidance in the laboratory, and for her patience when answering questions. Also, thank you to Ida Ormaasen, Morten Nilsen, and the rest of the Microbial Diversity group at NMBU for always being positive, friendly, and helpful – it has been a pleasure working at the MiDiv lab the past year!

I am very grateful to the salmon farmers providing me with the material necessary to study salmon gut microbiota. Without the generosity of Lerøy Sjøtroll in Bømlo, Lerøy Aurora in Skjervøy and Mowi in Chile, I would not be able to do this work. Also, thank you to Cyril Frantzen, who has corresponded with his contacts in Chile.

Finally, I want to thank the other master students under Professor Knut's supervision, Cecilie, Unni Lise, and Regina. Your support and friendship have meant a lot to me.

Kvinnherad, June 2020

Mari Raudstein

ABSTRACT

The fish gut microbiota has gotten considerable attention in recent years, and the microbes harboring the intestine of fish are thought to grant the host various effects related to size, metabolism, feeding behavior, and immune response. A *Mycoplasma* species has been discovered as highly abundant in the salmon gut. However, the resident strain has not yet been isolated. Knowledge regarding its colonization and the impact it may have on the host is, therefore, limited. This study aimed to map the prevalence and geographical distribution of *Mycoplasma* in the salmon gut and discover its potential role as part of the gut microbiota.

Salmon gut content was sampled for both cultivation purposes and direct DNA analyses in this project. Samples were collected from two salmon farms in Norway, Skjervøy (n = 23) and Bømlo (n = 19), and one in Chile (n = 20). A selection of Bømlo samples (n = 10) was cultivated in enriched growth medium. The prevalence of *Mycoplasma* at different geographical sites was investigated by analyzing the bacterial composition in the Bømlo and Chile samples using 16S rRNA gene sequencing. Moreover, selected samples from Bømlo (n = 4), Skjervøy (n = 7), and Chile (n = 1) were further processed for whole-genome shotgun sequencing to obtain genomic information of the salmon-associated *Mycoplasma*.

Mycoplasma was found abundantly in Norwegian salmon but was not detected in Chilean salmon. Thus, in this study, we observed a geographical difference (p = 0.00023) in the mycoplasmas' prevalence in the gut of farmed Atlantic salmon. The underlying reasons for the absence of *Mycoplasma* in Chilean salmon must be further investigated to explain our findings. Further, we found that the salmon-associated *Mycoplasma*'s DNA was most frequently classified as *M. penetrans*, which may suggest relatedness between these species. Whether the salmon *Mycoplasma* exhibits pathogenic or protective characteristics is not known. However, given the seemingly large prevalence of mycoplasmas in salmon, it is likely they exist in the gut microbiota as commensals. Further research is necessary to discover potential negative or positive impacts the salmon-associated *Mycoplasma* might have on the physiology and immunology of the fish.

SAMANDRAG

Tarmmikrobiotaen til fisk har fått auka merksemd dei siste åra, og mikroorganismane som utgjer denne har truleg innverknad på verten relatert til storleik, metabolisme, fôringsåtferd og immunrespons. Ein *Mykoplasma*-art har blitt oppdaga i rikelege mengder i laksetarmen. Arten er enno ikkje isolert, og det er lite kunnskap om denne bakterien si kolonisering, og om verknaden den kan ha på verten. Målet med denne studien var difor å kartlegge utbreiinga, samt den geografiske fordelinga av *Mykoplasma* i laksetarm, og å undersøke kva rolle denne bakterien potensielt har som del av tarmmikrobiotaen.

I dette prosjektet vart det samla inn tarminnhald frå laks for både kultivering i vekstmedium og direkte DNA analysar. Det vart henta prøvar frå to oppdrettsanlegg i Noreg, Skjervøy (n = 23) og Bømlo (n = 19), og eit i Chile (n = 20). Eit utval av Bømlo-prøvane (n = 10) vart dyrka i anrika vekstmedium. For å undersøke førekomsten av *Mykoplasma* på ulike geografiske stadar vart bakteriesamansetjinga i Bømlo- og Chile-prøvane analysert ved bruk av 16S rRNA gensekvensering. Vidare vart utvalde prøvar frå Bømlo (n = 4), Skjervøy (n = 7) og Chile (n = 1) prosessert for heilgenom-shotgunsekvensering for å skaffe informasjon om genomet til den lakse-assosierte *Mykoplasma*-arten.

Mykoplasma vart funne i rikelege mengder i tarmen hos norsk laks, men vart ikkje påvist i chilensk laks. I denne studien observerte vi difor ein geografisk skilnad (p = 0.00023) i utbreiinga av *Mykoplasma* i tarmen hos oppdrettslaks. For å forklare funna våre må dei underliggande årsakene til fråværet av *Mykoplasma* i chilensk laks undersøkast nærare. Vidare fann vi at DNA tilhøyrande den lakse-assosierte *Mykoplasma*-arten oftast blei klassifisert som *M. penetrans*, noko som kan tyde på slektskap mellom dei to artane. Om lakse-*Mykoplasma* har patogene eller beskyttande eigenskapar er ikkje kjent. Grunna den generelt hyppige førekomsten og utbreiinga av bakterien, er det mogleg at den eksisterer i tarmmikrobiotaen som ein kommensal bakterie. Det trengs likevel meir forsking for å oppdage potensielle negative eller positive effektar den kan ha på fisken sin fysiologi og immunsystem.

ABBREVIATIONS

ATP	–	Adenosine triphosphate
BLAST	–	Basic Local Alignment Search Tool
bp	–	Base pairs
ddNTP	–	Dideoxyribonucleotide triphosphate
DNA	–	Deoxyribonucleic acid
dNTP	–	Deoxyribonucleotide triphosphate
gDNA	–	Genomic DNA
kb	–	Kilobase pairs
MG-RAST	–	Metagenomic Rapid Annotations using Subsystems Technology
MycoBroth	–	Mycoplasma Growth Broth
OTU	–	Operational Taxonomic Unit
PCoA	–	Principal Coordinate Analysis
PCR	–	Polymerase chain reaction
QIIME	–	Quantitative Insights Into Microbial Ecology
rRNA	–	Ribosomal RNA
S.T.A.R	–	Stool Transport and Recovery
SPAdes	–	St. Petersburg genome assembler

TABLE OF CONTENTS

1. INTRODUCTION 1

1.1 Atlantic Salmon 2

1.1.1 Salmon Life Cycle 2

1.1.2 Habitat and Diet 3

1.1.3 Salmon Farming 4

1.2 The Fish Gut Microbiota 5

1.2.1 Colonization and Composition of the Salmon Gut Microbiota 5

1.3 The genus *Mycoplasma* 7

1.3.1 Phylogeny 7

1.3.2 Characteristics 7

1.3.3 Metabolism 8

1.3.4 Ecology and Habitat 9

1.3.5 Pathogenicity 9

1.3.6 Cultivation 10

1.4 Sequencing Approaches for Microbial Identification 11

1.4.1 Importance of Identifying and Characterizing Microbes 11

1.4.2 Targeted 16S rRNA Gene Sequencing 11

1.4.3 Whole-(meta)genome Shotgun Sequencing 12

1.5 Sequencing Technologies 14

1.5.1 First-generation Sequencing 14

1.5.2 Second-generation Sequencing 14

1.5.3 Third-generation Sequencing 16

1.6 Bioinformatic Software for Metagenome Analyses 17

1.6.1 Processing 16S rRNA Sequence Data 17

1.6.2 Processing Shotgun Metagenome Data 17

1.7 Aim of Thesis 18

2. MATERIALS AND METHODS 19

2.1 Samples Used in this Project and Sample Flow 19

2.2 Sampling Gut Content for Cultivation 21

2.2.1 Preparation of Enriched Mycoplasma Growth Broth and Agar Plates 21

2.2.2 Collecting Gut Content from Salmon 21

2.2.3 Cultivating Mycoplasmas in Mycoplasma Growth Broth and Agar Plates 22

2.2.4 Measuring pH value of Primary Cultures 22
2.3 Sampling Gut Content for DNA Analyses 23
2.3.1 Collecting Gut Content from Salmon 23
2.4 Extracting DNA for Downstream DNA Analyses 23
2.4.1 QIAamp DNA Mini Kit 23
2.4.2 Mag Midi LGC kit 24
2.4.3 Phenol-Chloroform-Isoamyl Alcohol Method 25
2.5 Quantification and Qualification of Extracted DNA 26
2.5.1 Measuring DNA Concentration on Qubit Fluorometer 26
2.5.2 Measuring DNA Fragment Size by Gel Electrophoresis 26
2.6 Polymerase Chain Reactions and DNA Sequencing 27
2.6.1 Quantitative PCR 27
2.6.2 Sanger Sequencing 27
2.6.3 16S rRNA Sequencing by Illumina MiSeq 28
2.6.4 Shotgun Metagenome Sequencing by Illumina HiSeq 29
2.6.5 Primer Overview 30
2.7 Data Analyses 30
2.7.1 Sanger Sequence Data 30
2.7.2 16S rRNA Sequence Data 31
2.7.3 Shotgun Metagenome Sequence Data 31
3. RESULTS 32
3.1 16S rRNA Analyses 32
3.1.1 Comparison of Gut Composition at Phylum Level 32
3.1.2 Comparison of Gut Composition at Genus Level 34
3.1.3 Microbial Community Comparison by Principal Coordinate Analysis 35
3.2 Shotgun Sequencing Analyses 37
3.2.1 Taxonomic Annotation of Sequences by MG-RAST 37
3.2.2 Recovery of Individual Genomes by MaxBin 40
3.3 Cultivation of Bacteria in Mycoplasma Growth Broth and Growth Agar 42
3.3.1 Growth in Primary Cultures 43
3.3.2 Growth on Agar Plates 44
3.3.3 DNA Quantification of Cell Pellets from Primary Cultures 45
3.3.4 pH Measurements of Supernatants from Primary Cultures 45
3.3.5 DNA Extraction and Verification of Mycoplasma Broth Cultures 45

4. DISCUSSION 46
4.1 Geographical Difference in *Mycoplasma* Prevalence 46
4.2 Salmon-associated *Mycoplasma's* relatedness to other *M.* spp. 47
4.3 Recovery of a Salmon-associated *Mycoplasma* Genome 48
4.4 The Potential Role of *Mycoplasma* in Salmon Gut 49
4.5 *Mycoplasma* Cultivation and DNA Extraction 50
4.6 Other Bacteria in the Salmon Gut 51
4.7 Technical Considerations 52
4.7.1 Small Sample-size 52
4.7.2 Lack of Information Regarding Sampled Fish 52
4.7.3 No Optimized Protocol for Mycoplasma Cultivation and DNA Extraction 52
4.8 Further Work 53
5. CONCLUSION 54
REFERENCES 55
APPENDIX 62
Appendix A: Overview of Samples Used in this Thesis 62
Appendix B: Final Concentrations of Enriched Mycoplasma Growth Broth 65
Appendix C: T-tests Conducted in Excel 66
Appendix D: Unweighted UniFrac PCoA plot 69
Appendix E: Taxonomic Annotation of Bacterial Shotgun Sequences at Species Level 70
Appendix F: Microbial Identification of Primary MycoBroth Cultures 72

1. INTRODUCTION

Seafood is the most traded food group globally, and Norway is the world's second-largest seafood producer (1). In 2018, the total wealth creation from the Norwegian seafood industry reached 100 billion NOK (2). Of all seafood exported, approximately 70% of the income came from salmon production, making Atlantic salmon the most important product within Norwegian aquaculture (1). Salmon farming allows large-scale production of Atlantic salmon and has made it an easily accessible food source. However, welfare and health issues warrant a challenge for the salmon farmers, leading to substantial economic losses. Infectious diseases attributable to pathogens like viruses, bacteria, and parasites, in addition to harsh treatment of said pathogens, cause approximately 15% of farmed salmon in Norway to perish during the production (3). Therefore, solving problems related to the protection of fish from pathogenic microbes is essential to advance the salmon industry further (4).

Studying the gut microbiota can hopefully provide useful insight into such problems: Does the gut microbiota affect the physiology of salmon? Does it have any immunological effects, and if so, how can it be used to benefit the salmon? There is a growing appreciation of the impact the gut microbiota conceivably has on fish health, and metagenomic research provides data on the potential physiological and immunological influences of the bacteria (4).

When mapping the microbes inhabiting the salmon gut, one of the most abundant bacteria in both farmed and wild salmon is *Mycoplasma*. The bacterium was first discovered as part of the salmon gut microbiota almost 20 years ago (5), but the resident strain has yet to be isolated. Little is therefore known regarding its colonization of the salmon gut and the impact it might have on the fish. A recent study found that most microbes colonize the gut based on a neutral model, meaning the surrounding environment is the most important factor in what bacteria are present (6). However, *Mycoplasma* seemed to be one of the main exceptions, suggesting that the host might influence the colonization of this bacterium. This practice further accentuates the question of *Mycoplasma's* role in the salmon gut and its interaction with the host – could it have properties that benefit the salmon?

1.1 Atlantic Salmon

Atlantic salmon (*Salmo salar*) is a species of ray-finned fish and part of the Salmonidae family. Along with salmon, the family includes char, trout, whitefish, and grayling (7). Salmon and trout – together forming the genus *Salmo* – show a notable anatomical resemblance and may sometimes be challenging to differentiate (8). Still, Atlantic salmon is considered the largest species in the genus, and male individuals can weigh up to 40 kg. When grown, the salmon has a small and pointy head, followed by a slim body with a slightly rounded tail fin at the end. However, the salmon's physiology changes remarkably throughout the phases of its life and during its maturation (8).

1.1.1 Salmon Life Cycle

Members of the salmonid family spend time in both freshwater and saltwater, a mode of life termed anadromy (9). Most Atlantic salmon are anadromous with a juvenile phase in freshwater, followed by migration to the sea for feeding and growth. When they become sexually mature, they return to freshwater to spawn.

Generally, anadromous Atlantic salmon spawn in rivers from September to February (9). The females dig nests in the gravel to deposit their eggs, which then hatch the following spring. The newly hatched fish, *alevins*, are still attached to their yolk sac, which they utilize for nutrition during the first weeks. When they finally emerge from the gravel and start feeding on plants and plankton, the fish are called *fry*. Eventually, the fish reach the *parr* stage. At this stage, they can remain in freshwater for 1-8 years, depending on environmental conditions and genetics (9). A physiological and morphological transformation then transpires, turning the parr into *smolts*. During this stage, the fish start migrating to the sea, swimming with the current instead of against it. When entering the sea, the fish are called *post-smolts.* This period is thought to be critical for the newly migrated fish as they are exposed to a novel environment containing higher salt concentrations, different food types, and dangerous predators. After 1-5 years at sea, the *adult* fish return to their native river to spawn, completing the salmon life cycle (9).

1.1.2 Habitat and Diet

The Atlantic salmon occur naturally along the east and west coasts of the North Atlantic Ocean (10). In the northeast Atlantic, salmon are found in watersheds from Portugal in the south, to the Barents and White Sea areas of Russia, in the northeast (9). In the northwest Atlantic, they are distributed from New England, the United States in the south, to Ungava Bay, Canada, in the north. During their marine life stage salmon increase their weight drastically, many over a 1000-fold or more (11). The primary factor that enables the growth and survival of salmon at sea is the increased availability of food compared to freshwater. Water temperature and other environmental factors may act indirectly, changing the production, and consequently, the availability of food.

Salmon are opportunistic feeders, feeding on a variety of available prey. As natural carnivores, their main prey is usually other species of fish and fish larvae, but also planktonic crustaceans (11). Post-smolt, pre-adult and adult Atlantic salmon have been reported to consume more than 40 different fish species from at least 19 families, and invertebrates from more than 10 major taxonomic groups. Independent of life stage, habitats and season, fish species such as eel, herring, capelin, and cod, as well as planktonic amphipods, are the primary components of a salmon's diet (11). Although salmon feed on several types of prey, it is uncommon to find more than 1 to 3 marine prey species in their stomach at the same time (11). This may indicate that prey availability varies depending on location and time of year, but also that individual salmon may prefer a specific type of prey. It is suggested that salmon select forage fish such as capelin due to their energy content (11). Capelins are usually higher in lipids than other potential prey organisms, making them more energy-dense and, thus, more beneficial to the consumer (12). However, the salmon must still be able to utilize a wide array of prey as the energy density of these species can change throughout the year (11).

1.1.3 Salmon Farming

Salmon production can be divided into three phases that resemble the life cycle of wild salmon. The first phase is broodstock production. Here, eggs from female fish and milt from male fish are collected from spawning brood fish and subsequently mixed for fertilization (13). In the second phase, the eggs are hatched in freshwater, and the fish goes into the fry stage. The purpose of this phase is to bring the fish to a certain size or biological state. The phase is over when the fish undergoes smoltification and can tolerate saltwater. The final phase, food fish production, is about raising and feeding the fish until it reaches a size that can be sold to the consumer (13).

The diet of wild and farmed salmon varies greatly. While wild salmon hunt various prey, farmed salmon are usually fed dry pellets (13). Commercial feed needs to contain an adequate nutritional composition, and fishmeal has frequently been used as the main source of protein (14). Recently, feed based on plant-meal, often containing soybeans, has been suggested as a substitute due to higher cost-efficiency. However, proteins and nutritional factors from plant-meals have an insufficient amino acid profile and are harder to digest for the carnivorous salmon (15). Several studies have reported that soybean protein feed may contribute to the development of intestinal disorders in salmonid fish by alternating the intestinal microbiota in the fish (14, 16).

1.2 The Fish Gut Microbiota

All vertebrates harbor complex microbial communities referred to as a microbiota (17). The microbiota can be defined as the group of microorganisms that reside within a specific habitat. An increased interest around the intestinal tract and its resident microbes has led to intriguing discoveries regarding this particular microbiota's functions. The fish gut microbiota has received less attention than that of mammals, but there is an increasing appreciation of the role it potentially has on fish health. The inhabiting microbes are thought to grant the fish various effects related to size, metabolism, feeding behavior, and immune response (4).

1.2.1 Colonization and Composition of the Salmon Gut Microbiota

Microbial colonization is thought to originate from the eggs, the surrounding water, and the first feed (18). Upon hatching, the sterile fish larvae take in the microbiota of the surrounding environment, which thus become the first colonizers of the fish intestine. The gut microbial community of newly hatched larvae tend to contain few bacteria, but then becomes diversified through feeding (18). Furthermore, the microbial community is found to change with life stage and habitat, indicating that the environment performs a significant part in influencing the subsequent colonization (4).

The fish gut microbiota is diverse and comprises fungi, yeasts, viruses, and members of the Bacteria and Archaea kingdoms (19). Bacteria are, however, dominating in the fish intestine (18). Presented in Table 1.1 is an overview of selected publications investigating the salmon gut microbiota. A summarization of the most abundant phyla found in the salmon gut is included in this table. It is worth noting that the studies have investigated fish from various locations and life stages, as well as material from different gut compartments. The studies have also been conducted using different approaches for microbial identification.

Table 1.1: Overview of publications on salmon gut microbiota used in this thesis. *The table includes the most dominant phyla found in their research and information about the salmon's habitat and domestication.*

Location	Domestication	Dominant bacterial phyla	Reference
Norway Scotland	Farmed/ Wild	Tenericutes, Proteobacteria	Holben et al., 2002
Norway	Farmed	Proteobacteria, Firmicutes	Hovda et al., 2007
Chile	Farmed	Proteobacteria	Navarrete et al., 2009
Canada Ireland West Greenland	Wild	Proteobacteria, Tenericutes	Llewellyn et al., 2015
Norway	Farmed	Proteobacteria, Firmicutes	Gajardo et al., 2016
Scotland	Farmed	Firmicutes, Proteobacteria	Dehler et al., 2016
Norway	Farmed	Firmicutes, Proteobacteria	Rudi et al., 2017

As seen in the overview in Table 1.1, the phylum Proteobacteria is observed in the salmon gut in all studies. Tenericutes and Firmicutes are sporadically detected. However, only the most dominating phyla observed in the studies are listed in this table, and others have been detected in smaller amounts. Taken together, the researchers have discovered that bacterial colonizers in the salmon gut include the already mentioned phyla, but also Actinobacteria, Fusobacteria, Bacteroidetes, Cyanobacteria, Planctomycetes, and more (5, 20-25). There is, however, seldom consensus on the abundance or even presence of any of these phyla.

At genus level, the studies in Table 1.1 revealed that *Delftia*, *Aliivibrio, Pseudomonas* and *Photobacterium* belonging in phylum Proteobacteria (20, 21, 24, 25), *Weissella*, *Lactococcus* and *Lactobacillus* in Firmicutes (20, 24), and *Mycoplasma* in the Tenericutes phylum (5, 21, 23) were found abundantly in the salmon gut.

1.3 The genus *Mycoplasma*

1.3.1 Phylogeny

The genus *Mycoplasma* is part of the phylum Tenericutes and class Mollicutes (*Mollis*, soft; *cutis*, skin). The class encompasses 5 families, 8 genera, and over 150 species (26, 27). *Mycoplasma* is the largest and most important genus in the class, with more than 100 identified species (27).

Phylogenetic analyses revealed that mycoplasmas originate from low G+C, gram-positive bacteria in the *Bacillus-Lactobacillus-Streptococcus* phylogenetic branch (28). This branch was later divided into many, eventually leading to the genera known today: *Asteroleplasma*, *Anaeroplasma*, *Acholeplasma*, *Spiroplasma*, *Entomoplasma*, *Mesoplasma*, *Ureaplasma*, *Mycoplasma* (27).

1.3.2 Characteristics

Mycoplasmas possess some unique characteristics compared to other prokaryotes. Firstly, measuring only 0.3-0.8 µM in diameter, they are one of the smallest free-living bacteria observed (27). Secondly, they completely lack a cell wall around their cell membrane, which causes them to be gram-negative and provides them with natural resistance to antibiotics that target cell wall synthesis. Considering the cell membrane is the only barrier between the extracellular environment and the cell, a higher sensitivity to osmotic shock and detergents is common for this bacterium. Finally, mycoplasmas can change their shape and size, being so-called pleomorphic. They may therefore appear different than other bacteria and often occur as spherical-, fried-egg-, or flask-shaped when studied (27). Their ability to keep such shapes indicates the presence of a cytoskeleton in the cell, which, in addition to the cell membrane and cytoskeleton, is built up of ribosomes and a circular double-stranded DNA molecule. Depending on the species, the genome ranges from 580 to 1350 kilobase pairs (kb) (29).

1.3.3 Metabolism

Much due to their limited genome, mycoplasmas are unable to exhibit the same wide-ranged metabolic activities found in other bacteria (27). Thus, their metabolic activities appear to be primarily associated with generating energy rather than supplying substrates for biosynthetic pathways. Depending on the mycoplasmas' ability to use carbohydrates as an energy source, they can be grouped into either fermentative or non-fermentative organisms (27).

Fermentative mycoplasmas usually differ in their ability to utilize other sugars than glucose, and some species favor fructose over glucose (27). Sequencing projects revealed that two of the most known species, *M. pneumoniae* and *M. genitalium*, carried all the enzymes of the Embden-Meyer-Parnas pathway (30, 31). However, the second pathway for metabolizing glucose (the pentose phosphate shunt) was truncated. Additionally, several of the enzyme activities related to the tricarboxylic acid (TCA) cycle are not known in Mollicutes, causing them to lack a complete TCA cycle (32). Pyruvate generated from glycolysis is further metabolized to either lactate or acetyl-CoA, which decreases the pH in the growth medium (27). Most of the non-fermentative species possess the arginine dihydrolase pathway. Hydrolysis of arginine produces ornithine, adenosine triphosphate (ATP), CO_2, and ammonia as end products, which increases the pH of the growth medium (33). Some mycoplasmas metabolize neither sugars nor arginine but can oxidize organic acids such as lactate and pyruvate to acetate and CO_2 (34).

None of the mycoplasmas investigated thus far possess any quinones or cytochromes, excluding oxidative phosphorylation as a mechanism for generating ATP (35). Consequently, the mycoplasmas' available energy-yielding pathways only produce low quantities of ATP, in addition to relatively high amounts of metabolic end products (27).

1.3.4 Ecology and Habitat

Mycoplasmas are ubiquitous in nature and found in humans, mammals, reptiles, fish, arthropods, and plants (27). Novel species are frequently discovered, along with previously uncharted habitats. The most common habitats of human and animal mycoplasmas are the mucous surfaces of the respiratory, urogenital, and gastrointestinal tracts (36). However, they are also found in the eyes, mammary glands, or joints. Moreover, mycoplasmas have been reported as being present in artificial habitats, for instance, contaminating cell cultures in a laboratory environment (27). Both pathogenic and commensal strains occur, and a mycoplasma-containing flora has been found in farm animals such as sheep, cattle, and horses, in household pets such as cats and dogs, and even in wild animals including elephants, turtles, and fish (27).

1.3.5 Pathogenicity

Mycoplasmas usually live with their host as commensals (27). In cases they are pathogenic, they generally follow a slow chronic course rather than that of an acute infection. Many pathogenic species reside within animals, where they are the causative agent of, e.g., pleuropneumonia, mastitis, and conjunctivitis in cattle, goats, and sheep, and chronic respiratory disease and arthritis in swine, chicken, and laboratory animals (27).

Most human and animal mycoplasmas are surface parasites, which means they usually adhere to the host's epithelial cells without invading surrounding tissues (37). Attachment is achieved by a specialized tip structure in cooperation with adhesion- and accessory proteins that facilitate movement and polarity (38). The most studied adhesin proteins are those of *M. pneumoniae*, called P1 and P30, and *M. genitalium*, called MgPa (37, 39). Adherence to epithelial mucosa is an essential virulence factor, and should the mycoplasma lose its ability to adhere; it consequently loses its infectivity (27). Furthermore, some species, such as *M. penetrans*, appear to not only adhere but to enter the eukaryotic host cells using the tip structure (40). Additionally, mycoplasmas' lack of a cell wall has been suggested to facilitate direct contact of the mycoplasma membrane with that of its eukaryotic host (27). This condition may lead to a fusion of the two membranes, enabling transfer or exchange of mycoplasmal cell components into the host cell.

There are several possible mechanisms of damage to the host once the mycoplasmas have adhered or entered the host cells. Because of their limited genes and energy for biosynthesis, they are highly dependent on the host. The host cell machinery supplies biochemical precursors required to synthesize amino acids, fatty acids, cofactors, and vitamins (38). Providing the mycoplasmas with supplies can lead to disruption of the host cell, depleting it for necessary components for cell growth, protein synthesis, and other essential functions. Another mechanism for possible injury is the production of cytotoxic metabolites released by adhering mycoplasmas (27). Hydrogen peroxide and superoxide radicals generated by mycoplasmas have been thought to cause oxidative damage to the host cell membrane. Still, because of the small amounts excreted, and the presence of detoxifying enzymes in host cells such as catalase – which converts hydrogen peroxide into water and oxygen – this has not yet been confirmed (27).

1.3.6 Cultivation

Despite mycoplasmas' omnipresent growth, they have strict nutritional requirements and are quite sensitive to environmental factors (26). Demanding conditions make cultivation challenging and time-consuming in vitro, and growth usually happens at a slow rate (27). Gradual and poor growth in bacterial cultures emphasizes mycoplasmas' parasitic approach, and their need of a host organism to supplement their lack of essential genes.

1.4 Sequencing Approaches for Microbial Identification

1.4.1 Importance of Identifying and Characterizing Microbes

Studying the genetic material of all microbes in a sample – the microbiome – has gained much popularity in the research field. Classical microbiome research relies on cultivation, but the recent evolution of several culture-independent techniques has significantly improved the qualitative and quantitative identification of microbes (41). This advancement has led to the development of metagenomics, defined as the direct study of genomes in an environmental sample (42).

Identifying microbes in a sample is valuable in physiological and ecological contexts. In microbial ecology, characterization of the microbial community, e.g., in soil, water, skin, or gut, may elucidate the function the microbes have in the habitat they reside in (41). Determining the functional attributes of the microbiota associated with hosts is essential for understanding their role in host metabolism and disease (43). There are usually two general approaches to identify microbes using genetic material: targeted 16S ribosomal RNA (rRNA) gene sequencing and whole-(meta)genome shotgun sequencing.

1.4.2 Targeted 16S rRNA Gene Sequencing

Targeted gene sequencing focuses on a select set of genes or regions in the microbe's genome. For bacterial identification, the most commonly used target gene is the 16S rRNA gene. Several features make this gene a valid target. Firstly, it is thought to exist throughout the prokaryotic domain and is, therefore, present in all bacteria (44). Secondly, the gene function has remained constant during evolution, so mutations are presumed to represent random changes rather than alter the molecular function (45). Finally, the gene is built up of approximately 50 functional domains, which is important because mutational changes over time in one domain does not greatly affect the sequence information in other domains (44). Thus, as illustrated in Figure 1.1, the 16S rRNA gene consists of conserved regions, which have been constant throughout evolution, and variable regions, which vary between bacteria as time has passed and mutations have occurred.

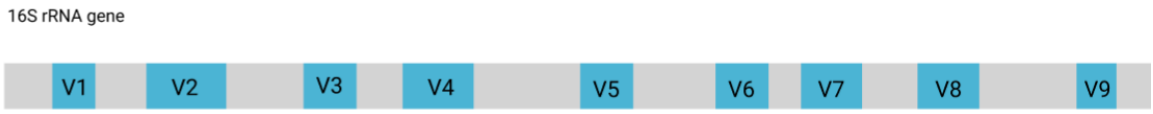

Figure 1.1: Illustration of the 16S ribosomal RNA gene in prokaryotes. *The gene is approximately 1550 bp long and consists of conserved (grey) and variable (blue) regions. The conserved regions have remained constant throughout evolution, while the variable regions vary between different bacteria as mutations have occurred. By sequencing the 16S rRNA gene, all microbes in a sample can be identified.*

In the process of 16S rRNA sequencing, amplification is achieved by designing primers that attach to the conserved regions of the gene (46). Utilizing universal primers allows the gene to be amplified for every prokaryotic organism in a sample. Regions containing both conserved and variable genetic information are thusly sequenced, and the sequence information in the variable regions of the gene is used to distinguish microbes from one another (46). However, 16S rRNA gene sequencing is not a perfect measure for bacterial identification (44). A limitation of this approach is the lacking ability to identify microbes at species level due to the method's low resolution. Additionally, as the 16S rRNA gene is nonexistent in viruses and fungi, other target genes must be utilized to obtain sequences from these microbes, e.g., marker genes specific for the virus investigated or the 18S rRNA gene in eukaryotes (47, 48).

1.4.3 Whole-(meta)genome Shotgun Sequencing

Application of whole-(meta)genome shotgun sequencing allows the entire genome(s) of the microbe(s) present in a sample to be sequenced. Acquiring whole-genome information allows for bacterial identification at a lower taxonomic level, as well as the detection of fungi, viruses, and novel microorganisms (49). The most efficient way to sequence a large DNA molecule is to break it into smaller pieces (50). As illustrated in Figure 1.2, the process of shotgun sequencing shears the DNA molecule randomly into short fragments, which thereafter are individually sequenced (51). Resulting sequences are analyzed by computer programs, scanning for identical regions in the fragments (50). When identical regions are identified, they are overlapped with one another, allowing the two sequence reads to be connected. The connection of overlapping fragments is executed by assembly algorithms, and generates a genomic sequence encompassing all the fragments in one long, known sequence, called a contig (51). Generally, multiple contigs are produced, separated by unknown sequences.

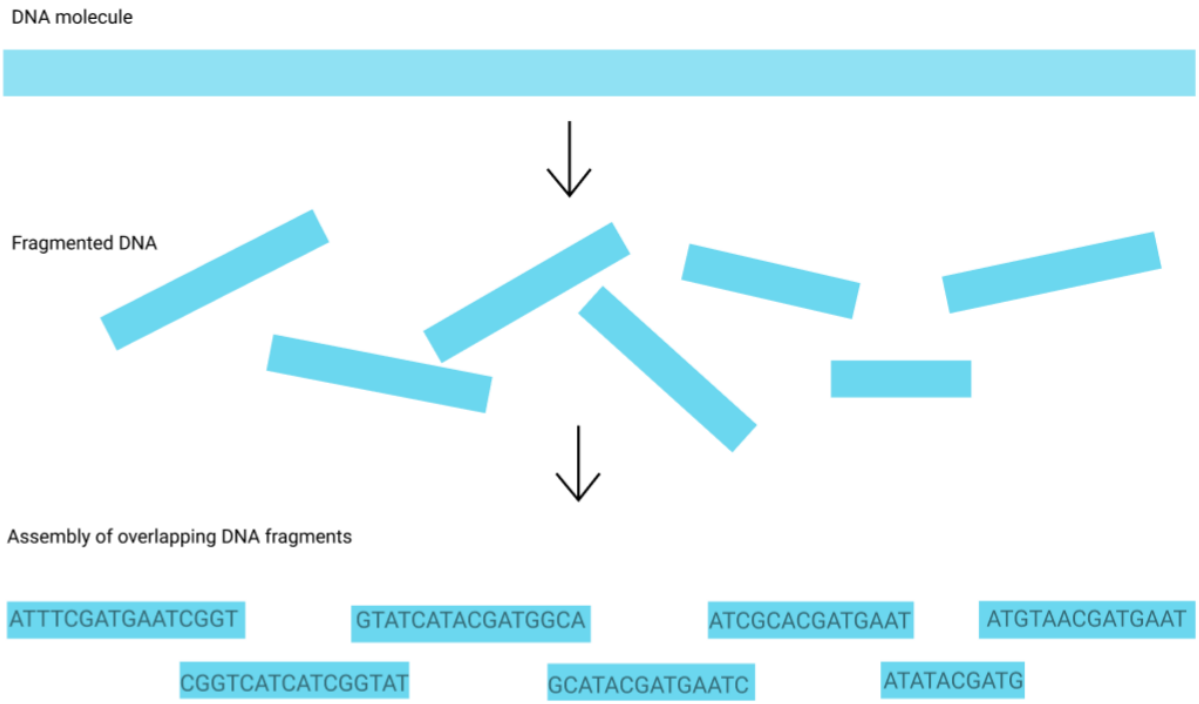

Figure 1.2: Illustration of shotgun sequencing. *In the process of shotgun sequencing, the large DNA molecule is sheared randomly into smaller fragments. The fragments are then individually sequenced and overlapping regions of the sequences are identified by computer programs. Computer programs are further used to reconnect the sequences into the correct order, representing the original DNA molecule.*

Since fragments are randomly sequenced, the approach requires a certain level of coverage to ensure that the majority of the original DNA molecule will be represented by the overlapping fragments (51). Thus, a large number of reads must be generated when using this method.

Choosing whether to apply 16S or shotgun sequencing depends on the nature of the study conducted (43). If the number of samples is high, and resolution is not of great importance, 16S sequencing is an efficient and low-priced approach. Shotgun sequencing, on the other hand, offers increased resolution, but is generally more expensive and requires more data processing. In addition to deciding whether the 16S or shotgun approach is preferred, it is also necessary to determine how the sequencing itself should be performed. Multiple commercial companies have developed their own sequencing platforms, each using a particular technology.

1.5 Sequencing Technologies

Over the last fifty years, researchers have invested time and resources in developing and improving technologies related to DNA sequencing (52). Innovations have led to an advancement in the amount of sequencing data being generated, and in addition to increased throughput, the cost and time spent has greatly decreased.

1.5.1 First-generation Sequencing

In 1977 Fred Sanger and his colleagues developed what is known as the chain-termination or dideoxy technique (53). It was considered a breakthrough as it greatly simplified the former sequencing techniques (52). Sanger's method of chain-termination is based on the incorporation of dideoxynucleotides (ddNTPs). These are chemical analogs to the monomers of DNA strands – deoxyribonucleotides (dNTPs) – but lack the 3' hydroxyl group. Without this group, no bond with the 5' phosphate group of the next dNTP can be formed, resulting in termination of DNA polymerization (54). When mixing radio-labeled ddNTPs in a polymerization reaction with standard dNTPs, the ddNTPs are randomly incorporated during strand extension, stopping further progress (52). DNA strands of each possible length are this way produced. By performing four parallel reactions containing each individual ddNTP base and run the results on four lanes of a polyacrylamide gel, the resulting DNA strands are distributed in the gel based on varying lengths. The nucleotide sequence of the original template can then be confirmed by autoradiography (52). Several changes have been applied to the method throughout the years. The most notable ones were replacing radio-labeling with fluorescent-labeling, allowing the reaction to happen in one container instead of four, and detection through capillary-based electrophoresis (52). Both improvements contributed to the development of automated Sanger sequencing machines which are commonly used today (55).

1.5.2 Second-generation Sequencing

Parallelization of sequencing reactions significantly increased the amount of DNA that could be sequenced in one run and was recognized as a paradigm shift in the study of genomics (52). The pioneer company that first utilized parallelization was 454 Life Sciences, later owned by Roche. Their technology uses adapter sequences to attach DNA molecules to beads, which then undergo

a water-in-oil emulsion PCR (emPCR) (56). The emulsion process creates a droplet ideally containing only one DNA template and one bead. DNA-coated beads amplify in their droplets during the PCR, and so-called pyrosequencing occurs as dNTPs are subsequently washed over the beads. When a dNTP is incorporated into the DNA strand, a pyrophosphate molecule is released (57). The released pyrophosphate molecule is then converted into ATP, which is further used as a substrate for luciferase, an enzyme that produces light proportional to the amount of pyrophosphate (58). During pyrosequencing, a detector is used to pick up the light emitted, and the light intensity is used to determine the number of dNTPs incorporated (57). If no light is emitted, the nucleotide on the template DNA strand is not complementary to the dNTPs currently washed over the beads. Any unused dNTPs are removed, allowing the process to be repeated with the other dNTPs until synthesis is complete, and the DNA sequence is determined.

After the advances of 454, other companies developed parallelization technologies. One technology gaining popularity was Solexa sequencing, today known as Illumina sequencing (59). Rather than parallelizing using emPCR, this technology utilizes a flow cell coated with oligonucleotides. Adapters ligated to the DNA molecules attach the DNA to complementary oligonucleotides on the flow cell (52). A PCR phase is used to amplify the original flow cell-binding DNA strands, creating clusters of replicates by "bridge amplification." The name comes from DNA strands having to arch over and bind to neighboring oligonucleotides in order to prime the next round of polymerization (59). Like Sanger sequencing and pyrosequencing, Illumina requires a DNA polymerase to produce an observable output (52). Fluorescent reversible-terminator dNTPs are used, where a fluorescent molecule occupies the 3' hydroxyl position making further extension impossible before the molecule eventually is cleaved off – hence the "reversible-terminator" term (60). The identity of the fluorescent reversible-terminator dNTP incorporated is detected by exiting the fluorescent molecule with an appropriate laser (52). After recording the nucleotide, the fluorescent molecule is removed by enzymatic reactions, and incorporation of a new fluorescent reversible-terminator dNTP can occur. Then, the process repeats in a stepwise manner, resulting in the DNA sequence of the original strand bound to the flow cell (52). An advantage of Illumina's technology is the utilization of paired-end reads (52). After the sequence read of the original strand is obtained, it is washed away, and a second round of DNA polymerization of the reverse strand takes place. Paired-end reads greatly improves the accuracy and credibility of sequences generated (52).

1.5.3 Third-generation Sequencing

Defining second- and third-generation techniques have been subject to discussion (52). However, publications suggest that third-generation sequencing should include features like single-molecule sequencing and real-time sequencing (61, 62). Pacific BioSciences (PacBio) has led the development of a single-molecule real-time technology (62). Such technology allows for the sequencing of non-amplified DNA, thus eliminating possible template errors caused by PCR (63). PacBio's technology is based on light detection at the bottom of zero-mode waveguide (ZMW) nanostructures, which are essentially wells embedded in a metallic film covering a chip (52). ZMW nanostructures utilize the property of light that causes it to decay exponentially as it passes through a hole smaller than its wavelength. When a fluorescently labeled dNTP is incorporated into the DNA template by a DNA polymerase attached to the bottom of the ZMW, a unique light pulse is produced that identifies the nucleotide (64). The signal will diminish relatively fast and not interfere with the fluorescent signal from the next dNTP being incorporated (52). Every light pulse is recorded, resulting in a sequence read complementary to the template DNA. PacBio's technology is widely used as the machines can produce long reads, over 10 kb in length, in a short amount of time.

A unique single-molecule technology with promising prospects is nanopore technology (52). Oxford Nanopore Technologies has successfully utilized synthetic nanopores in their platforms, in a system based on "lab on a chip" technology (62). Here, an array chip is coated with multiple wells, each containing a single protein nanopore. On top of these wells lies a lipid bilayer equipped with electrodes, which allows a voltage to be applied. When DNA is introduced to the chip, it is denatured by enzymes (52). Further, enzymes threads one of the strands through the nanopores. As the nucleotides of the DNA strand encounter the pore, the ionic flow that is applied to the chip is prevented. Each nucleotide disrupts the flow in a distinctive manner, and can this way be identified (62). The disruption signal is detected, and the order of nucleotides in the nucleic acids is recorded, providing a full sequence read of the template when finished. Neither amplification nor labeling is needed, making it an inexpensive and rapid sequencing technology (65).

1.6 Bioinformatic Software for Metagenome Analyses

Bioinformatical applications are tools used for processing data acquired from the different sequencing platforms. There are multiple applications available, which aids in taxonomic assignment, assembly of genomes, annotation of gene function, and more.

1.6.1 Processing 16S rRNA Sequence Data

Data obtained from 16S sequencing encompasses multiple reads of the 16S rRNA gene of all bacteria in a sample. Applications such as BLAST (Basic Local Alignment Search Tool) or QIIME (Quantitative Insights Into Microbial Ecology) allows alignment of obtained sequences (query sequences) to known reference sequences in a database, and are this way used to identify the individual microbes in a sample (66, 67). These applications also generate statistical information, which further aids the user in interpreting output data.

1.6.2 Processing Shotgun Metagenome Data

Data obtained from shotgun sequencing will generally be large and require more processing compared to 16S data. The application SPAdes (St. Petersburg genome assembler) analyzes the shotgun sequences and constructs contigs based on overlapping regions of the fragments (68). The contigs can be compared to reference sequences in databases in order to identify the organism(s) in a sample. MG-RAST (Metagenomic Rapid Annotations using Subsystems Technology) assigns taxonomy to sequences and compares input sequences to databases on both nucleotide- and amino acid levels (69). Additionally, MG-RAST produces functional assignments to sequences and can thus be used to investigate gene functions of the metagenome.

If the goal is to reconstruct the genome of all bacteria present in a sample, it is also possible to group contigs associated with a single organism in a process called binning. Binning allows individual genomes to be recovered from metagenome data and can be performed by an application called MaxBin (70).

1.7 Aim of Thesis

Although a novel *Mycoplasma* species was discovered as an abundant resident in the intestinal tract of Atlantic salmon almost 20 years ago, information is still lacking regarding its colonization and potential influence on the fish.

The principal aim of this project was to obtain and examine genomic sequence information from gut content of farmed salmon from different geographical regions, to both investigate the prevalence of *Mycoplasma* and to discover the potential role it has as part of the gut microbiota.

Methods to achieve these goals included cultivation, polymerase chain reactions (PCR), DNA quality and quantity checks, as well as various sequencing approaches. Detecting mycoplasmal DNA was achieved by quantitative PCR, gel electrophoresis, and Sanger sequencing. Investigating the bacterial composition in the salmon gut was done using 16S rRNA gene sequencing, while whole-genome shotgun sequencing was applied to obtain genetic information of the salmon-associated *Mycoplasma* species. Bioinformatical tools were thereafter used to process and interpret output data.

2. MATERIALS AND METHODS

2.1 Samples Used in this Project and Sample Flow

This work is an extension of a former project executed at the Microbial Diversity laboratory at NMBU. Therefore, the sampled gut content used in this project were collected on different occasions. Table 2.1 shows key information about the samples, while a full overview containing comprehensive sample information is found in Appendix A.

Table 2.1: Key information of samples used in this thesis. *Gut content from fish was collected from Norway and Chile on different occasions, illustrated by using different colors: Blue represents samples from Skjervøy; Orange/yellow samples from Bømlo; Green samples from Chile; and grey samples from Dagali. The table includes information about location, time of sampling, type of species, sample ID and storage buffer/medium used for the samples.*

Location	Time	Sample ID	Species	Description
Skjervøy[1]	Oct. 2018	i1-i20 (excluding no. 2, 3, 8, 13, 16)	Salmon	Non-pooled, individual samples stored on S.T.A.R. buffer
Skjervøy[1]	Mar. 2019	S1-S8	Salmon	Pooled samples stored on eight different storage buffers/mediums
Bømlo	Sept. 2019	B1-B19	Salmon	Pooled samples stored on MycoBroth (for cultivation purposes), RNA later and S.T.A.R buffer
Filtered Bømlo	Sept. 2019	F1-F10	Salmon	Primary bacterial cultures made by filtering Bømlo samples through a 0.45 µM filter and transferring filtrate to tubes containing fresh MycoBroth
Chile	Nov. 2019	C1-C10	Salmon	Samples sent from Chile by mail stored on RNA later
Chile	Feb. 2020	C11-C20	Salmon	Samples sent from Chile by mail stored on RNA later
Dagali[2]	Summer 2019	D1-D8	Trout	Individual intestines from wild trout in freshwater, stored on S.T.A.R buffer

[1] Samples from Skjervøy were obtained before this thesis was initiated and were processed by Laboratory Engineer Inga Leena Angell.

[2] Samples from Dagali were used for practice purposes.

Samples were collected for both cultivation purposes and direct DNA analyses. Therefore, the processing of samples differs somewhat. Figure 2.2 illustrates the workflow of the samples. DNA was extracted using different extraction protocols; QIAGEN, LGC Genomics, or the phenol-chloroform-isoamyl alcohol method. Following extraction, genomic DNA (gDNA) was quality controlled and quantified before preparing the samples for different sequencing approaches. Bioinformatic applications such as QIIME, BLAST, MG-RAST, and MaxBin were further used to process the sequencing output.

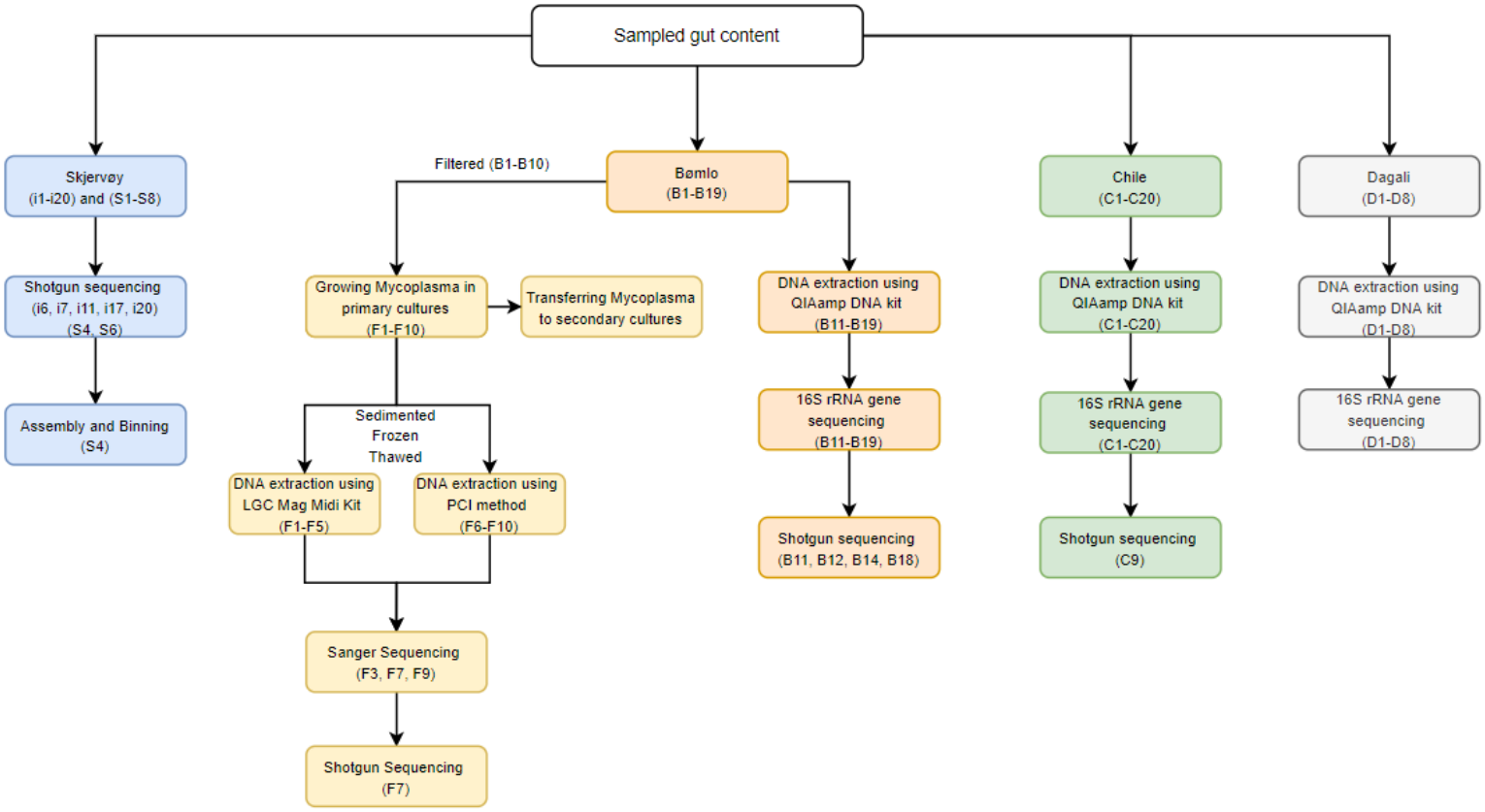

Figure 2.2: Flow chart showing the processing of fish gut content collected from different sites. *Each color represents a different site: blue being Skjervøy; orange/yellow is Bømlo; green is Chile; grey is Dagali. Sample identification is written in parenthesis and refers to individual samples as they are described in Appendix A and Table 2.1. Samples from Skjervøy (n = 23) were obtained before this thesis was initiated and were processed by Laboratory Engineer Inga Leena Angell. A selection of the Skjervøy samples (n = 7) was further processed for shotgun metagenome sequencing in this project. Samples from Bømlo (n = 19) were split based on the medium they were stored in: samples stored on Mycoplasma Growth Broth (n = 10) were filtered and cultivated before sequencing (yellow), while samples stored on storage buffer (n = 9) were directly prepared for 16S rRNA sequencing (orange). Selected samples from Bømlo (n = 4) were processed for shotgun sequencing (orange). Samples from Chile (n = 20) were directly prepared for 16S rRNA sequencing. One sample from Chile was further processed for shotgun sequencing. Trout samples from Dagali (n = 8) were used for practice purposes.*

2.2 Sampling Gut Content for Cultivation

2.2.1 Preparation of Enriched Mycoplasma Growth Broth and Agar Plates

Mycoplasmas have strict nutritional requirements, and the culture medium needs to contain the necessary components to support the growth and maintenance of the bacterium. An antibiotic, usually Penicillin, should also be present to inhibit the growth of gram-positive bacteria.

Mycoplasma growth broth (hereafter referred to as MycoBroth) was prepared by adding a volume of Mycoplasma broth powder (Sigma-Aldrich, USA) to Milli-Ro water in a 100 mL flask to a concentration of 2.55 g per 100 mL water. The broth powder was dissolved by stirring, and the pH adjusted to 7.8 ± 0.2 by adding NaOH. To create agar plates, 1.5% agar powder was added and dissolved by boiling. The flask was autoclaved at 121^{o}C and 15 psi for 15 min to ensure a sterile environment. After cooling the flask to 50-60^{o}C, 30 mL MycoBroth was removed and replaced with an equal volume of Mycoplasma enrichment supplement (Sigma-Aldrich, USA). DNA sodium from salmon testes (Sigma-Aldrich, USA) was added to the flask at a concentration of 20 mg/L. Further, agar plates were created by pouring the mixture to 5 x 1 cm Petri dishes in a sterile bench to avoid contamination. The final concentrations of the MycoBroth medium are listed in Appendix B.

2.2.2 Collecting Gut Content from Salmon

The prepared MycoBroth medium was transferred to sample tubes before the sampling of gut content. The medium allows microbial growth to continue after sampling.

An incision was made along the abdomen of the fish. The gut was removed from the abdomen, and a section was made right after the small intestine. Gut content was pressed towards the section, transferring raw material from each fish into the same empty sampling tube. The collected gut material was mixed before being distributed to individual sampling tubes containing MycoBroth. Tubes were kept cold (4-8^{o}C) during transport.

2.2.3 Cultivating Mycoplasmas in Mycoplasma Growth Broth and Agar Plates

Three series of cultures were set up to cultivate mycoplasmas: one primary and two secondary cultures. The primary cultures were made by filtering a volume of the gut samples through a 0.45 µm filter to ensure the removal of bacteria larger than mycoplasmas. The filtrate was supplemented with fresh MycoBroth medium and left to sit at room temperature. Microbial growth in the primary cultures was regularly checked. When setting up secondary cultures, a volume of the primary culture was 1) transferred to fresh MycoBroth medium, and 2) frozen in Cryotubes for two days before transferring a volume to fresh medium. Additionally, one of the primary cultures was plated on Petri dishes containing MycoBroth agar.

The cultures were kept for approximately 1.5-2 months before it was decided to freeze the cells in order to preserve them. The samples were centrifuged at 7500 rpm for 5 min to sediment the cells. DNA concentration of the pellets was measured before storing them at -80^{o}C awaiting DNA extraction. The supernatants were kept for pH analysis.

2.2.4 Measuring pH value of Primary Cultures

The mycoplasma species' ability to use carbohydrates as an energy source determines their fermentative or non-fermentative character. Fermentative mycoplasmas produce metabolites from glycolysis that decrease the pH of the medium, whereas non-fermentative species produce end-products that increase the pH value.

MColorpHast^{TM} pH strips (Merck KGaA, Germany) were used according to the manufacturer's instructions to determine the pH value of the supernatants derived from the primary MycoBroth cultures.

2.3 Sampling Gut Content for DNA Analyses

2.3.1 Collecting Gut Content from Salmon

S.T.A.R (Stool Transport and Recovery) buffer (Roche, Switzerland) and RNA later (Invitrogen, USA) were transferred to sample tubes before the sampling of gut content. The two mediums reflect the gut microbiota of the fish at the time of sampling.

The extraction of gut content was done as previously described in Chapter 2.2.2. However, at some of the sites, gut content was transferred directly into individual sampling tubes containing medium without pooling the gut content in the same tube. See Appendix A for information about the different samples. Tubes were kept at room temperature or cold (4-8^{o}C) during transport.

2.4 Extracting DNA for Downstream DNA Analyses

DNA extraction and purification are necessary to ensure that contaminations such as proteins, RNA, or other disturbing molecules are removed from the final product. Contaminants may affect downstream analyses and cause unwanted biases, impairing the results.

2.4.1 QIAamp DNA Mini Kit

The QIAamp DNA Mini Kit (QIAGEN, Germany) uses enzymatic tissue lysis of the cells. Released DNA binds to the silica-gel membrane in a spin column while contaminants are washed off. Finally, the purified DNA is eluted in water or buffer.

From the gut samples, 1 mL was transferred to 1.5 mL Eppendorf tubes and centrifuged at 7500 rpm for 5 min. Samples stored on RNA later were not properly pelleted, and in order to thin out the reagent, ice-cold 1 x Phosphate-buffered saline (PBS) was added to a 1.6x concentration of the sample. The tubes were centrifuged at 10 000 rpm for 5 minutes to sediment the cells. Hereafter they were treated identically to the other samples.

Following centrifugation, the supernatant was discarded, and Buffer ATL (tissue lysis buffer) was added to a total volume of 180 µL. Further, to degrade proteins in the samples, 20 µL Proteinase K was added. The tubes were set to incubate at 56^{o}C with shaking at 1000 rpm for 1

hour. After a brief spin down to remove drops from the lid, 200 µL Buffer AL (lysis buffer) was added. The mixture was then pulse-vortexed for 15 seconds before incubating the samples at 70°C for 10 min.

In order to enhance DNA binding to the spin column, 200 µL ethanol (96%) was added, and the mixture was pulse-vortexed for 15 sec. The tubes were spun down before transferring the content to QIAamp Mini spin columns sitting in collection tubes. The columns were centrifuged at 8000 rpm for 1 min. After placing the columns in new collection tubes, DNA was washed by adding 500 µL Buffer AW1 and centrifuging the samples at 8000 rpm for 1 min. The washing step was repeated, using 500 µL Buffer AW2 and centrifugation at 14 000 rpm for 3 min. Once more, the columns were placed in new collection tubes and centrifuged at 14 000 rpm for 1 min to avoid Buffer AW2 carryover.

Finally, the columns were placed in Eppendorf tubes, and 200 µL Buffer AE (elution buffer) was added to elute the DNA. The columns were incubated at room temperature for 5 min before centrifuging them at 8000 rpm for 1 min. In some cases, the elution step was repeated to increase the yield further. The eluate was stored at -20°C.

2.4.2 Mag Midi LGC kit

The Mag Midi LGC kit (LGC Genomics, UK) utilizes superparamagnetic particles to capture nucleic acids from a sample. The nucleic acid/particle complex is then washed to remove contaminants before the nucleic acid is eluted from the particles in a buffer.

Cell pellets were resuspended in 300 µL S.T.A.R buffer. The samples were prepared for mechanical lysis by bead-beating by transferring them to specialized crushing tubes (SARSTEDT, Germany) containing approximately 0.2 g of both 9-13 µm and 100 µm glass beads (Sigma-Aldrich, Germany). The samples were processed twice in the MagNAlyser (Roche, Switzerland) for 20 sec at 6500 rpm before centrifuging them at 13 000 rpm for 10 min to collect supernatants. The supernatants were mixed with 50 µL Lysis buffer BLm and 5 µL Protease before incubation at 55°C for 10 min.

After cooling the samples down to room temperature, 50 µL ethanol and 16 µL Mag particle suspension BLm was added. All samples were then incubated for 2 min at room temperature to

allow sufficient binding. The samples were placed on a magnet, and the supernatant was removed and discarded. After removing the samples from the magnet, the pellet was washed by adding 170 µL Wash buffer BLm 1. The samples were incubated with shaking for 10 min, ensuring thorough washing. The samples were placed back on the magnet, and the supernatant discarded. The washing step was repeated twice using Wash buffer Blm 2.

After discarding the supernatant, the pellet was air-dried at 55°C for 6 min. 63 µL Elution buffer BLm was used to resuspend the pellet, and the samples were incubated at 55°C for 10 min while vortexing regularly. After pelleting the particles on the magnet, the eluate was transferred to a new Eppendorf tube and stored at -80°C.

2.4.3 Phenol-Chloroform-Isoamyl Alcohol Method

The phenol-chloroform-isoamyl alcohol DNA extraction method separates molecules based on their solubility in organic solution or water. The protein and lipid components of the cell are denatured and removed by separating them from the DNA. The DNA, which is soluble in water, can then be retrieved from the water phase.

Cell pellets were resuspended in 500 µL Tris-NaCl-EDTA buffer (0.01 M). The cells were lysed by adding 10 µL of both 10% sodium dodecyl sulfate and 10% N-lauryl sarcosine sodium salt. Additionally, 10 µL Proteinase K was added before incubating the lysate at 37°C for 1.5 hours. RNAse was added to a concentration of 100 µg/mL, and the mixture was incubated at 37°C for an additional half-hour. Further, 500 µL Phenol was added, and the samples were centrifuged at 13 000 rpm for 10 min. The aqueous phase was transferred to a new 1.5 mL Eppendorf tube, where 500 µL phenol-chloroform-isoamyl was added. The mixture was again centrifuged at 13 000 rpm for 10 min, and the step repeated.

For DNA precipitation, 40 µL Sodium acetate was added together with 800 µL ethanol (96%) before freezing the mixture at -20°C for 16 hours. DNA was sedimented by centrifuging the tubes at 13 000 rpm for 10 min. The supernatant was discarded, and the pellet washed by adding 500 µL ethanol (80%) before centrifugation as previously. The supernatant was discarded, and the DNA air-dried for 2-5 min. The tubes were stored at -80°C.

2.5 Quantification and Qualification of Extracted DNA

2.5.1 Measuring DNA Concentration on Qubit Fluorometer

The amount of extracted DNA must be satisfactory before sequencing. The gDNA quantity was therefore measured using Quant-iT High-Sensitivity dsDNA Assay Kit (Thermo Fisher, USA). This assay is highly selective for DNA in the range of 0.2-100 ng.

A Quant-iT Working Solution was made by diluting the Quant-iT reagent 1:200 in Quant-iT buffer. For each sample, 200 µL of Working Solution is required. The amount of Working Solution was prepared according to the number of samples. Assay Tubes were then prepared by adding 190 µL Working Solution + 10 µL Standard for standards, and 198 µL Working Solution + 2 µL user sample for the samples. All tubes were vortexed for 2-3 seconds and incubated for 2 min. The Qubit Fluorometer (Invitrogen, USA) was used to detect the fluorescent signal.

2.5.2 Measuring DNA Fragment Size by Gel Electrophoresis

An agarose gel electrophoresis was performed to verify the quality of gDNA. In an Erlenmeyer flask, the gels were prepared by adding agarose powder to 1 x Tris-Acetate-EDTA buffer at a concentration of 1:100. The mixture was heated to near boiling until becoming transparent. In order to visualize the DNA fragments, Peq-green (VWR Peqlab, Germany) was added at a concentration of 2 µL per 50 mL.

Purple Loading Dye (New England BioLabs, USA) was mixed with 6x purified DNA before loading the gel. The gel was run for 35-40 minutes at 80 volts. As DNA is negatively charged, the fragments travel towards the positive electrode of the gel chamber. The smaller DNA fragments are less hindered by the resistance of the gel and thus travel farther. The fragments appear as bands that were visualized by UV light using the Gel Doc™ XR instrument (Bio-Rad, USA).

2.6 Polymerase Chain Reactions and DNA Sequencing

2.6.1 Quantitative PCR

Quantitative PCR (qPCR) was performed on extracted gDNA using *Mycoplasma* specific primers (MycoRev and MycoFrw) and 16S primers (341F and 806R) to compare the amount of mycoplasmal DNA to other prokaryotic DNA. Additionally, 18S rRNA primers (3NDF and V4EukR2) were used to quantify eukaryotic DNA in the samples. Two parallels of each sample were set up.

1x HOT FIREPol EvaGreen® qPCR supermix (Solis BioDyne, Estonia), 0.2 µM forward- and reverse primer, and nuclease-free water (VWR, USA) were added to 1 µL template DNA to a total volume of 20 µL. The PCR was performed in the thermal cycler of a CFX96^{TM} Real-Time System (Bio-Rad, USA). The reactions containing 16S and *Mycoplasma* primers were amplified using the same conditions, starting with an initial denaturation stage at 95^{o}C for 15 min, followed by 40 cycles of denaturation at 95^{o}C for 30 sec, annealing at 55^{o}C for 30 sec and extension at 72^{o}C for 45 sec. The reaction containing 18S primers had similar thermal cycling conditions, except for the annealing step, which was set to 59^{o}C due to higher primer efficiency at this temperature.

Fluorescence was measured by the optics module of the CFX96^{TM} Real-Time System. Raw data was exported to the Bio-Rad CFX Maestro software and included information such as Ct-values and melting curves.

2.6.2 Sanger Sequencing

Sanger sequencing was performed to identify and verify the presence of bacteria in selected Primary MycoBroth cultures.

1x HOT FIREPol DNA Polymerase RTL (Solis BioDyne, Estonia), 0.2 µM forward- and reverse primer, and nuclease-free water were added to 4 µL template DNA to a total volume of 50 µL. The PCR was performed in a thermal cycler (Applied Biosystems, USA). The samples were amplified under the thermal cycling conditions of initial denaturation at 95^{o}C for 15 min, followed by 30 cycles at 95^{o}C for 30 sec, 55^{o}C for 30 sec, and 72^{o}C for 1 min 20 sec.

The PCR products were cleaned with AMPure XP beads (Beckmann Coulter, USA) at a 1x concentration of the template DNA. Clean-up is necessary for the removal of primer dimers and excess nucleotides from the reaction. 10 µL clean PCR product was distributed equally in two Eppendorf tubes, one containing 5 µL forward primer, and the other 5 µL reverse primer, matching the primers used during the PCR. The Eppendorf tubes were labeled and sent to Eurofins, Germany, for Sanger sequencing.

2.6.3 16S rRNA Sequencing by Illumina MiSeq

The 16S rRNA gene was sequenced using the Illumina MiSeq platform. Following Illumina's recommendations, two PCRs were executed, one for amplification of the 16S rRNA gene and one for adding primer indexes to the amplified DNA. The reactions contained polymerase, 0.2 µM forward- and reverse primer, and nuclease-free water. Different polymerases were added for the two PCR steps: 1x HOT FIREPol Blend Master Mix RTL (Solis BioDyne, Estonia) for the first PCR, and 1x FIREPol Master Mix RTL (Solis BioDyne, Estonia) for the second.

In the first PCR, 341F forward primer and 806R reverse primer were used to target and amplify the V3-V4 regions of the 16S rRNA gene. Initial denaturing occurred at 95^{o}C for 15 min, followed by 25 cycles at 95^{o}C for 30 sec, 55^{o}C for 30 sec, and 72^{o}C for 45 sec, before a final extension step for 7 min at 72^{o}C. The PCR products were cleaned with AMPure XP beads at a 1.5x concentration of the template DNA. The cleaned PCR products were used in the second PCR step, where Illumina adapter indexes attaches to the primers added in the previous PCR. Indexes are necessary to identify and associate the reads generated to the correct sample. Unique reverse and forward primers were added, careful not to use the same combination twice. The product was amplified using the same cycling conditions as in the previous PCR, except initial denaturing was set to 5 min, followed by 10 cycles. Additionally, the annealing was doubled to 1 min.

DNA quantity of the PCR products was measured using the Qubit Fluorometer. Normalization and pooling of the samples were done accordingly. AMPure XP beads at a 1x concentration were used to clean the pooled library before storing it in the freezer awaiting sequencing. Further preparations and loading on the Illumina MiSeq platform were executed by the laboratory personnel.

2.6.4 Shotgun Metagenome Sequencing by Illumina HiSeq

Selected samples were prepared for shotgun metagenome sequencing using the Nextera Flex DNA Kit (Illumina, USA). This kit utilizes a bead-based transposome complex to tagment gDNA. Tagmentation is a process that, in one step, both fragments and tags the gDNA with adapter sequences.

Bead-Linked Transposomes and Tagmentation Buffer are responsible for the tagmentation process. Samples containing these components were placed in a thermal cycler for 15 min at 55^{o}C. Immediately after tagmentation, Tagmentation Stop Buffer was added, and the mix was incubated at 37^{o}C for 15 min. The products were cleaned twice with Tagment Wash Buffer. A PCR step for index ligation and amplification was performed by adding enhanced PCR Mix and nuclease-free water to the cleaned tagmented gDNA. Index primers i5 and i7 were added, creating unique combinations. The PCR was performed under the thermal conditions of two pre-PCR steps, one at 68^{o}C for 3 min and one at 98^{o}C for 3 min, followed by X cycles at 98^{o}C for 45 sec, 62^{o}C for 30 sec and 68^{o}C for 2 min, before a final step at 68^{o}C for 1 min. The number of cycles depended on the total DNA input (ng) and varied from 5-12 cycles. The library was cleaned using Sample Purification Beads at a concentration of 1.8x the template DNA.

Library normalization and pooling were done based on DNA quantification measurements by the Qubit Fluorometer. The final library was measured on Qubit Fluorometer before sending it to Norsk Sekvenseringssenter, Norway, for shotgun sequencing on the Illumina HiSeq platform.

2.6.5 Primer Overview

Primers used for PCR are presented in Table 2.2.

Table 2.2: An overview of primers and their target sequence used for PCR.

Primer	Sequence (5'-3')	Target region/gene	Reference
PRK341F PRK806R	CCTACGGGRBGCASCAG GGACTACYVGGGTATCTAAT	The V3-V4 region of the 16S rRNA gene	Yu et al., 2005
MycoFrw MycoRev	GCAATCCCGCGTGAATGAATG CCTTCGCCTCTGGTGTTCTT	Mycoplasma specific 16S rRNA gene	In house
3NDF V4EukR2	GGCAAGTCTGGTGCCAG ACGGTATCTRATCRTCTTCG	V4 region of eukaryotic 18S rRNA gene	Cavalier-Smith et al., 2009 Bråte et al., 2010

2.7 Data Analyses

2.7.1 Sanger Sequence Data

Files containing Sanger sequence data in FASTA format were received from Eurofins, Germany. The data included sequence information of both the forward and reverse primer. The sequences were identified using BLAST, choosing the "Highly similar sequences" option and nucleotide (nt/nr) database. The match with the lowest E-value was reported in this thesis. The E-value is a calculated parameter that helps evaluate whether the query sequences are, in fact, homologous to the reference sequences, or if the alignment happened by chance (66). In general, E-values will be low if the query sequence and reference sequence are very similar.

2.7.2 16S rRNA Sequence Data

The QIIME pipeline was used to process the 16S rRNA sequence data. The sequences were filtered to ensure sufficient quality, and samples with less than a set number of sequences were discarded. The sequences were clustered at a 97% homology level and aligned to the SILVA database to create an OTU (Operational Taxonomic Unit) table. Additionally, QIIME provided Principle Coordinate Analysis (PCoA) plots, which are used to explore and visualize similarities and dissimilarities in the data set. QIIME automatically generates plots based on, e.g., Binary Jaccard index and UniFrac distance.

Lastly, unpaired T-tests were used to determine if there was a statistical significance between data sets obtained from 16S rRNA gene sequencing. The unpaired T-tests were conducted in Excel using the two-tailed and unequal variance options.

2.7.3 Shotgun Metagenome Sequence Data

For taxonomic assignment, shotgun metagenome data was uploaded to MG-RAST servers. An Excel file was created containing metadata about the samples. MG-RAST automatically performed a quality check, discarding sequences inadequate for analysis. Shotgun sequences were aligned to the M5nr database using a minimum of 60% percent identity and E-value 0.00005. The database contains sequences and annotations from multiple available databases (71). Data obtained from MG-RAST was visualized by creating bar graphs and pie charts using Excel.

Genome assembly and binning were executed by secondary supervisor Lars-Gustav Snipen. When processing the shotgun metagenome data, shotgun sequences were assembled into contigs using the software SPAdes. Further, the assembled contigs were binned using MaxBin. The FASTA file received from Lars-Gustav Snipen containing all binned contigs was uploaded to MG-RAST servers to classify the bins. Contigs were aligned to the M5nr database using a minimum of 60% percent identity and E-value 0.00005. Pie charts were created in Excel using data obtained from MG-RAST's taxonomic assignments.

3. RESULTS

3.1 16S rRNA Analyses

Determining the bacterial composition in the fish gut was done by sequencing the 16S rRNA gene and analyzing the sequences in QIIME. A varying number of reads were generated for the individual samples during the sequencing runs, ranging from 9502 to 174 253. After filtration samples with less than 10 000 reads (n = 2) were discarded to ensure adequate quality. In total, 3 013 170 sequences were generated during the sequencing runs. The sequences were clustered at a 97% homology level and aligned to the SILVA database to create an OTU table comprising 1366 different OTUs.

3.1.1 Comparison of Gut Composition at Phylum Level

Taxonomic composition in farmed salmon from Bømlo and Chile is presented at phylum level in Figure 3.1. Overall, the gut microbiota showed little diversity, and only a few phyla were detected. The phylum Proteobacteria dominated at both sites, comprising 97% of the gut microbiota in Chilean salmon and 66% of the microbiota in salmon from Bømlo. The differences in means were calculated by an unpaired t-test and revealed that Proteobacteria was significantly less abundant in Bømlo samples ($p = 0.00019$). The phylum Tenericutes was the second most abundant phyla found, comprising 30% of the Bømlo samples. Tenericutes was found in negligible amounts in the Chilean samples (0.05%) and was significantly less established in these fish ($p = 0.00023$). The output derived from Excel's t-test analyses on phylum level can be viewed in Appendix C.

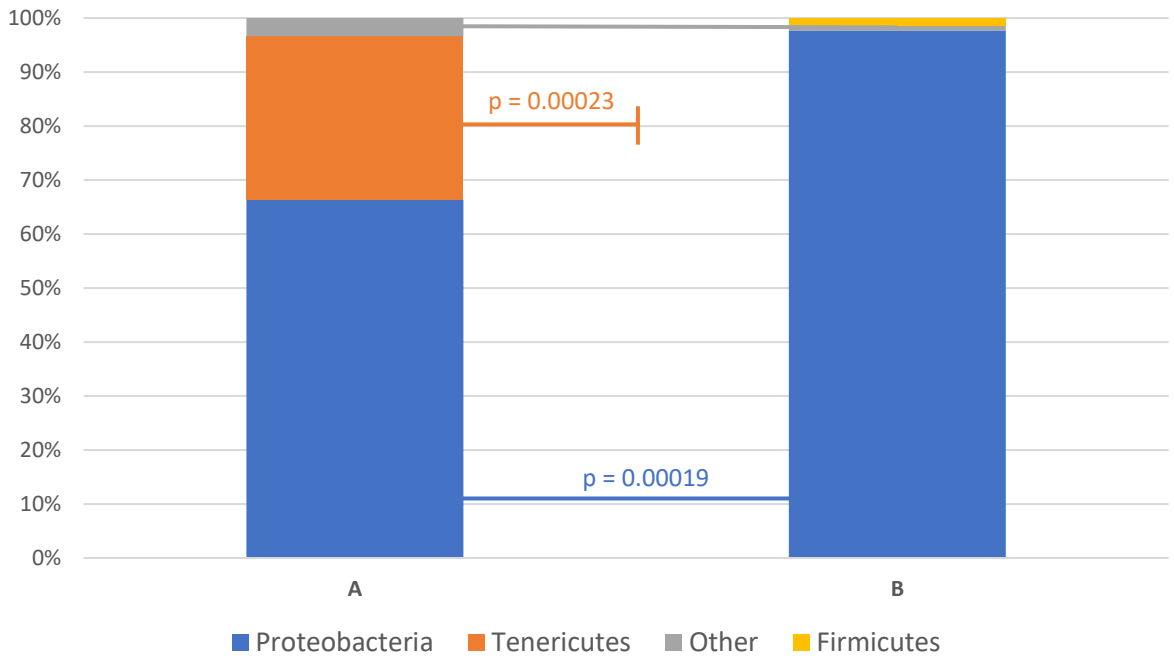

Figure 3.1: Overall gut microbiota composition in farmed salmon at phylum level. *Gut content collected from salmon located in Bømlo, Norway (A) and Chile (B) were investigated using 16S rRNA gene sequencing. The lines link phyla shared across sites. P-values generated from unpaired t-tests are included in the figure. Phyla below an abundance of 1% are not shown but summarized in the "other" group. Graphics were created in Excel using data derived from QIIME.*

3.1.2 Comparison of Gut Composition at Genus Level

Taxonomic composition in farmed salmon from Bømlo and Chile is presented at genus level in Figure 3.2. The analyses revealed that *Aliivibrio* (73%), *Photobacterium* (14%), and *Vibrio* (6%), all belonging to phylum Proteobacteria, were most abundant in the gut of Chilean salmon. In the gut of Bømlo salmon, *Aliivibrio*, *Photobacterium*, and *Mycoplasma* were equally abundant, making up approximately 30% of the microbiota each. Some *Vibrio* (7%) was also observed in these samples. Unpaired t-tests were used to determine significant differences in means and revealed significant differences in the prevalence of *Mycoplasma* ($p = 0.00023$) and *Aliivibrio* ($p = 0.000078$). No significant differences between sites were detected for *Photobacterium* and *Vibrio*. The output derived from Excel's t-test analyses on genus level can be viewed in Appendix C.

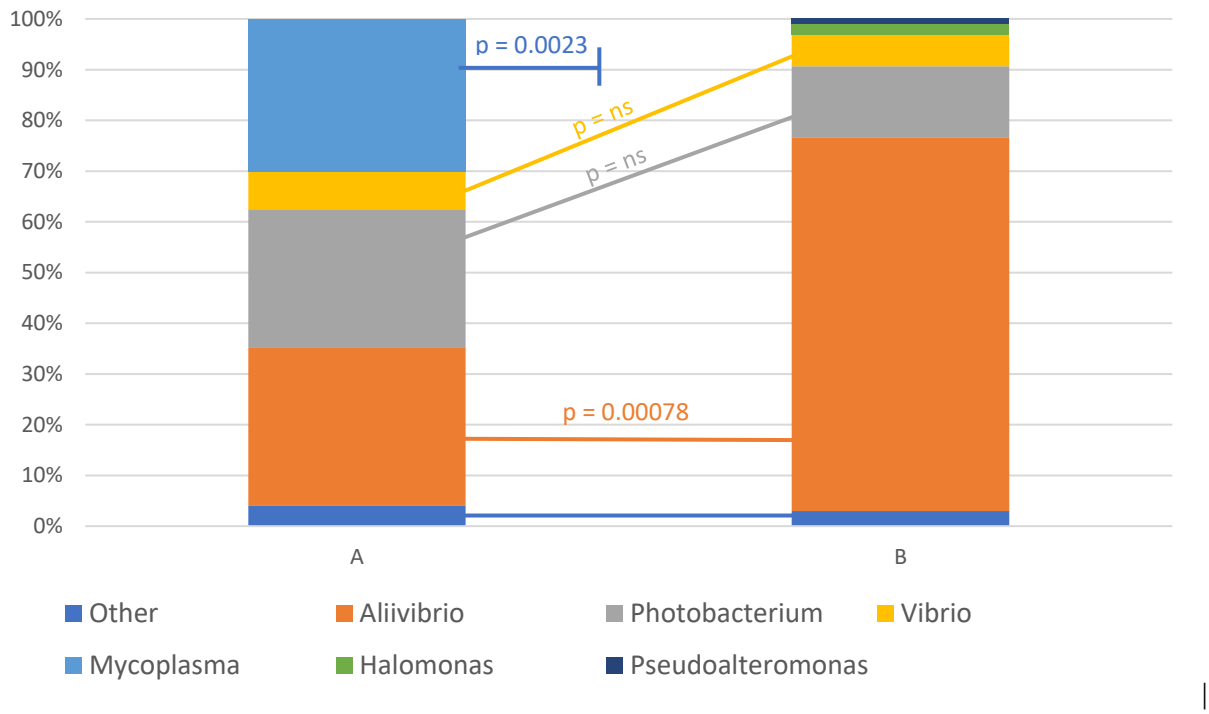

Figure 3.2: Overall gut microbiota composition in farmed salmon at genus level. *Gut content collected from salmon located in Bømlo, Norway (A) and Chile (B) were investigated using 16S rRNA gene sequencing. The lines link phyla shared across sites. P-values generated by t-tests are included in the figure, and ns means there was no statistical significance between the data sets. Genera below an abundance of 1% are not shown but summarized in the "other" group. Graphics were created in Excel using data derived from QIIME.*

3.1.3 Microbial Community Comparison by Principal Coordinate Analysis

Diversity between the microbial communities in farmed salmon from Bømlo and Chile were analyzed using UniFrac and Binary-Jaccard PCoA plots derived from QIIME.

The weighted UniFrac PCoA plot (Figure 3.3A) showed clustering of the Chile samples (blue). However, Bømlo samples (red) were moderately scattered, and one of the Bømlo samples overlapped with the Chilean cluster. Still, a separation of the microbial communities in farmed salmon from Bømlo and Chile was observed.

In the Binary-Jaccard plot shown in Figure 3.3B, the clustering was more distinct, clearly separating the two locations with no overlapping between samples. This was also observed in the unweighted UniFrac PCoA plot, which can be found in Appendix D.

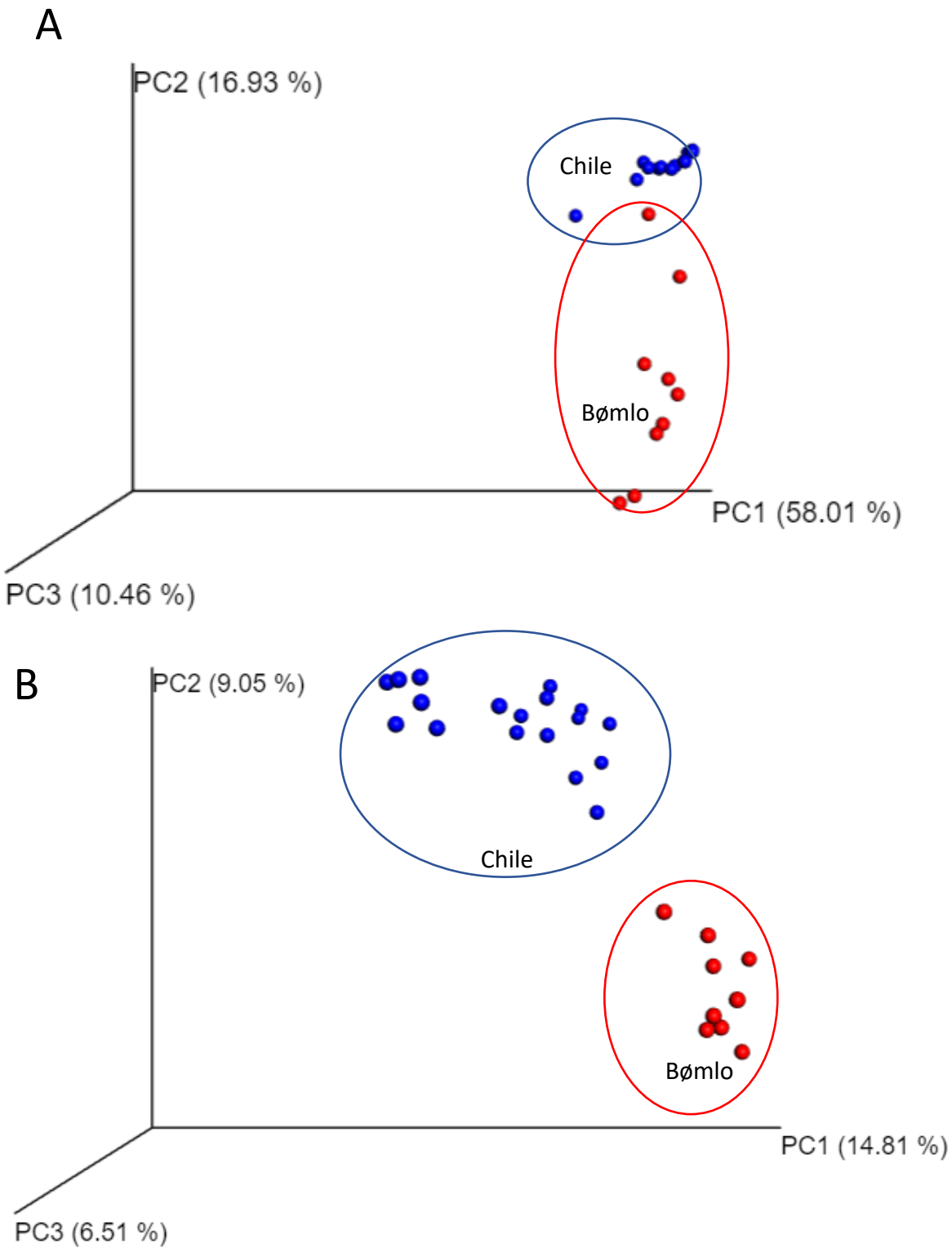

Figure 3.3: Beta-diversity between microbial communities in the gut of farmed salmon*. Differences in the gut microbiota of salmon from Bømlo, Norway, and Chile were analyzed by Principal Coordinate Analysis (PCoA) plots generated by QIIME. Panel A shows a weighted UniFrac PCoA plot, while panel B shows Binary-Jaccard distance. Each dot represents one sample; red dots represent salmon from Bømlo, and blue dots salmon from Chile.*

3.2 Shotgun Sequencing Analyses

Shotgun metagenome sequencing was performed to investigate the metagenomic composition of selected samples and to obtain genetic information of *Mycoplasma*. A selection of samples from Skjervøy and Bømlo in Norway, and Chile was processed for shotgun sequencing. Selected samples can be viewed in Appendix A and are expressed in bold writing. The obtained shotgun sequences were assigned taxonomy, and the sequences from one Skjervøy sample were further assembled and binned.

3.2.1 Taxonomic Annotation of Sequences by MG-RAST

The obtained shotgun data was analyzed by the MG-RAST application to identify DNA in the samples. After the application's built-in quality check, a total of 489 516 336 DNA sequences were obtained, with an average of 30 594 771 sequences per sample.

Figure 3.4 shows how the most abundant genera were distributed across the three sites. Large portions of the obtained sequences were classified as eukaryotic DNA. *Salmo* (salmon), *Danio* (zebrafish), and *Schistosoma* (parasitic flatworm) were recovered consistently and abundantly, and together comprised more than 40% of the sequences from each site.

Compared to eukaryotic DNA, bacterial sequences were less abundant. As seen in Figure 3.4, the most abundant bacteria detected in the Skjervøy samples was *Mycoplasma* (16%). Individual differences were observed, however, and *Mycoplasma* varied from 2-52% between samples. *Mycoplasma* was also the most abundant bacteria recovered from Bømlo samples (4%). However, the individual differences were smaller, ranging from 2-7%. The genera *Photobacterium* (2%) and *Aliivibrio* (1%) were also detected in the Bømlo samples. In the Chile sample, *Aliivibrio* was found to comprise 16% of the sequences.

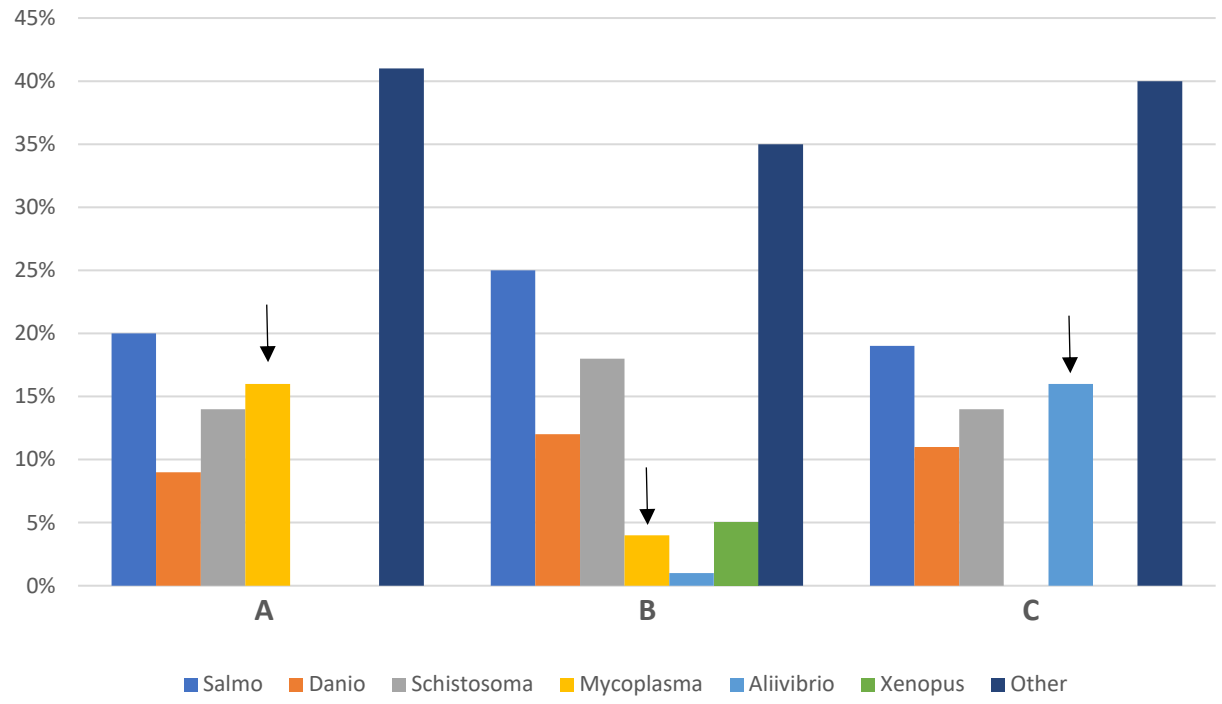

Figure 3.4: Average distribution of the most abundant DNA recovered from salmon gut. *Gut content collected from Skjervøy (A), Bømlo (B), and Chile (C) were processed for shotgun sequencing, and sequences were assigned taxonomy by MG-RAST. Genera below 5% are not shown but summarized in the "Other" column. The arrows mark what bacterial genus was most abundantly detected in the three respective sites. The graphic was created in Excel.*

Sequences belonging to the most abundant bacterium recovered from each site were further investigated at a lower taxonomic level. Figure 3.5 thus visualizes species classification of *Mycoplasma* sequences from Skjervøy and Bømlo, and *Aliivibrio* sequences from Chile. *Mycoplasma* sequences were annotated a total of 20 different species, all of which are listed in Appendix E. However, a few species were observed more frequently, including *M. penetrans* and *M. gallisepticum*. In Skjervøy samples, 51% of the sequences were classified as *M. penetrans*, and 14% as *M. gallisepticum*. A similar distribution was observed in Bømlo samples, where 52% of the sequences were annotated *M. penetrans* and 13% *M. gallisepticum*. *Aliivibrio* sequences obtained from Chile were classified as *Aliivibrio salmonicida* or *Vibrio fischeri*, where the latter dominated, comprising 78% of the sequences.

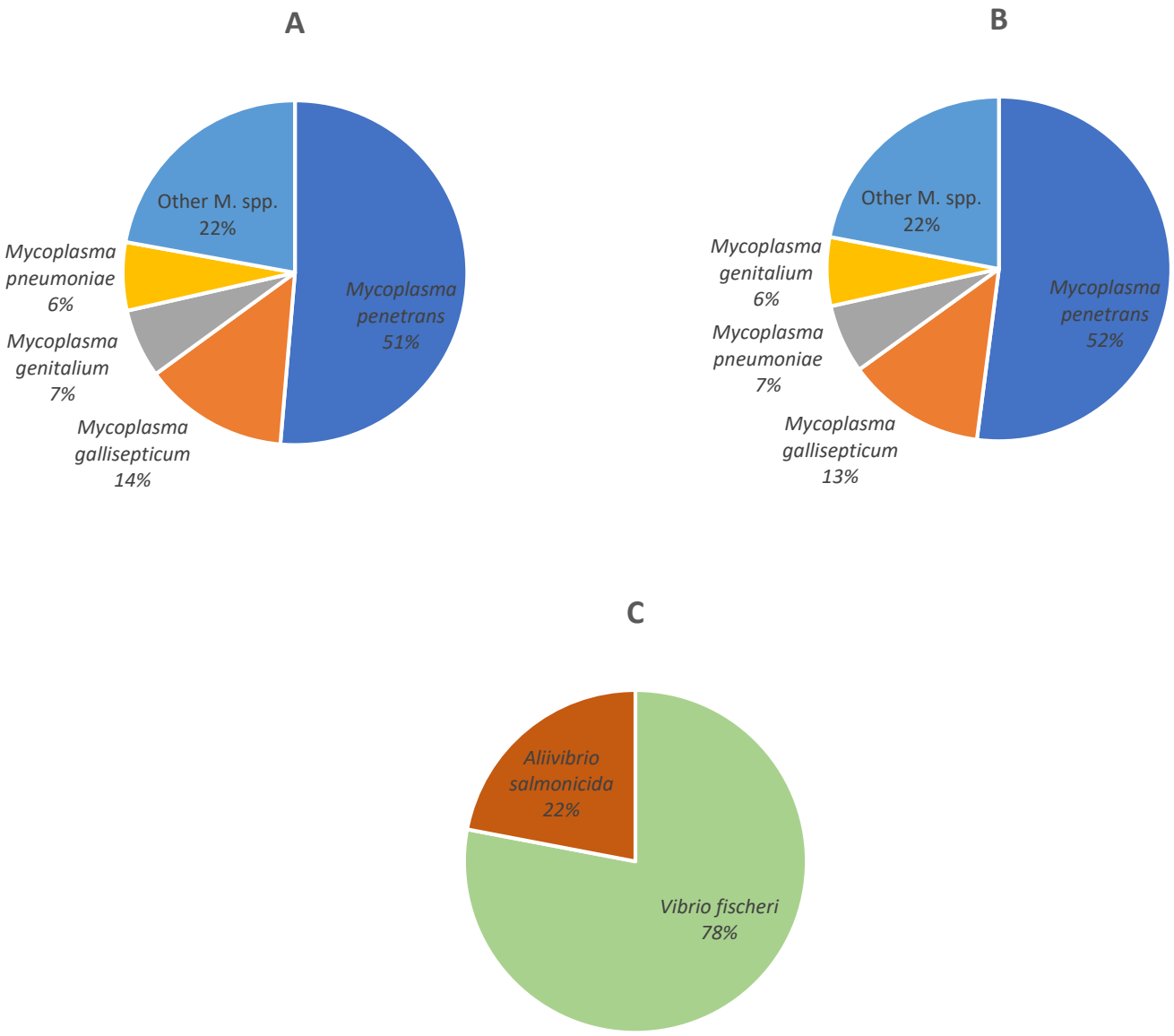

Figure 3.5: Taxonomic annotation of Mycoplasma and Aliivibrio sequences. *The pie charts each show MG-RAST's species classification of the most abundant bacterial genus recovered from three locations: Skjervøy (A), Bømlo (B), and Chile (C). Species comprising less than 5% of the sequences are not shown but summarized in the "Other M. spp." wedge. The pie charts were created in Excel.*

3.2.2 Recovery of Individual Genomes by MaxBin

One sample from Skjervøy containing high amounts of *Mycoplasma* was processed in SPAdes to investigate the salmon-associated *Mycoplasma* genome. SPAdes' algorithm assembled the shotgun sequences to a total of 7554 contigs, with length varying from 300 to 81 154 bp.

Contigs with less than 500 bp were discarded as noise, and 3964 contigs were further processed in MaxBin. Approximately 25% of the contigs were unclassified and, therefore, not binned. MaxBin grouped the remaining contigs into two bins. In Table 3.1, key information about the bins is summarized.

Table 3.1: Key information about proposed genomes generated by MaxBin. *Shotgun sequences obtained from farmed salmon were assembled to contigs, which thereafter were grouped into bins encompassing sequences belonging to the same organism. The table includes information derived from the binning process such as number of contigs, completeness of the genome, genome size, and G+C content.*

Bin	Total contigs	Completeness	Genome size	G+C content
bin1	1016	83.2%	2 562 264	27.9%
bin2	1953	81.3%	2 741 914	38.9%

As seen from Table 3.1, two large genomes were recovered from the sample. Both consisted of over 2500 kb and were estimated to be just over 80% complete. G+C content was relatively low for both genomes: ~28% and ~39% in bin1 and bin2, respectively.

The binned contigs were further uploaded to MG-RAST to identify the genomes. Bin1 contained 1016 contigs with an average length of 2522 bp, while bin2 contained 1953 contigs with an average length of 1404 bp. Bin1 was classified as *Mycoplasma*, containing about 62% *Mycoplasma*-annotated contigs, while Bin2 was classified as *Photobacterium*, containing 81% contigs annotated this genus. Figure 3.6 visualizes the taxa distribution of contigs comprising the two bins.

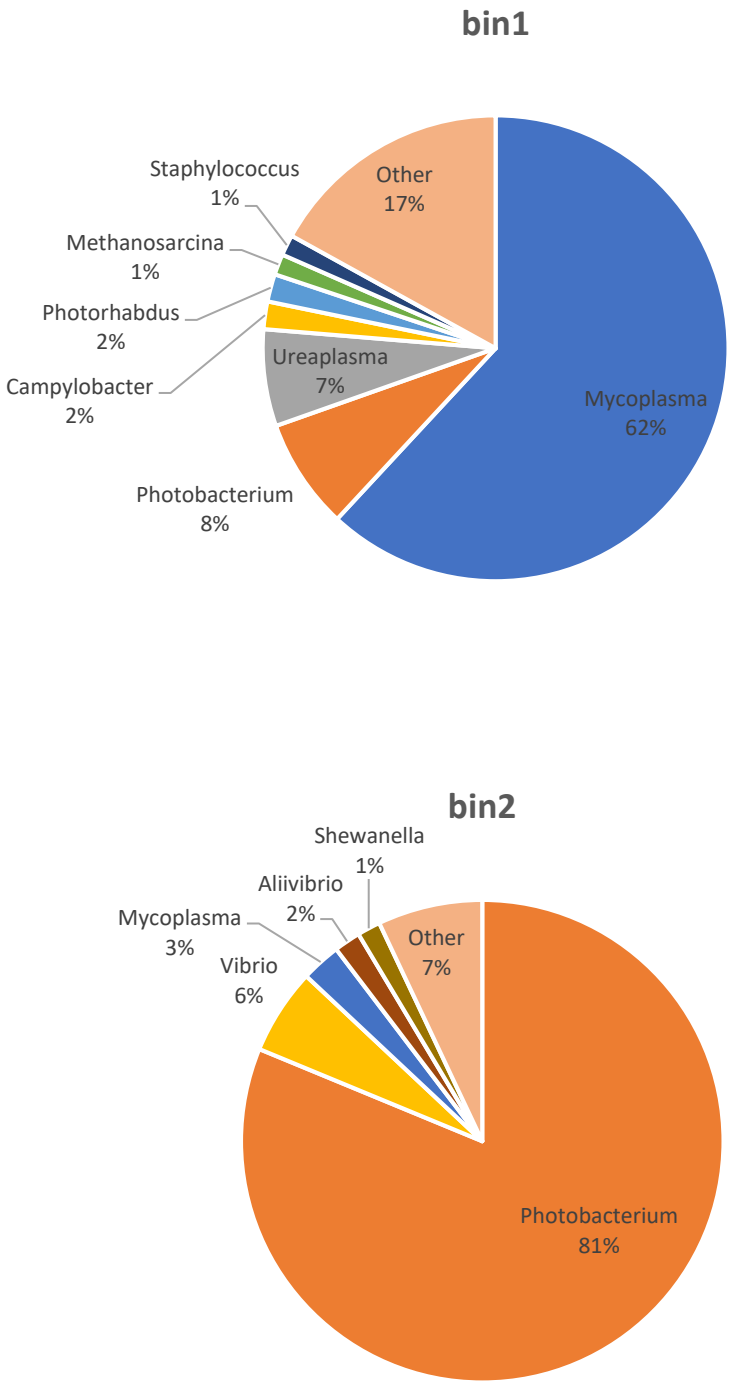

Figure 3.6: Taxonomic assignment of contigs belonging to bin1 and bin2 derived from MaxBin. *Contigs were assembled from shotgun data obtained from the gut of farmed salmon. The contigs were binned to recover individual genomes before classifying the bins using MG-RAST. Species comprising less than 1% of the contigs are not shown but summarized in the "other" wedge. The graphics were made in Excel using data derived from MG-RAST.*

3.3 Cultivation of Bacteria in Mycoplasma Growth Broth and Growth Agar

Cultivating *Mycoplasma* included setting up primary and secondary cultures, as well as agar plates. Microbial growth was observed in all primary cultures, while the secondary cultures were contaminated by fungi. Table 3.2 presents an overview of data generated from the cultivation process. The table includes observed growth in agar plates, the DNA concentration of cell pellets and the pH value of supernatants derived from the sedimentation of primary cultures, as well as quantification of extracted DNA and identification of said DNA.

Table 3.2: Overview of data gathered from attempting to cultivate Mycoplasma. *The table shows observed growth in agar plates, DNA concentration of cell pellets and pH value of supernatants derived from sedimentation of primary cultures, as well as DNA concentration of extracted DNA and the identity of said DNA.*

Primary culture (ID)	**Growth on agar plate**	**DNA conc. (ng/µL) cell pellet**	**pH value supernatant**	**DNA conc. (ng/µL) extracted DNA**	**Identification by BLAST**
F1	N/A	4.26	7.5	Too low	N/A
F2	N/A	4.18	7.0	Too low	N/A
F3	N/A	43.5	5.5	5.6	*Staphylococcus* sp.
F4	N/A	4.89	7.0	Too low	N/A
F5	N/A	4.60	7.0 - 7.5	Too low	N/A
F6	N/A	4.79	7.0 - 7.5	Too low	N/A
F7	N/A	6.80	7.0	0.6	*Cinara fresei* Uncultured bacterial/*Mycoplasma* sp.
F8	N/A	4.66	7.0	Too low	N/A
F9	N/A	29.0	7.0	27.9	*Micrococcus* sp.
F10	Yes	4.4	7.0	Too low	N/A

Note: Microbial growth was observed in all primary cultures

3.3.1 Growth in Primary Cultures

Approximately 2-3 weeks after filtering and transferring gut samples to fresh medium, microbial growth appeared in a few of the primary cultures. Growth was observed as a white ring-like structure at the bottom of the tubes. Over the next weeks, the ring established at varying intensity in all cultures. Figure 3.7 shows the cultures with the least and most growth.

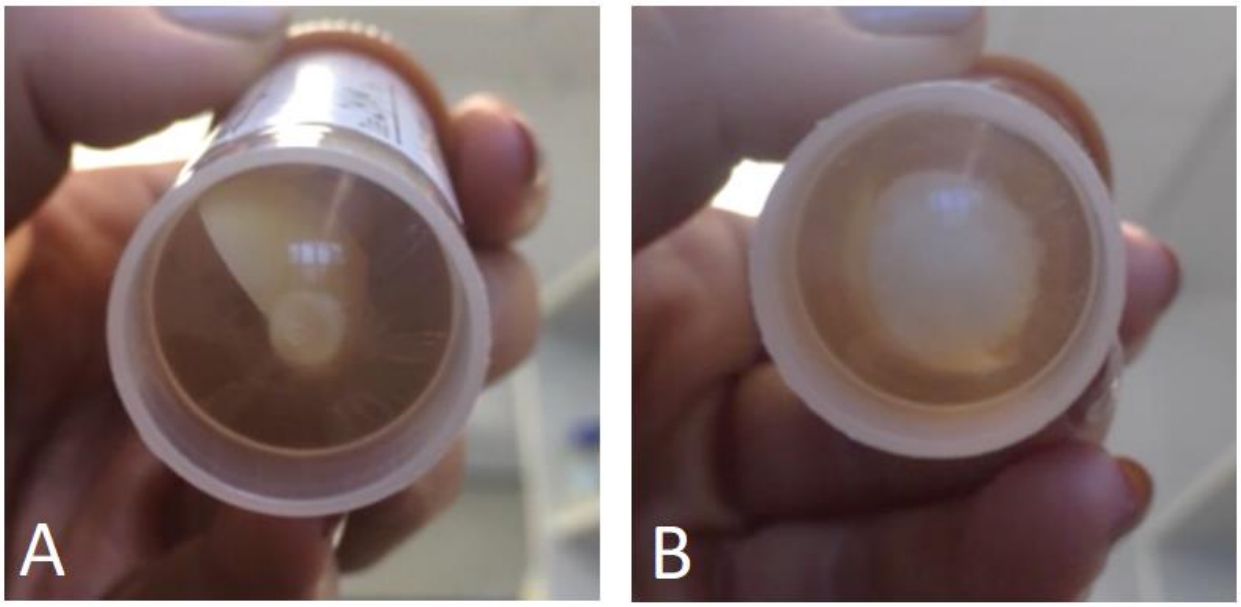

Figure 3.7: Microbial growth in Mycoplasma Growth Broth cultures. *Gut content obtained from salmon located in Bømlo, Norway, was filtered and transferred to growth medium. Picture A shows the sample tube with the least growth, while picture B shows the tube with most the growth after 1.5-2 months.*

3.3.2 Growth on Agar Plates

Colony-forming bacteria were observed on the agar plates after nearly 3 months. The bacteria appeared as white or brown circular colonies, slightly elevated from the agar (Figure 3.8A). Many of the colonies had a darker or denser middle. When picking colonies, they felt solid, almost stuck to the agar. Viewing the agar plate under a magnifier revealed clear fried egg-shaped colonies (Figure 3.8B).

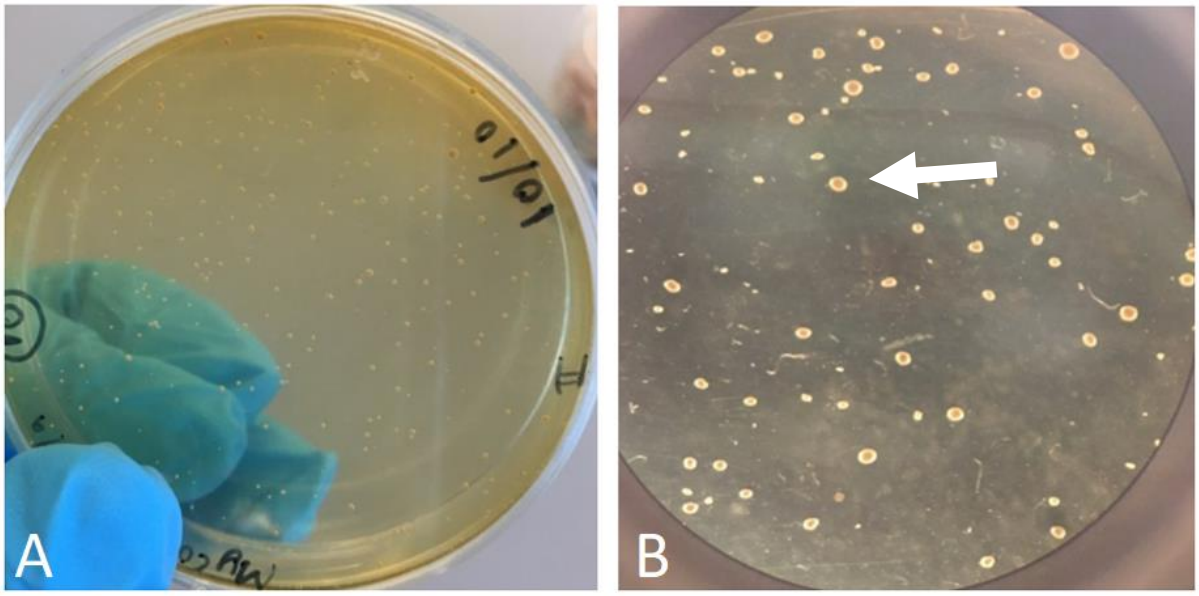

***Figure 3.8**: **Microbial growth in Mycoplasma Growth Agar.** Agar plates were plated with filtered gut content from salmon. Picture A shows bacterial growth on the agar plate after 3 months, while picture B shows the same agar plate under a magnifier. The arrow points to one of the many colonies that are shaped like a fried egg.*

3.3.3 DNA Quantification of Cell Pellets from Primary Cultures

Cell pellets derived from the primary cultures were quantified before freezing them for preservation. As seen in Table 3.2, DNA was detected in all cell pellets, and the concentration ranged from 4.18 - 43.5 ng/µL between the individual pellets.

3.3.4 pH Measurements of Supernatants from Primary Cultures

The supernatant's pH value was measured in primary cultures to identify possible fermentative or non-fermentative mycoplasmas. Table 3.2 lists the pH value in each culture. The pH varied from 7.0 - 7.5, with the exception of culture F3, in which the pH was measured to 5.5. As the MycoBroth originally had a pH value calibrated to 7.8 ± 0.2, the supernatant seemingly became slightly more acidic in most of the cultures.

3.3.5 DNA Extraction and Verification of Mycoplasma Broth Cultures

Identifying the microbes growing in the primary cultures was done by extracting and sequencing the obtained gDNA. As seen from Table 3.2, DNA extraction was only successful in three out of the ten cultures. Using BLAST, *Staphylococcus* and *Micrococcus* were identified in cultures F3 and F9, respectively. In culture F7, the forward and reverse sequences had varying results. The alignment of the forward sequence resulted in a species from the *Cinara* genus (insect). The result of the reverse sequence was inconclusive and suggested an uncultured bacterial species, but also an uncultured *Mycoplasma* species. Information regarding E-value, query coverage, and percent identity of the BLAST alignments can be viewed in Appendix F.

4. DISCUSSION

4.1 Geographical Difference in *Mycoplasma* Prevalence

The most conspicuous observation in the present study was the geographical difference of mycoplasmal prevalence revealed by 16S analyses. *Mycoplasma* was detected in Norwegian salmon but was completely absent in Chilean salmon (Figure 3.2). Corresponding to our findings, *Mycoplasma* has previously been detected in Norwegian salmon (5, 72). The genus has also been found in salmon from Scotland and Greenland (5, 21), indicating that it does not exclusively reside in Norwegian salmon. Additionally, several other fish species seem colonized by mycoplasmas (4), including salmonid fish such as Chinook salmon (17) and rainbow trout (73). However, no literature on mycoplasmal establishment in Chilean salmon was found.

Colonization of the fish gut is thought to occur immediately after hatching and is influenced by the first feed and surrounding water (18). Whether the bacteria encounter the fish as part of the feed or are free-living in the water, the gut microbiota is seemingly shaped by the bacteria present in the fish's environment. The fjords of Norway and Chile were found to have different characteristics, where water chemistry and structure of microbial communities varies (74). Even though the fjords investigated are not necessarily those our fish were collected from, it does state a difference between the two environments. Despite mycoplasmas' omnipresence, we cannot exclude that they are nonexistent in Chilean water environments, and consequently, Chilean salmon. However, differences in management practices, e.g., antibiotics usage, between the salmon farms in Norway and Chile, should also be considered as reasons for this absence.

Chilean salmon farms apply an enormous amount of antibiotics during the production, and a recent study claims that Chile currently has the highest use of antibiotics per ton of harvested fish in the world (75). Our contacts in Chile could not say if the sampled fish were exposed to any antibiotics. However, they did reveal that the fish might have been treated with lufenuron and/or florfenicol, as this was common practice for the fish farmers. The former is used as a treatment for salmon lice and targets chitin in insects, while the latter is an antibiotic targeting the protein synthesis of bacteria. Although the salmon-associated *Mycoplasma's* susceptibility to florfenicol is unknown, other *Mycoplasma* species have been found sensitive to this antibiotic (76).

4.2 Salmon-associated *Mycoplasma's* relatedness to other *M.* spp.

The *Mycoplasma*-annotated sequences were most frequently identified as *M. penetrans,* with more than 50% of the sequences classified as this species (Figure 3.5). Furthermore, *M. gallisepticum, M. pneumoniae,* and *M. genitalium* were identified in descending order. Interestingly, all four seem to originate from the same common ancestor and are clustered in the *Mycoplasma* phylogenetic three (29). Considering the taxonomic assignment in this study, it is conceivable that the salmon-associated *Mycoplasma* has some relatedness to these species. Moreover, as *M. penetrans* dominated, shared genes between this and salmon *Mycoplasma* is likely. Our findings correspond with previous studies, where salmon *Mycoplasma* was found to have similar gene content to that of *M. penetrans* (5, 72).

Despite the predominance of *M. penetrans*-annotated sequences, a total of 20 different species were identified by MG-RAST. A list containing all the species can be seen in Appendix E. Essentially, as the sequences were matched with many species rather than one specific, it may imply they were erroneously assigned to these species. To my knowledge, no genome sequence of the salmon *Mycoplasma* has yet been established, and no reference genome was thus available at the time the alignment analyses were performed. As the results in the present study suggest, salmon *Mycoplasma* may share multiple genes with *M. penetrans,* but also with other members of the genus.

4.3 Recovery of a Salmon-associated *Mycoplasma* Genome

Bins generated by MaxBin are considered to represent individual genomes recovered from a microbial community (70). One of the bins derived from the Skjervøy sample was classified as *Mycoplasma*. Thus, the proposed genome of salmon *Mycoplasma* was in this study found to contain over 2500 kb (Table 3.1). Additionally, the genome was only ~80% complete, indicating an even larger size. The largest *Mycoplasma* genome characterized – which interestingly belongs to *M. penetrans* – consists of just over 1350 kb (29). It seems improbable that the salmon *Mycoplasma* genome surpasses this to such an extent. Taxonomic assignments by MG-RAST revealed that other organisms contaminated large portions of the bin, and only 62% of the contigs were classified as *Mycoplasma* (Figure 3.6). Therefore, the constructed genome is probably excessively large because contigs not belonging to *Mycoplasma* were wrongfully placed in this bin.

The *Mycoplasma* genome is generally low in G+C content and usually ranges between 24-33% (27). Accordingly, the G+C content of the proposed genome in this study was calculated to 28%, (Table 3.1). However, seeing the genome was contaminated by other organisms, the actual G+C content might be slightly higher or lower.

4.4 The Potential Role of *Mycoplasma* in Salmon Gut

Whether the mycoplasmas have a protective or pathogenic role in the salmon gut remains an unanswered question. However, mycoplasmas are the causative agents of many diseases in other animals (27). The findings in the present study suggest that the salmon *Mycoplasma* has a gene content similar to that of *M. penetrans*, a human pathogen infecting the respiratory and urogenital tract (77). Therefore, the two species may exhibit some of the same virulence factors. For instance, the *M. penetrans* genome includes genes encoding cytotoxic proteins, which are thought to damage the host tissue (77). In addition, some strains of this species were found to penetrate the epithelial cells of its host, and while some cells remained intact, extensive invasion resulted in cell disruption and necrosis (40). However, there are, to my knowledge, currently no reports on mortality in the salmon industry attributable to gut mycoplasmas. If salmon *Mycoplasma* was associated with severe infection, the disease would likely be considered a common health issue in salmon farms given the seemingly large prevalence of the bacterium.

According to a recent study, mycoplasmas presumably colonize the gut in a non-neutral fashion, suggesting that the host might be exerting an influence over this bacterium (6). Whether this is due to mycoplasmas having beneficial characteristics or that they are highly adapted to the salmon gut is unknown. However, as *Mycoplasma* is consistently found in the gut microbiota of fish, the bacterium might have a protective role in the gut. Most mycoplasmas adhere to the epithelial cells of its host (37). Therefore, it could be possible that the adhering mycoplasmas outcompete other pathogenic bacteria as the tissue is already colonized. The 16S analyses in the present study revealed that *Aliivibrio* was more abundant when *Mycoplasma* was absent. Thus, our findings suggest that the abundance of *Aliivibrio* is negatively correlated with the presence of *Mycoplasma.* As seen in Figure 3.5, one of our observations was that *Aliivibrio salmonicida* inhabited the gut of Chilean salmon. *A. salmonicida* is known as the causative agent for cold-water vibriosis, a disease causing the fish to experience general bleeding and eventually death (78). In support, a study on Chinook salmon found that the abundance of potential pathogenic *Vibrio* species appeared to be negatively correlated with the presence of *Mycoplasma* (17).

4.5 *Mycoplasma* Cultivation and DNA Extraction

Many mycoplasmas grow slowly and sparingly in growth mediums, and novel species are likely to require more cultivation time due to a less well-adapted medium (27). Although microbial growth – slow and sparing – was observed in the primary cultures set up in this project, there were not necessarily mycoplasmas growing in the cultures. Still, the first indication of *Mycoplasma* growth in primary cultures is often a slight to moderate pH change of the medium (27). As seen in Table 3.2, the pH value decreased in all primary cultures. Additionally, when plating out one of the cultures on agar, colonies appeared fried egg-shaped (Figure 3.8B), which is a characteristic colony shape of mycoplasmas (27).

To identify the microbe(s) present in the cultures, we attempted to extract DNA for sequencing. However, isolating DNA from the cultures proved to be challenging. In the cases DNA was recovered, the results revealed other organisms growing in the cultures. Several protocols for DNA extraction from colonies on the agar plates were performed as well, but nor here was any DNA recovered. Mycoplasmas have been suggested to penetrate the agar (79), making DNA extraction from agar plates non-optimal as DNA will be lost when picking colonies. This may be a valid reason to why extraction from solid media did not yield any DNA, though it does not explain why extraction from cultures was unsuccessful.

Mycoplasmas require nucleases for their metabolism, which work by digesting DNA and RNA of the host cells (80). However, when mycoplasmas are filtered and cultivated in growth medium, ideally, no DNA from a host cell is available. Therefore, it could be possible the mycoplasmas' DNA is degraded by their own nucleases when the cell is lysed during extraction. In this case, we would be unable to recover the DNA. If DNA from other organisms is available to the nucleases, the Mycoplasma's DNA might not be digested. Supporting this theory is our observation that mycoplasmal DNA was able to be recovered when we applied 16S rRNA sequencing using mixed cultures. With that said, we still do not know if *Mycoplasma*, in fact, was growing in the cultures. However, we do know that DNA quantification of cell pellets indicated DNA presence in all cultures, even if we were unable to extract DNA from them (Table 3.2).

4.6 Other Bacteria in the Salmon Gut

The phylum Proteobacteria was in this project found to dominate the gut microbiota of farmed salmon. As shown in Figure 3.1, the gut microbiota of Chilean salmon was entirely made up of Proteobacteria (97%), while in Norwegian salmon, the phylum was slightly less dominant (~70%). Our findings correspond well with previous studies on salmon gut microbiota, where Proteobacteria is the phylum most consistently detected. Some of the publications considered in this thesis, listed in Table 1, all report observing Proteobacteria.

The major reason for Proteobacteria's abundance is possibly due to the phylum's vast size. It encompasses a wide variety of genera, of which many are facultative or obligately anaerobic (81). The preference for an oxygen-free environment makes the bacteria included in this phylum plausible inhabitants of the intestinal microbiota. The most abundant genera detected in the present study were *Aliivibrio* and *Photobacterium* (Figure 3.2), which both encompass potential fish pathogenic species (82, 83). However, *Photobacterium* has been detected in several projects investigating salmon gut microbiota (20, 23, 24), suggesting that this genus is a normal inhabitant of the salmon intestine. Observations of *Aliivibrio* have also occurred (21), and one study proposes this genus as one of the core genera' in the salmon gut (23). Therefore, it is likely they are part of the general gut microbiota of salmon, living there either as commensals or symbionts.

4.7 Technical Considerations

4.7.1 Small Sample-size

One of the main weaknesses of the present study was the small number of samples investigated. In addition to 23 previously-obtained samples from Skjervøy, we collected 19 samples from Bømlo and 20 samples from Chile. Therefore, the results cannot be applied on a larger scale to affirm that *Mycoplasma* is only found in Norwegian salmon and not in Chilean salmon. With that said, the samples used in this project were sampled at different times, on two occasions in Norway, and two in Chile. This shows the results are somewhat reproducible and strengthens the findings.

4.7.2 Lack of Information Regarding Sampled Fish

The farmed fish sampled for this thesis were adults and large enough for harvest. Besides this, little information regarding the fish and its health status was received. We were not thoroughly informed about management practices at the different salmon farms, such as medical or mechanical treatments, antibiotics used, or nutritional components in the feed. It is important to mention that these factors may influence the gut microbiota and are likely to differ between the locations we collected our fish.

4.7.3 No Optimized Protocol for Mycoplasma Cultivation and DNA Extraction

As gut *Mycoplasma* to this date has not been successfully isolated from salmon, no protocol for cultivating this bacterium is currently available. Additionally, no protocol used for DNA extraction of possible *Mycoplasma* growing in filtered MycoBroth cultures or agar plates were found to be optimal.

4.8 Further Work

Mycoplasmas rely on the host for several biosynthetic pathways and may not efficiently metabolize or utilize certain products without a host cell available. Supplying the growth medium with host cells could, therefore, aid in the growth and maintenance of mycoplasmas in a culture. Thus, creating a refined cultivation protocol that includes epithelial cells from salmon should be considered for further research of the salmon *Mycoplasma*.

Mycoplasmas have atypical characteristics compared to other bacteria. Extraction kits made for bacteria may thus not necessarily be suitable for mycoplasmal DNA extraction. Therefore, it could be an idea to attempt extracting DNA using a fungal extraction kit. However, one of the cultures we were able to extract DNA from, indicated presence of an unidentified bacterial species (Table 3.2). Although with a higher E-value, BLAST also suggested an unidentified *Mycoplasma* sp., and the sample was processed for shotgun sequencing. The shotgun library was not analyzed due to a lack of time. Finish analyzing the library and identifying the bacterium present in the sample would determine if the cultivation and extraction, to some extent, were successful. Should the bacterium turn out to be *Mycoplasma*, it would, to my knowledge, be the first known isolation of *Mycoplasma* from salmon gut.

The genome contains information about mycoplasmas' coding genes and their function. Therefore, further investigations of the genetic material obtained in this study should be considered. Removing contigs belonging to contaminating microorganisms from the *Mycoplasma*-bin could reveal the true size of the genome and allow us to examine the gene functions. Gene annotation could further help us elucidate possible negative or positive impacts *Mycoplasma* has on the salmon's immune system and physiology.

5. CONCLUSION

In the present study, we found a geographical difference in mycoplasmas' prevalence in the gut microbiota of farmed Atlantic salmon. *Mycoplasma* was detected in Norwegian salmon but was absent from Chilean salmon. Differences in the environment and salmon farms in Norway and Chile possibly explain the lacking colonization of *Mycoplasma* in Chilean salmon. However, the underlying reasons for this absence must be further investigated to clarify our findings.

Mycoplasma was attempted cultivated, but unsuitable cultivation and DNA extraction protocols made it challenging to obtain genetic material from a pure culture. The genome of the salmon-associated *Mycoplasma* was, therefore, investigated using mixed cultures. Most of the obtained mycoplasmal DNA was classified as *M. penetrans*, indicating that the salmon *Mycoplasma* is closely related to this species. Furthermore, a genome consisting of over 2500 kb was suggested. This is larger than any previously sequenced *Mycoplasma* genome, and it was discovered that DNA from microorganisms not belonging to *Mycoplasma* was included in the genome.

Mycoplasmas are usually dependent on a host and are often associated with disease in animals. However, given the seemingly large prevalence of *Mycoplasma* in salmon, its existence as a commensal inhabitant in the intestinal tract is likely. Further research is required to elucidate the *Mycoplasma's* influence on salmon physiology and immunology, and its role as part of the salmon gut microbiota. The harboring microbes' impact on fish health becomes progressively acknowledged, and future utilization of the gut microbiota in aquaculture could hopefully increase overall health and welfare conditions amongst farmed salmon and lead to a more sustainable industry.

REFERENCES

1. Straume HM, Landazuri-Tveteraas U, Oglend A. Insights from transaction data: Norwegian aquaculture exports. Aquac. Econ. Manag. [Internet]. 2019 Oct 31 [cited 2020 May]. doi: 10.1080/13657305.2019.1683914
2. Richardsen R, Myhre MS, Tyholt IL, Johansen U. Nasjonal betydning av sjømatnæringen. [internet] Trondheim, Norway: SINTEF; 2019 [cited 2020 Apr]. Available from: https://www.sintef.no/globalassets/sintef-ocean/pdf/nasjonal-verdiskapning_2018tall_endelig_200619.pdf
3. Sommerset I, Walde CS, Jensen BB, Bornø G, Haukaas A, Brun E. Fiskehelserapporten 2019. Rapport 5a/2020. [internet] Norway: Veterinærinstituttet; 2020 [cited 2020 Apr]. Available from: https://www.vetinst.no/rapporter-og-publikasjoner/rapporter/2020/fiskehelserapporten-2019
4. Yukgehnaish K, Kumar P, Sivchandran P, Marimuthu K, Arshad A, Paray BA, et al. Gut microbiota metagenomics in aquaculture: factors influencing gut microbiome and its physiological role in fish. Aquacult. 2020 Jan 29 [cited 2020 Apr];1-25. Available from: https://doi.org/10.1111/raq.12416
5. Holben WE, Williams P, Saarinen M, Särkilahti LK, Apajalahti JHA. Phylogenetic Analysis of Intestinal Microflora Indicates a Novel Mycoplasma Phylotype in Farmed and Wild Salmon. Microb. Ecol. 2002 Sep;44:175-85. doi: 10.1007/s00248-002-1011-6
6. Heys C, Cheaib B, Busetti A, Kazlauskaite R, Maier L, SLoan W, et al. Neutral processes dominate microbial community assembly in Atlantic salmon, *Salmo salar*. Appl. Environ. Microbiol. [Internet]. 2020 Apr 1 [cited 2020 May];86(8):e02283-19. Available from: https://aem.asm.org/content/86/8/e02283-19. doi: 10.1128/AEM.02283-19
7. Vøllestad A. laksefamilien [Internet]. Oslo: Store norske leksikon; 2009 [updated 2019; cited 2020 May]. Available from: https://snl.no/laksefamilien
8. Vøllestad A. laks [Internet]. Oslo: Store norske leksikon; 2009 [updated 2019; cited 2020 May]. Available from: https://snl.no/laks
9. Thorstad EB, Whoriskey F, Rikardsen AH, Aarestrup K. Aquatic Nomads: The Life and Migrations of the Atlantic Salmon. In: Aas Ø, Einum S, Klemetsen A, Skurdal J, editors. Atlantic Salmon Ecology. United Kingdom: Wiley-Blackwell; 2010. p. 1-32
10. Klemetsen A, Amundsen P-A, Dempson J, Jonsson B, Jonsson N, O'Connell MF, et al. Atlantic salmon *Salmo salar* L., brown trout *Salmo trutta* L. and Arctic charr *Salvelinus alpinus* (L.): a review of aspects of their life histories. Ecol Freshw Fish. 2003 Feb 13;12(1):1-59. doi: 10.1034/j.1600-0633.2003.00010.x
11. Rikardsen AH, Dempson JB. Dietary Life-support: The Food and Feeding of Atlantic Salmon at Sea. In: Aas Ø, Einum S, Klemetsen A, Skurdal J, editors. Atlantic Salmon Ecology. United Kingdom: Wiley-Blackwell; 2010. p. 115-43

12. Van Pelt TI, Piatt JF, Lance BK, Roby DD. Proximate composition and energy density of some North Pacific forage fishes. Comp. Biochem. and Physiol. 1997 Dec;118(4):1393-8 doi: 10.1016/S0300-9629(97)00240-5
13. Misund B. fiskeoppdrett [Internet]. Oslo: Store norske leksikon; 2009 [updated 2019; cited 2020 May]. Available from: https://snl.no/fiskeoppdrett
14. Green TJ, Smullen R, Barnes AC. Dietary soybean protein concentrate-induced intestinal disorder in marine farmed Atlantic Salmon, *Salmo salar* is associated with alterations in gut microbiota. Vet. Microbiol. 2013 Sep;166(1-2):286-92. doi: 10.1016/j.vetmic.2013.05.009
15. Naylor RL, Goldburg RJ, Primavera JH, Kautsky N, Beveridge M, Clay J, et al. Effect of aquaculture on world fish supplies. Nature [Internet]. 2000 Jun [cited 2020 May];405:1017-24. Available from: https://www.nature.com/articles/35016500
16. Desai AR, Links MG, Collins SA, Mansfield GS, Drew MD, Van Kessel AG, et al. Effects of plant-based diets on the distal gut microbiome of rainbow trout (*Oncorhynchus mykiss*). Aquaculture. 2012 Jun;350-353:134-42. doi: 10.1016/j.aquaculture.2012.04.005
17. Ciric M, Waite D, Draper J, Jones JB. Characterisation of gut microbiota of farmed Chinook salmon using metabarcoding. bioRxiv [Internet] 2018 Mar [cited 2020 May]. Available from: https://doi.org/10.1101/288761
18. Egerton S, Culloty S, Whooley J, Stanton C, Ross RP. The Gut Microbiota of Marine Fish. Front. Microbiol. 2018 May;9:873. doi: 10.3389/fmicb.2018.00873
19. Merrifield DL, Rodiles A. The fish microbiome and its interaction with mucosal tissues. In: Beck BH, Peatman E, editors. Mucosal Health in Aquaculture. USA: Academic Press; 2015. p. 273-95
20. Gajardo K, Rodiles A, Kortner TM, Krogdahl Å, Bakke AM, Merrifield DL, et al. A high-resolution map of the gut microbiota in Atlantic salmon (Salmo salar): A basis for comparative gut microbial research. Scientific Reports. 2016 Aug;6:30893. doi: 10.1038/srep30893
21. Llewellyn MS, McGinnity P, Dionne M, Letuourneau J, Thonier F, Carvalho GR, et al. The biogeography of the Atlantic salmon (Salmo Salar) Gut Microbiome. ISME J. 2016 May;10(5):1280-4. doi: 10.1038/ismej.2015.189
22. Rudi K, Angell IL, Pope PB, Vik JO, Sandve SR, Snipen L-G. Stable Core Gut Microbiota across the Freshwater-to-Saltwater Transition for Farmed Atlantic Salmon. Appl. Environ. Microbiol. 2018 Jan;84(2):e01974-17. doi: 10.1128/AEM.01974-17
23. Dehler CE, Secombes CJ, Martin SAM. Environmental and physiological factors shape the gut microbiota of Atlantic salmon parr (*Salmo salar* L.). Aquaculture. 2016 Jan;467:149-57. doi: 10.1016/j.aquaculture.2016.07.017
24. Hovda MB, Lunestad BT, Fontanillas R, Rosnes JT. Molecular characterisation of the intestinal microbiota of farmed Atlantic salmon (*Salmo salar* L.). Aquaculture. 2007;272:581-8. doi: 10.1016/j.aquaculture.2007.08.045

25. Navarrete P, Espejo RT, Romero J. Molecular Analysis of Microbiota Along the Digestive Tract of Juvenile Atlantic Salmon (*Salmo salar* L.). Microb. Ecol. 2009 Apr;57(3):550-61. doi: 10.1007/s00248-008-9448-x
26. Waites KB, Xiao L, Paralanov V, Viscardi RM, Glass JI. Mycoplasma and Ureaplasma. In: de Filippis I, McKee ML, editors. Molecular Typing in Bacterial Infections: Totowana: Humana Press; 2013. p. 229-81
27. Razin S. The Genus Mycoplasma and Related Genera (Class Mollicutes). In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, editors. The Prokaryotes. 3rd ed. New York: Springer; 2006. p. 836-904
28. Woese CR, Maniloff J, Zablen LB. Phylogenetic analysis of the mycoplasmas. Proc. Natl. Acad. Sci. 1980 Jan;77(1):494-8: 10.1128/jb.171.12.6455-6467.1989
29. Thompson CC, Vieira NM, Vincente ACP, Thompson FL. Towards a genome based taxonmy of Mycoplasmas. Infect. Genet. Evol. 2011 Oct;11(7):1798-804. doi: 10.1016/j.meegid.2011.07.020
30. Himmelreich R, Hilbert H, Plagens H, Pirkl E, Li BC, Herrmann R. Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae. Nucleic Acids Res. 1996 Nov;24(22):4420-49. doi: 10.1093/nar/24.22.4420
31. Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, et al. The Minimal Gene Complement of *Mycoplasma genitalium*. Science. 1995 Nov;270(5235):397-404. doi: 10.1126/science.270.5235.397
32. Manolukas JT, Barlie MF, Chnadler DK, Pollack JD. Presence of anaplerotic reactions and transamination, and the absence of the tricarboxylic acid cycle in mollicutes. J. Gen. Microbiol. 1988 Mar;134(3):791-800.
33. Razin S. The mycoplasmas. Microbiol. Rev. 1978 Jun;42(2):414-70
34. Taylor RR, Varsani H, Miles RJ. Alternatives to arginine as energy sources for the non-fermentative *Mycoplasma gallinarum*. FEMS Microbiol. Lett. 1994 Jan;115(2-3):163-7. doi: 10.1111/j.1574-6968.1994.tb06632.x.
35. Pollack JD, Williams MV, McElhaney RN. The Comparative Metabolism of the Mollicutes (*Mycoplasmas):* The Utility for Taxonomic Classification and the Relationship of Putative Gene Annotation and Phylogeny to Enzymatic Function in the Smallest Free-Living Cells. Crit. Rev. Microbiol. 1997 Feb;23(4):269-354. doi: 10.3109/10408419709115140
36. Tully JG. Mollicute-host interrelationships: Current concepts and diagnostic implications. In: Tully JG, Razin S, editors. Molecular and Diagnostic Procedures in Mycoplasmology. 2nd ed. U.S: Academic Press; 1996. p. 1-21
37. Razin S, Jacobs E. Mycoplasma adhesion. J. Gen. Microbiol. 1992 Mar;138(3):407-22. doi: 10.1099/00221287-138-3-407
38. Rottem S. Interaction of mycoplasmas with host cells. Physiol. Rev. 2003 Apr;83(2):417-32. doi: 10.1152/physrev.00030.2002
39. Chaudhry R, Varshney AK, Malhotra P. Adhesion proteins of Mycoplasma pneumoniae. Front. Biosci. 2007 Feb;12(2):690-9. doi: 10.2741/2093

40. Lo S-C, Hayes MM, Tully JG, Wang RY, Kotani H, Pierce PF, et al. *Mycoplasma penetrans* sp. nov., from the Urogenital Tract of Patients with AIDS. Int. J. Syst. Bacteriol. 1992 Jul;42(3):357-64.
41. Malla MA, Dubey A, Kumar A, Yadav S, Hashem A, Abdallah EF. Exploring the Human Microbiome: The Potential Future Role of Next-Generation Sequencing in Disease Diagnosis and Treatment. Front. Immunol. 2019 Jan; 9:2868. doi: 10.3389/fimmu.2018.02868
42. Thomas T, Gilbert J, Meyer F. Metagenomics - a guide from sampling to data analysis. Microb. Inform. Exp. 2012 Feb; 2(3). doi: 10.1186/2042-5783-2-3
43. Jovel J, Patterson J, Wang W, Hotte N, O'Keefe S, Mitchel T, et al. Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. Front. Microbiol. 2016 Apr;7(459). doi: 10.3389/fmicb.2016.00459
44. Patel JB. 16S rRNA Gene Sequencing for Bacterial Pathogen Identification in the Clinical Laboratory. Mol. Diagn. 2001 Dec;6(4):313-21. doi: 10.1054/modi.2001.29158
45. Woese CR. Bacterial Evolution. Microbiol. Rev. 1987 Jun;51(2):221-71
46. Clarridge JE. Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases. Clin. Microbiol. Rev. 2004 Oct;17(4):840-62. doi: 10.1128/CMR.17.4.840-862.2004
47. Hadziavdic K, Lekang K, Lanzen A, Jonassen I, Thompson EM, Troedsson C. Characterization of the 18S rRNA Gene for Designing Universal Eukaryote Specific Primers. PLoS One. 2014 Feb;9(2). doi: 10.1371/journal.pone.0087624
48. Vibin J, Chamings A, Collier F, Klaassen M, Nelson TM, Alexandersen S. Metagenomics detection and characterisation of viruses in faecal samples from Australian wild birds. Sci. Rep. [Internet] 2018 Jun [cited 2020 May];8(8686). Available from: https://www.nature.com/articles/s41598-018-26851-1
49. Shelton S. 16S sequencing vs. Shotgun metagenomics: Which one to use when it comes to microbiome studies [Internet]. [place unknown]: Genohub; 2018 [cited 2020 May]. Available from: https://blog.genohub.com/2018/04/12/16s-sequencing-vs-shotgun-metagenomics-which-one-to-use-when-it-comes-to-microbiome-studie/
50. Green ED. Shotgun Sequencing [Internet]. U.S: National Human Genome Research Institute; 2020 [cited 2020 May] Available from: https://www.genome.gov/genetics-glossary/Shotgun-Sequencing
51. Clark DP, Pazdernik NJ. Molecular Biology [Internet]. U.S: Academic Cell; 2013. Chapter 9, Genomics and Systems Biology. [cited 2020 May]. Available from: https://doi.org/10.1016/B978-0-12-378594-7.00009-3.
52. Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA. Genomics. 2016 Jan;107(1):1-8. doi: 10.1016/j.ygeno.2015.11.003
53. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. 1977 Dec;74(12):5463-7. doi: 10.1073/pnas.74.12.5463

54. Chdgeavadze ZG, Beabealashvilli RS. 2', 3'-Dideoxy-3' aminonucleoside 5'-triphosphate are the terminators of DNA synthesis catalyzed by DNA polymerases. 1984 Feb;12(3):1671-86. doi: 10.1093/nar/12.3.1671
55. Prober JM, Trainor GL, Dam RJ, Hobbs FW, Robertson CW, Zagursky RJ, et al. A System for Rapid DNA Sequencing With Fluorescent Chain-Terminating Dideoxynucleotides. Science. 1987 Oct;238(4825):336-41. doi: 10.1126/science.2443975
56. Margulies M, Egholm M, Altman W, Attiya S. Genome sequencing in microfabricated high-density picolitre reactors [Internet]. Nature. 2005 Jul [cited 2020 May];437:376-80. Available from: https://www.nature.com/articles/nature03959
57. Greenwood M. What is Pyrosequencing? [Internet] [place unknown]: News Medical Life Sciences; 2018 [cited 2020 May]. Available from: https://www.news-medical.net/life-sciences/What-is-Pyrosequencing.aspx
58. Nyrén P, Lundin A. Enzymatic method for continuous monitoring of inorganic pyrophosphate synthesis. Anal. Biochem. 1985 Dec;151(2):504-9. doi: 10.1016/0003-2697(85)90211-8
59. Voelkerding KV, Dames Sa, Durtschi JD. Next-generation sequencing: from basic research to diagnostics. Clin. Chem. 2009 Apr;55(4):641-58. doi: 10.1373/clinchem.2008.112789
60. Turcatti G, Romieu A, Fedurco M, Tairi A-P. A New Class of Cleavable Fluorescent Nucleotides: Synthesis and Optimization as Reversible Terminators for DNA Sequencing by Synthesis. Nucleic Acids Res. 2008 Mar; 36(4):e25. doi: 10.1093/nar/gkn021
61. Schadt EE, Turner S, Kasarskis A. A window into third-generation sequencing. Hum. Mol. Genet. 2010 Oct;19(R2):R227-R40. doi: 10.1093/hmg/ddq416
62. Niedringhaus TP, Milanova D, Kerby MB, Snyder MP, Barron AE. Landscape of Next-Generation Sequencing Technologies. Anal. Chem. 2011 Jun;83(12):4327-41. 10.1021/ac2010857
63. Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I, et al. Single-Molecule DNA Sequencing of a Viral Genome. Science. 2008 Apr;320(5872):106-9. doi: 10.1126/science.1150427
64. Rhoads A, Au KF. PacBio Sequencing and Its Applications. Genomics Proteomics Bioinformatics. 2015 Oct;13(5):278-89. doi: 10.1016/j.gpb.2015.08.002
65. Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, et al. The potential and challenges of nanopore sequencing. Nat. Biotechnol. [Internet]. 2008 Oct [cited 2020 May];26:1146-53. Available from: https://www.nature.com/articles/nbt.1495
66. Kerfeld CA, Scott KM. Using BLAST to Teach "E-value-tionary" Concepts. PLOS Biol. 2011 Feb; 9(2)e1001014. doi: 10.1371/journal.pbio.1001014
67. Caparaso GJ, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods. 2010 May;7(5):335-6. doi: 10.1038/nmeth.f.303

68. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 2017 May;27(5):824-34. doi: 10.1101/gr.213959.116
69. Wilke A, Bischof J, Gerlach W, Glass E, Harrison T, Keegan KP, et al. The MG-RAST metagenomics database and portal in 2015. Nucleic Acids Res. 2015 Jan;44(D1):D590-D4. doi: 10.1093/nar/gkv1322
70. Wu Y-W, Tang Y-H, Tringe SG, Simmons BA, Singer SW. MaxBin: an automated binning method to revover individual genomes from metagenomes using an expectation-maximization algorithm. Microbiome [Internet]. 2014 Aug [cited 2020 May];2(26). Avaliable from: https://microbiomejournal.biomedcentral.com/articles/10.1186/2049-2618-2-26
71. Wilke A, Harrison T, Wilkening J, Field D, Glass E, Kyrpides N, et al. The M5nr: a novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. BMC Bioinformatics. 2012 Jun;13(141). doi: 10.1186/1471-2105-13-141
72. Jin Y, Angell IL, Sandve SR, Snipen L-G, Olsen Y, Rudi K. Atlantic salmon raised with diets low in long-chain polyunsaturated n-3 fatty acids in freshwater have a Mycoplasma-dominated gut microbiota at sea. Aquacult. Env. Interac. 2019 Jan;11:31-9. doi: 10.3354/aei00297
73. Brown RM, Wiens GD, Salinas I. Analysis of the gut and gill microbiome of resistant and susceptible lines of rainbow trout (*Oncorhynchus mykiss*). Fish Shellfish Immunol. 2019 Mar;86:497-506. doi: 10.1016/j.fsi.2018.11.079
74. Bailey J, Ardelan MV, Hernández KL, González HE, Iriarte JL, Olsen LM, et al. Interdisciplinarity as an Emergent Property: The Research Project "CINTERA" and the study of Marine Eutrophication. Sustainability. 2015 Jul;7(7):9118-39. doi: 10.3390/su7079118
75. Miranda CD, Godoy FA, Lee MR. Current Status of the Use of Antibiotics and the Antimicrobial Resistance in the Chilean Salmon Farms. Front. Microbiol. 2018 Jun;9:1284. doi: 10.3389/fmicb.2018.01284
76. Siugzdaite J, Gabinaitiene A, Kerziene S. Susceptibility of *Mycoplasma bovis* field isolates to antimicrobial agents. Veterinarni Medicina [Internet]. 2012 [cited 2020 May];57(11):575-82. Available from: http://vri.cz/docs/vetmed/57-11-575.pdf
77. Sasaki Y, Ishikawa J, Yamashita A, Oshima K, Kenri T, Furuaya K, et al. The complete genomic sequence of *Mycoplasma penetrans*, an intracellular bacterial pathogen in humans. Nucleic Acids Res. 2002 Dec;30(23):5293-300. doi: 10.1093/nar/gkf667
78. Håstein T. Kaldtvannsvibriose [Internet]. Place unknown: Store norske leksikon; 2009 [cited 2020 May]. Available from: https://snl.no/kaldtvannsvibriose
79. Sakmanoglu A, Sayin Z, Ucan US, Pinarkara Y, Uslu A, Erganis O. Comparison of five methods for isolation of DNA from *Mycoplasma cynos*. J. Microbiol. Methods. 2017 Sep;140:70-3. 10.1016/j.mimet.2017.07.003
80. Rivera A, Cedillo L. Mycoplasmas and Nucleases. J. Pure Appl. Microbio. 2015 Oct;9(3):2115-25

81. Madigan M, Martinko J, Bender K, Buckley D, Stahl D. Diversity of Bacteria. In: Churchman K, McFadden N, Williams A, editors. Brock Biology of Microorganisms. 14 ed. England: Pearson; 2015. p. 503-40
82. Urbanczyk H, Ast JC, Higgins MJ, Carson J, Dunlap PV. Reclassification of *Vibrio fischeri, Vibrio logei*, *Vibrio salmonicida* and *Vibio wodanis* as *Aliivibrio fisheri gen. nov., comb. nov., Aliivibrio logei comb. nov., Aliivibrio salmonicida comb nov and Aliivibrio wodanis comb nov.* Int. J. Syst. Evol. Micr. 2007;57:2823-9. doi: 10.1099/ijs.0.65081-0
83. Romalde JL. Photobacterium damselae subsp. piscicida: an integrated view of a bacterial fish pathogen. Int. Microbiol.. 2002 Mar;5(1):3-9. doi: 10.1007/s10123-002-0051-6

APPENDIX

Appendix A: Overview of Samples Used in this Thesis

***Table A.1: Extensive information regarding the samples used in this thesis**. The colors indicate different locations: Blue represents samples from Skjervøy; Orange/yellow samples from Bømlo; Green samples from Chile; and Grey samples from Dagali. Color codes correspond to Table 2.1 and Figure 2.1. Samples were collected for cultivation purposes and DNA analyses. Samples expressed in bold were selected for shotgun sequencing.*

Location	Time of sampling	Individual/Mixed	Sample ID	Storage Buffer /Medium	Kit used DNA extraction	Myco Miseq %	16S qPCR	18S qPCR	DNA conc. ng/µl	Comment
Skjervøy	Oct. 2018	Individual	i1	STAR	LGC Mag Midi	3	29.78	30.11	0.1	
Skjervøy	Oct. 2018	Individual	i4	STAR	LGC Mag Midi	55	28.18	33.05	0.1	
Skjervøy	Oct. 2018	Individual	i5	STAR	LGC Mag Midi	19	28.34	38.35	Too low	
Skjervøy	**Oct. 2018**	**Individual**	**i6**	**STAR**	**LGC Mag Midi**	**93**	**14.49**	**16.87**	**3.5**	**Shotgun**
Skjervøy	**Oct. 2018**	**Individual**	**i7**	**STAR**	**LGC Mag Midi**	**57**	**7.16**	**10.48**	**44.4**	**Shotgun**
Skjervøy	Oct. 2018	Individual	i9	STAR	LGC Mag Midi	13	33.01	38.24	0.1	
Skjervøy	Oct. 2018	Individual	i10	STAR	LGC Mag Midi	64	28.16	29.86	0.2	
Skjervøy	**Oct. 2018**	**Individual**	**i11**	**STAR**	**LGC Mag Midi**	**92**	**10.3**	**15.14**	**20.4**	**Shotgun**
Skjervøy	Oct. 2018	Individual	i12	STAR	LGC Mag Midi	17	30.73	35.74	Too low	
Skjervøy	Oct. 2018	Individual	i14	STAR	LGC Mag Midi	67	28.32	39.03	0.1	
Skjervøy	Oct. 2018	Individual	i15	STAR	LGC Mag Midi	63	29.69	28.61	0.2	
Skjervøy	**Oct. 2018**	**Individual**	**i17**	**STAR**	**LGC Mag Midi**	**99**	**18.26**	**18.29**	**4.4**	**Shotgun**
Skjervøy	Oct. 2018	Individual	i18	STAR	LGC Mag Midi	19	28.57	37.47	0.1	
Skjervøy	Oct. 2018	Individual	i19	STAR	LGC Mag Midi	41	11.25	15.22	17.2	
Skjervøy	**Oct. 2018**	**Individual**	**i20**	**STAR**	**LGC Mag Midi**	**97**	**18.4**	**15.18**	**17.5**	**Shotgun**
Skjervøy	Mar. 2019	Mixed	S1	PBS	Qiagen	20	13.4	32.32	4.9	
Skjervøy	Mar. 2019	Mixed	S2	10 x TE	Qiagen	35	17.9	32.36	0.9	
Skjervøy	Mar. 2019	Mixed	S3	96% EtOH	Qiagen	85	26.53	26.45	0.3	
Skjervøy	**Mar. 2019**	**Mixed**	**S4**	**STAR**	**Qiagen**	**83**	**18.8**	**27.88**	**1.0**	**Shotgun**
Skjervøy	Mar. 2019	Mixed	S5	DNA shield	Qiagen	94	22.4	N/A	1.0	
Skjervøy	**Mar. 2019**	**Mixed**	**S6**	**RNA later**	**Qiagen**	**94**	**20.84**	**26.18**	**1.2**	**Shotgun**
Skjervøy	Mar. 2019	Mixed	S7	MycoBroth	Qiagen	1.6	10.88	35.55	14.0	
Skjervøy	Mar. 2019	Mixed	S8	MycoPlate	Qiagen	0	11.34		34.0	

Bømlo	Sept. 2019	Mixed	B1	MycoBroth	Qiagen	2	24.96	19.01	8.9	Filtered
Bømlo	Sept. 2019	Mixed	B2	MycoBroth	Qiagen	6.7	8.54	21.69	19.1	Filtered
Bømlo	Sept. 2019	Mixed	B3	MycoBroth	Qiagen	0.3	14.41	20.72	5.3	Filtered
Bømlo	Sept. 2019	Mixed	B4	MycoBroth	Qiagen	0.2	12.19	20.39	3.9	Filtered
Bømlo	Sept. 2019	Mixed	B5	MycoBroth	Qiagen	0.9	29.01	18.81	8.0	Filtered
Bømlo	Sept. 2019	Mixed	B6	MycoBroth	Qiagen	2.4	17.84	19.64	2.7	Filtered
Bømlo	Sept. 2019	Mixed	B7	MycoBroth	Qiagen	0.3	13.38	20.51	4.3	Filtered
Bømlo	Sept. 2019	Mixed	B8	MycoBroth	Qiagen	0.1	12.25	19.61	6.5	Filtered
Bømlo	Sept. 2019	Mixed	B9	MycoBroth	Qiagen	0.1	33.33	19.48	7.0	Filtered
Bømlo	Sept. 2019	Mixed	B10	MycoBroth	Qiagen	0.3	11.97	19.56	6.8	Filtered
Bømlo	**Sept. 2019**	**Mixed**	**B11**	**RNA later**	**Qiagen**	**52.3**	**19.87**	**19.93**	**2.0**	**Shotgun**
Bømlo	**Sept. 2019**	**Mixed**	**B12**	**RNA later**	**Qiagen**	**28**	**19.48**	**19.83**	**1.6**	**Shotgun**
Bømlo	Sept. 2019	Mixed	B13	RNA later	Qiagen	30.4	36.19	20.02	1.7	
Bømlo	**Sept. 2019**	**Mixed**	**B14**	**RNA later**	**Qiagen**	**28**	**21.75**	**19.96**	**2.0**	**Shotgun**
Bømlo	Sept. 2019	Mixed	B15	RNA later	Qiagen	13.4	21.81	19.58	1.6	
Bømlo	Sept. 2019	Mixed	B16	RNA later	Qiagen	6.1	20.22	20.25	1.3	
Bømlo	Sept. 2019	Mixed	B17	RNA later	Qiagen	37.3	34.51	20.59	1.3	
Bømlo	**Sept. 2019**	**Mixed**	**B18**	**RNA later**	**Qiagen**	**45.7**	**20.41**	**18.2**	**3.6**	**Shotgun**
Bømlo	Sept. 2019	Mixed	B19	RNA later	Qiagen	31.7	22.02	18.95	2.5	
Filtered Bømlo	Sept. 2019	Filtered	F1	MycoBroth	LGC Mag Midi	N/A	N/A	N/A	Too low	
Filtered Bømlo	Sept. 2019	Filtered	F2	MycoBroth	LGC Mag Midi	N/A	N/A	N/A	Too low	
Filtered Bømlo	Sept. 2019	Filtered	F3	MycoBroth	LGC Mag Midi	N/A	N/A	N/A	5.6	Sanger
Filtered Bømlo	Sept. 2019	Filtered	F4	MycoBroth	LGC Mag Midi	N/A	N/A	N/A	Too low	
Filtered Bømlo	Sept. 2019	Filtered	F5	MycoBroth	LGC Mag Midi	N/A	N/A	N/A	Too low	
Filtered Bømlo	Sept. 2019	Filtered	F6	MycoBroth	Phenol-chloroform	N/A	N/A	N/A	Too low	
Filtered Bømlo	**Sept. 2019**	**Filtered**	**F7**	**MycoBroth**	**Phenol-chloroform**	**N/A**	**37.64**	**N/A**	**0.6**	**Sanger, Shotgun**
Filtered Bømlo	Sept. 2019	Filtered	F8	MycoBroth	Phenol-chloroform	N/A	N/A	N/A	Too low	
Filtered Bømlo	Sept. 2019	Filtered	F9	MycoBroth	Phenol-chloroform	N/A	24.41	N/A	27.9	Sanger
Filtered Bømlo	Sept. 2019	Filtered	F10	MycoBroth	Phenol-chloroform	N/A	N/A	N/A	Too low	Plated on agar
Chile	Nov. 2019	N/A	C1	RNA later	Qiagen	0	26.92	16.33	34.9	
Chile	Nov. 2019	N/A	C2	RNA later	Qiagen	0	25.23	21.2	2.0	
Chile	Nov. 2019	N/A	C3	RNA later	Qiagen	0	22.18	18.63	19.5	
Chile	Nov. 2019	N/A	C4	RNA later	Qiagen	0	22.52	17.66	2.2	
Chile	Nov. 2019	N/A	C5	RNA later	Qiagen	0	25.96	15.35	31.3	

Chile	Nov. 2019	N/A	C6	RNA later	Qiagen	0	24.03	15.15	35.7	
Chile	Nov. 2019	N/A	C7	RNA later	Qiagen	0	18.42	13.78	21.4	
Chile	Nov. 2019	N/A	C8	RNA later	Qiagen	0	23.49	13.33	51.0	
Chile	**Nov. 2019**	**N/A**	**C9**	**RNA later**	**Qiagen**	**0**	**17.15**	**17.47**	**6.6**	**Shotgun**
Chile	Nov. 2019	N/A	C10	RNA later	Qiagen	0	26.68	14.95	43.6	
Chile	Feb. 2020	N/A	C11	RNA later	Qiagen	0	26.47	19.41	6.8	
Chile	Feb. 2020	N/A	C12	RNA later	Qiagen	0	30.09	21.99	5.2	
Chile	Feb. 2020	N/A	C13	RNA later	Qiagen	0	30.67	23.96	0.7	
Chile	Feb. 2020	N/A	C14	RNA later	Qiagen	0	24.31	18.78	13.1	
Chile	Feb. 2020	N/A	C15	RNA later	Qiagen	0	26.23	18.5	15.6	
Chile	Feb. 2020	N/A	C16	RNA later	Qiagen	0	28.61	20.63	2.8	
Chile	Feb. 2020	N/A	C17	RNA later	Qiagen	0	26.92	15.9	60.0	
Chile	Feb. 2020	N/A	C18	RNA later	Qiagen	0	23.63	15.36	55.0	
Chile	Feb. 2020	N/A	C19	RNA later	Qiagen	0	25.04	18.16	17.8	
Chile	Feb. 2020	N/A	C20	RNA later	Qiagen	0	19.71	18.65	11.7	
Dagali	Summer 2019	Individual	D1	STAR	Qiagen	N/A	30.14	26.73	0.5	
Dagali	Summer 2019	Individual	D2	STAR	Qiagen	N/A	N/A	N/A	0.7	
Dagali	Summer 2019	Individual	D3	STAR	Qiagen	9	31.55	31.74	1.2	
Dagali	Summer 2019	Individual	D4	STAR	Qiagen	N/A	32.37	55.24	2.0	
Dagali	Summer 2019	Individual	D5	STAR	Qiagen	N/A	30.78	29.28	0.2	
Dagali	Summer 2019	Individual	D6	STAR	Qiagen	N/A	31.12	29.65	0.7	
Dagali	Summer 2019	Individual	D7	STAR	Qiagen	N/A	30.59	33.57	0.4	
Dagali	Summer 2019	Individual	D8	STAR	Qiagen	15	57.63	27.415	1.1	

Appendix B: Final Concentrations of Enriched Mycoplasma Growth Broth

Table B.1: Final concentrations of Enriched Mycoplasma Growth Broth.

Component	**g/L**
Bacterial peptone	10.0 g/L
Lab Lemco powder	10.0 g/L
Sodium Chloride	5.0 g/L
Mineral Supplement	0.5 g/L
Horse Serum	20 %
Yeast Extract (25 % w/v solution)	10 %
Thallous Acetate	250 mg/L
Penicillin G	200 000 IU/L (=200 mg)
DNA from Salmon	20 mg/L

Appendix C: T-tests Conducted in Excel

Phylum: Proteobacteria

Table C.1: T-test of the Proteobacteria data set conducted by Excel.

Proteobacteria	*Variable 1*	*Variable 2*
Mean	0.975791504	0.663505584
Variance	0.001881607	0.023135117
Observations	18	9
Hypothesized Mean Difference	0	
df	9	
t Stat	6.037848507	
P(T<=t) one-tail	9.66664E-05	
t Critical one-tail	1.833112933	
P(T<=t) two-tail	**0.000193333**	
t Critical two-tail	2.262157163	

Phylum: Tenericutes

Table C.2: T-test of the Tenericutes data set conducted by Excel.

Tenericutes	*Variable 1*	*Variable 2*
Mean	5.3674E-05	0.30271528
Variance	9.2995E-09	0.02062785
Observations	18	9
Hypothesized Mean Difference	0	
df	8	
t Stat	-6.3219577	
P(T<=t) one-tail	0.00011368	
t Critical one-tail	1.85954804	
P(T<=t) two-tail	**0.00022735**	
t Critical two-tail	2.30600414	

Genus: Aliivibrio

Table C.3: T-test of the Aliivibrio data set conducted by Excel.

Aliivibrio	*Variable 1*	*Variable 2*
Mean	0.736774	0.314393
Variance	0.092442	0.025981
Observations	18	9
Hypothesized Mean Difference	0	
df	25	
t Stat	4.715757	
P(T<=t) one-tail	3.89E-05	
t Critical one-tail	1.708141	
P(T<=t) two-tail	**7.78E-05**	
t Critical two-tail	2.059539	

Genus: Mycoplasma

Table C.4: T-test of the Mycoplasma data set conducted by Excel.

Mycoplasma	*Variable 1*	*Variable 2*
Mean	0	0.30271
Variance	0	0.020626
Observations	18	9
Hypothesized Mean Difference	0	
df	8	
t Stat	-6.32325	
P(T<=t) one-tail	0.000114	
t Critical one-tail	1.859548	
P(T<=t) two-tail	**0.000227**	
t Critical two-tail	2.306004	

Genus: Vibrio

Table C.5: T-test of the Vibrio data set conducted by Excel.

Vibrio	*Variable 1*	*Variable 2*
Mean	0.061486	0.073814
Variance	0.007985	0.004329
Observations	18	9
Hypothesized Mean Difference	0	
df	21	
t Stat	-0.40544	
P(T<=t) one-tail	0.344628	
t Critical one-tail	1.720743	
P(T<=t) two-tail	**0.689256**	
t Critical two-tail	2.079614	

Genus: Photobacterium

Table C.6: T-test of the Photobacterium data set conducted by Excel.

Photobacterium	*Variable 1*	*Variable 2*
Mean	0.140569	0.273003
Variance	0.070063	0.060926
Observations	18	9
Hypothesized Mean Difference	0	
df	17	
t Stat	-1.28257	
P(T<=t) one-tail	0.108428	
t Critical one-tail	1.739607	
P(T<=t) two-tail	**0.216856**	
t Critical two-tail	2.109816	

Appendix D: Unweighted UniFrac PCoA plot

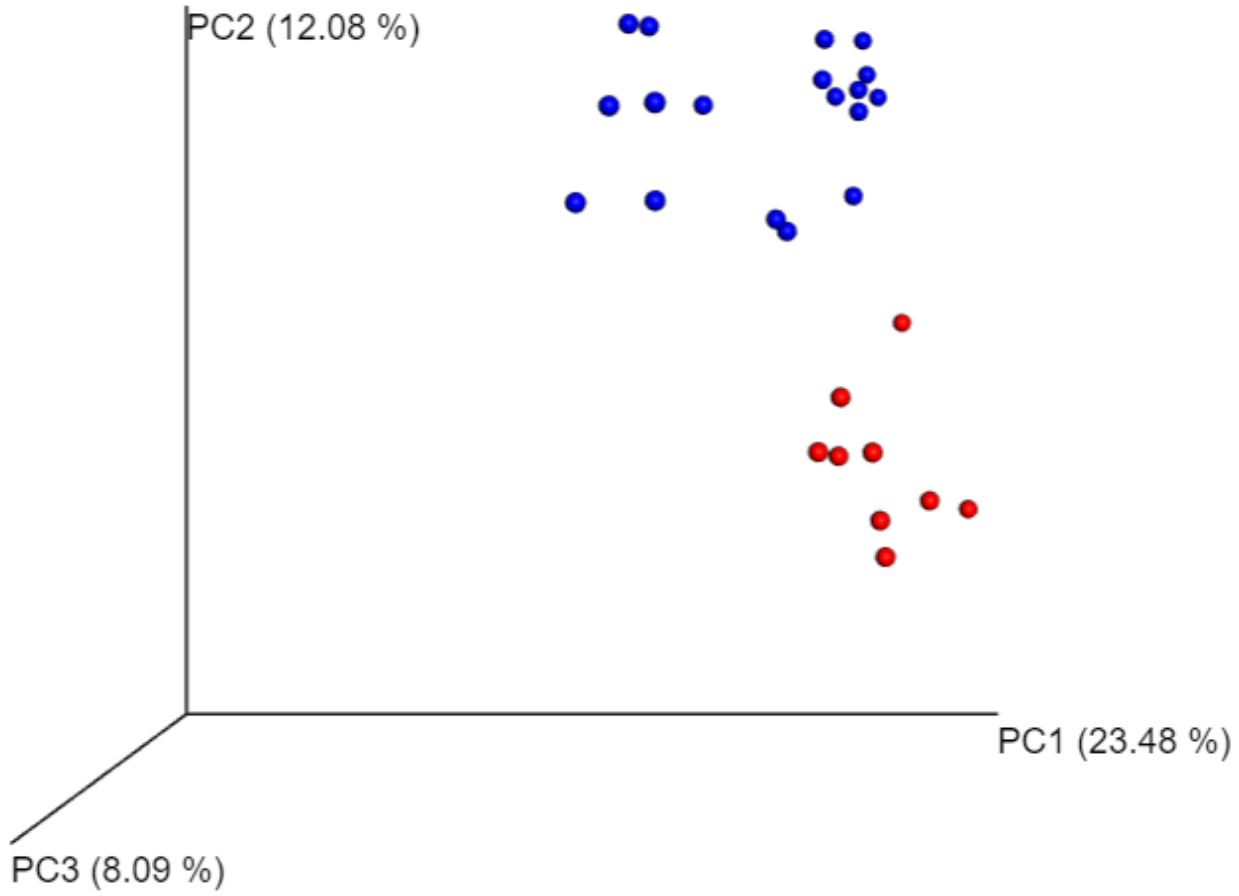

Figure D.1: Beta-diversity between microbial communities in the gut of farmed salmon. *Differences in the gut microbiota of salmon from Bømlo and Chile were analyzed by Principal Coordinate Analysis (PCoA) plots generated by QIIME. The PCoA depicted in this figure is calculated by unweighted UniFrac measures. Each dot represents one sample; red dots represent salmon from Bømlo, and blue dots represent salmon from Chile.*

Appendix E: Taxonomic Annotation of Bacterial Shotgun Sequences at Species Level

Table E.1: Taxonomic annotation of bacterial shotgun sequences at species level. *Shotgun sequences obtained from gut content of salmon from three different locations were assigned taxonomy using the MG-RAST application.*

Sample location	Species	Percentage
Bømlo	*Mycoplasma agalactiae*	0.9%
	Mycoplasma alligatoris	1.4%
	Mycoplasma arthritidis	1.2%
	Mycoplasma capricolum	1.2%
	Mycoplasma crocodyli	1.8%
	Mycoplasma fermentans	2.0%
	Mycoplasma gallisepticum	13.1%
	Mycoplasma genitalium	6.5%
	Mycoplasma hominis	0.9%
	Mycoplasma hyopneumoniae	0.8%
	Mycoplasma leachii	0.4%
	Mycoplasma mobile	3.4%
	Mycoplasma mycoides	1.4%
	Mycoplasma penetrans	51.9%
	Mycoplasma pneumoniae	6.5%
	Mycoplasma pulmonis	2.0%
	Mycoplasma synoviae	1.3%
	Mycoplasma conjunctivae	1.2%
	Mycoplasma hyorhinis	1.0%
	Mycoplasma bovis	0.9%
Chile	*Aliivibrio salmonicida*	22.0%
	Vibrio fischeri	78.0%
Skjervøy	*Mycoplasma agalactiae*	0.9%
	Mycoplasma alligatoris	1.4%
	Mycoplasma arthritidis	1.1%
	Mycoplasma capricolum	1.2%
	Mycoplasma crocodyli	1.8%

	Mycoplasma fermentans	2.0%
	Mycoplasma gallisepticum	13.6%
	Mycoplasma genitalium	6.5%
	Mycoplasma hominis	0.8%
	Mycoplasma hyopneumoniae	0.8%
	Mycoplasma leachii	0.4%
	Mycoplasma mobile	3.8%
	Mycoplasma mycoides	1.3%
	Mycoplasma penetrans	51.3%
	Mycoplasma pneumoniae	6.4%
	Mycoplasma pulmonis	2.0%
	Mycoplasma synoviae	1.3%
	Mycoplasma conjunctivae	1.4%
	Mycoplasma hyorhinis	1.0%
	Mycoplasma bovis	1.0%

Appendix F: Microbial Identification of Primary MycoBroth Cultures

Table F.1: Identity of primary Mycoplasma Broth cultures determined by BLAST*. DNA extracted from cultures were sent to Eurofins, Germany, for Sanger sequencing. The sequence output in FASTA format was aligned to the Nucleotide Collection (nr/nt) database to identify the microbe(s). The best hit (lowest E-value) for each primer is included in the table.*

Sample	Primers	BLAST hit	% Query Cover	E-value	% Identity
F3	341F	*Staphylococcus* sp.	98	0.0	99
	806R	No significant similarity found	N/A	N/A	N/A
	MycoFrw	*Staphylococcus* sp.	97	4e-141	99
	MycoRev	*Staphylococcus* sp.	98	1e-146	99
F7	MycoFrw	*Cinara fresei*	49	2e-177	100
	MycoRev	Uncultured bacterial clone	98	4e-146	98
		Uncultured *Mycoplasma* sp.	98	2e-134	95
F9	341F	*Micrococcus luteus*	99	0.0	99
	806R	*Micrococcus* sp.	60	0.0	99

Norges miljø- og biovitenskapelige universitet
Noregs miljø- og biovitskapelege universitet
Norwegian University of Life Sciences

Postboks 5003
NO-1432 Ås
Norway