Norwegian University of Life Sciences

Master's Thesis 2018 60 ECTS

Faculty of Chemistry, Biotechnology and Food Science
Supervisor: Knut Rudi

Maternal probiotic consumption and its effects on children's oral microbiota

Siril Malene Isaksen

Biotechnology
Molecular Biology

Acknowledgments:

This master thesis was performed at the Norwegian University of Life Science, at the Faculty of Chemistry, Biotechnology and Food Sciences, with Professor Knut Rudi as supervisor.

I would first like to thank Professor Knut Rudi for introducing me to this thesis. Thank you so much for all your guidance and for always being available to answer my questions. Your knowledge and positivity are a big inspiration.

I would also like to give a special thanks to both Inga Leena Angell and Jane Ludvigsen for always being able to give me a helping hand in the laboratory or answering my questions. You have both taught me so much.

Finally, a big thanks to my friends and family for your support, thank you for always encouraging me. A special thanks to my boyfriend Andreas for always believing in me.

Ås, 2019

Siril Malene Isaksen

Sammendrag

Omtrent halvparten av alle cellene i kroppen vår er av prokaryotisk opprinnelse, og vi vet at den mikrobielle sammensetningen påvirker helsen vår. Denne kunnskapen kan brukes til å endre den mikrobielle sammensetningen slik at den fungerer til vår fordel. Den orale mikrobiotaen er den nest mest omfattende etter mage-tarmkanalens mikrobiota. Den beskytter kroppene mot uønsket stimulering fra utsiden, og tidligere studier har vist at probiotisk behandling av gravide fører til en reduksjon i tilfeller av atopisk eksem hos deres barn. En av de mulige veiene de probiotiske bakteriene overføres fra mor til barn er gjennom munnen. Formålet med denne oppgaven er derfor å avgjøre om inntak av probiotika kan endre den orale mikrobiotaen hos mødre og deres barn.

De 878 prøvene som ble brukt i denne oppgaven var fra Prevention of Allergy among Children in Trondheim (Pro-PACT) studien. Spyttprøver ble samlet inn fra både mødre og barn både 10 dager og 3 måneder etter fødsel. Sekvensering ble brukt til å bestemme sammensetningen til den orale mikrobiotaen i disse prøvene.

Den probiotiske behandlingen til mødrene ser ikke ut til å ha noen større effekt på den orale mikrobiotaen hos barna. Sammensetningen av barns orale mikrobiota korresponderte med funn i tidligere studier, de mest omfattende phylaene var *Firmicutes*, *Proteobacteria*, *Fusobacteria* og *Actinobacteria*. En av de mindre skilnadne som ble funnet hos barna var i slekten *Gemella*. Barna fra de mødrene som fikk placebobehandling hadde dobbelt så mange sekvenser fra denne slekten enn de andre barna. Det probiotiske forbruket hos mødre synes å påvirke deres orale mikrobiota basert på de 20 artene det var mest av i alle prøvene.

Abstract

Approximately half of the cells in our bodies are of prokaryotic origin, and we know that the microbial composition influences our health. This knowledge can be used to work out how we can alter the microbial composition to our benefit. The oral microbiota is the second most abundant one after the gastrointestinal gut, and it protects our bodies from unwanted stimulation from the outside. Earlier studies have shown that probiotic consumption by pregnant women lead to a decrease in cases of atopic dermatitis in their children. One of the possible transmission paths of the probiotic bacteria from mother to child is though the oral cavity. The aim of this thesis is therefore to determine if probiotic intervention can alter the oral microbiota in mothers and their children.

The 878 samples used in this thesis were from the Prevention of Allergy among Children in Trondheim (Pro-PACT) study. Saliva samples were collected at 10 days postpartum and 3 months postpartum from both the mothers and their children. Sequencing were used to determine the oral microbiota in these samples.

The probiotic treatments of the mothers did not seem to have any major effect on the oral microbiota in the children. The composition of the children's oral microbiota corresponded to findings in earlier studies, with the most abundant phyla being *Firmicutes, Proteobacteria, Fusobacteria* and *Actinobacteria.* One small difference found in the children were the genus *Gemella* having twice as high quantities in the children from mothers who received placebo treatment, then in the children whose mothers were treated with probiotics. The probiotic consumption in mothers seems to influence their oral microbiota based on the 20 most abundant species in the community.

Table of contents

1 Introduction

1.1 The development of oral microbiome

Bacteria were first discovered by the Dutch scientist Antonie van Leeuwenhoek, also known as "the father of microbiology". In the 17th century, van Leeuwenhoek studied dental plaque in self-made microscopes and discovered small objects we now know to be bacteria. Our methods for finding and analysing bacteria has developed in the over 300 years since van Leeuwenhoek did his discovery. We are now able find and analyse all the bacteria present in our bodies, also known as our microbiome. The Human Microbiome project was launched in 2007. This project aimed to define our microbiome and to investigate how these microorganisms are involved in human health (Turnbaugh et al., 2007). A study from 2016 stated that over 200g of the bodyweight of an average sized human (70kg, 170cm) is microorganisms. The study also redefined the number of microbial cells in the microbiome to be ~3.8 x 10^{13} (Sender et al., 2016). Another study stated that the number of human cells in our bodies equals the number of microorganisms present. The study used a mathematical approach to count the number of microorganism (Bianconi et al., 2013).

The human oral microbiome is defined as all the microorganisms that are found on or in the human oral cavity and its contiguous extensions. The microbiota in the oral cavity is the second most abundant one after the gastrointestinal tract. Its composition and interaction are important for the protection of the human body against unwanted stimulation from the outside. After all, the mouth and the nose are the gateway to both the gut and the respiratory system, so the oral microbiota has a direct influence on both the immune response and our metabolism (Verma et al., 2018).

The Human Oral Microbiome Database (HOMD) is an organ-specific microbial database inspired by the Human Microbiome Project. The site was last updated on 20/2-2019 and included at that time the information on 770 prokaryotic species. 57% of these species are officially named, 13% are cultivated but unnamed and 30% uncultivated phylotypes. The

HOMD site also contains genomes for 475 taxa, about 62% of all taxa described on the site. 85% of the cultivated taxa has their genomes on the site (Escapa et al., 2018).

Distinct microbial communities colonize different oral structures and tissues, and a core microbiome is established at an early age. The core human oral microbiome is made up of 6 different phylum and constitutes 96% of the total oral bacteria in a healthy oral cavity. The 6 different phyla are *Firmicutes* (36.7%), *Bacteriodetes* (17.1%), *Proteobacteria* (17.1%), *Actinobacteria* (11.6%), *Spirochaetes* (7.9%) and *Fusobacteria* (5.2%) (Dewhirst et al., 2010).

There are many factors which can alter the oral microbiota. Maternal influence plays a major role in the formation of the new-born's microbiome, but diet and medication later on can also have a huge influence on its composition. Another important factor is the environment (climate, geographical conditions and air quality) (Verma et al., 2018).

1.1.1 Colonization of the oral microbiome from birth to adulthood

A significant number of the microorganisms that a new born is exposed to is of maternal origin. Type of delivery greatly influences what kind of microorganisms the new born encounters to start with (Dominguez-Bello et al., 2010). Children born with Caesarean section tends to have bacterial communities similar to the mother's skin to start with, and usually has increased abundances of *Propionbacterium, Veillonella* and *Staphylococcus* spp. in the oral cavity. Those born vaginally similarly have bacterial communities resembling the mother's vaginal microbiota, such as *Prevotella, Lactobacillus* and *Bacteroides* spp. (Gomez & Nelson, 2017), (Sampaio-Maia, B. & Monteiro-Silva, F., 2014).

A new born is exposed to a large number of microorganisms in the first hours after birth, but only a few of these are able to colonize the baby. The first few microorganisms that are able to colonize the oral cavity do so after only 24 hours. The most frequent initial oral colonizers are Gram positive cocci, including *Streptococcus, Actinomyces* and *Staphylococcus*, all three

genera are facultative anaerobes. These initial colonizers play a major role in the development of the oral microbiota due to their metabolism. The products that the colonizers generate and excrete promotes an alteration of the environment. This alteration will then benefit the growth of other species such as *Fusobacteria* and *Veillonella,* two strictly anaerobic genra. The change in the oral environment will eventually lead to a more complex and stable microbial community (Sampaio-Maia, B. & Monteiro-Silva, F., 2014).

The infant's microbial community grows and changes with the baby, and will be significantly different from the mother already at around 5 months of age. This difference in microbiota from the mothers is mostly due to different environmental exposure that the child encounters in the first months after birth. The environmental factors that plays the biggest roles in this period includes contact with both humans and animals, ingestion of food, teething and hygienic habits. The microbiota at approximately 5 months of age is mostly comprised of bacteria from the six phyla: *Proteobacteria, Firmicutes, Actinobacteria, Bacteriodetes, Fusobacteria* and *Spirochaetes*, with *Streptococcus, Haemophilus, Neisseria* and *Veillonella* being the most frequent genera (Dzidic et al., 2018).

At around 5 months of age, the children often have a greater microbial diversity in the oral cavity than their parents, even though they have fewer oral microorganisms in total (Cephas et al., 2011). Many of the microorganisms found in the oral cavity at around 5 months of age produces immunoglobin A (IgA) proteases that degrade the IgA which is secreted from the breast milk. It is believed that this trait ensures the survival of these species while the oral cavity is an IgA-rich environment (Sampaio-Maia, B. & Monteiro-Silva, F., 2014) (Cole et al., 1994).

The bacterial composition in the infants continues to change and develop as the child grows. One important event in the maturing process is the teething period. New teeth give the microorganisms new adhesion surfaces, promoting the growth of other microorganisms than those already present in the oral cavity. The bacterial community changes from facultative gram-positive or aerobic until it contains mostly anaerobic gram-negative bacteria as the child becomes an adult (Tanner et al., 2002), (Sampaio-Maia, Benedita & Monteiro-Silva, Filipa, 2014).

1.2 DNA based approaches to study microbiota

About 30% of all the oral microorganisms are uncultivable, making molecular methods that are cultivation-independent the preferable approach for getting the best possible representation of the microbiome. One such method is the 16S rRNA gene-analysis (Sampaio-Maia, B. & Monteiro-Silva, F., 2014).

1.2.1 The 16S rRNA gene as a genetic marker

The prokaryotic cell has, as most living cells do, a ribosome which is responsible for all protein synthesis in the cell. The ribosome consists of two subunits, a large one (50S) and a smaller one (30S). The small subunit binds and reads the DNA while the large subunit translates mRNA into proteins. The 16S rRNA gene and 19 proteins together make up the 30S subunit of the prokaryotic ribosome. The gene is normally about 1500 bp long, and its structure is highly conserved.

The conserved structure of the 16S rRNA gene, and its conserved sequence-areas is exactly what makes it ideal for analysis of bacterial communities.
It is vital that a cell's protein synthesis is executed correctly, so any major mutation to the ribosome will most likely be fatal to the organism. This means that the genetic structure and function of the gene has been relatively unchanged over time. Small, random changes to the sequence which has not compromised the 16S rRNA gene's functionality can therefore be

used as an evolutionary measurement. There are certain areas on the gene in which these small random changes can be found, and it is these areas that are targeted for microbial studies. Closely related taxa tend to have similar mutations, making it possible to construct phylogenetic trees from this information (Willey et al., 2014), (Janda & Abbott, 2007).

The 16S rRNA gene consists of 9 highly conserved regions and 9 regions that vary among different bacterial species. The conserved regions can be used to analyse higher taxa, as well as design primers for amplification for multiple taxa at the same time. The variable regions give information which makes it possible to differentiate between different taxa. These variable regions are shown in figure 1 where the two hypervariable regions targeted in this thesis are highlighted. The primers used in this thesis encompasses the variable regions V3 and V4.

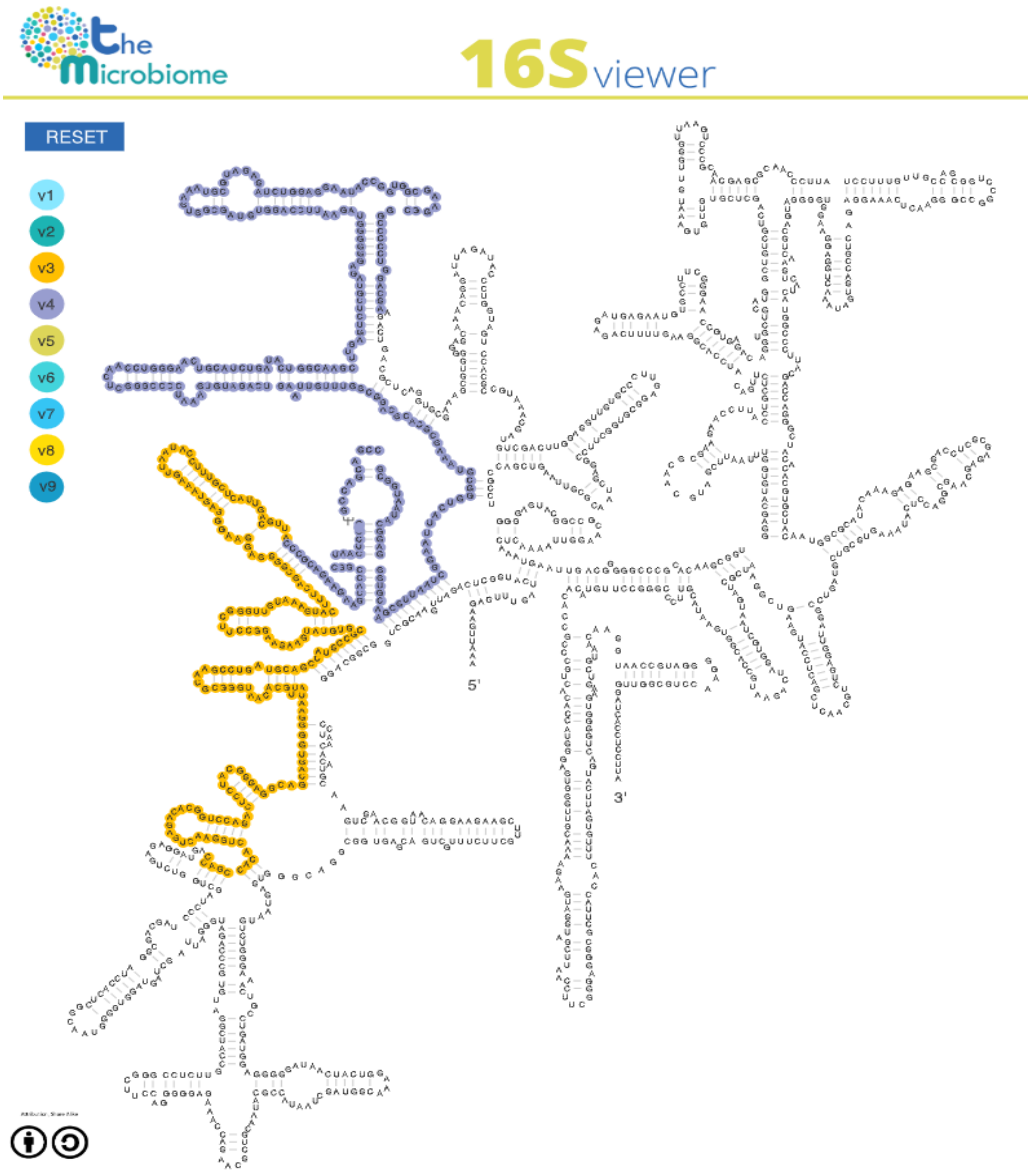

*Figure 1: The 16S rRNA gene with the variable regions V3 and V4 (**http://themicrobiome.com/media/16S_viewer.cfm**).*

1.2.2 Polymerase Chain Reaction

Polymerase Chain Reaction (PCR) is a process that amplifies DNA. The process is often used to get a large enough sample of a particular DNA-sequence for further analysis. The reaction-sample consists of four components: target DNA, DNA primers, deoxynucleotide triphosphate (dNTPs) and heat-stable DNA polymerase enzymes. The PCR process itself consists of 3 different steps; denaturation, annealing and elongation, and exposing the sample to different, specific temperatures brings on the next step in the process.

Denaturation is the part of the process where the target DNA's double strand is cleaved into two single strands (ssDNA). This is achieved by using a relatively high temperature (95°C) to break the hydrogen bonds between the two DNA strands, leaving two separate DNA strands. The solution is then cooled to a temperature where the annealing-step can start.

Annealing is the step where the DNA primers binds to the part of the sequence we are interested in amplifying. The DNA primers chosen for this task are specific to this part of the DNA sequence. The location on the DNA sequence where the primers attach is the starting point of the next step in the process: the elongation. The solution is heated once again in order to bring on this next step.

Elongation is the step where the heat-stable DNA polymerase enzymes attaches to the area on the sequence where the primers are bounded. The polymerases then use the dNTPs to synthesise a new DNA strand. This strand is now complementary to the targeted DNA sequence, and so the amount of this sequence present in the sample is at this point doubled. These three steps are then repeated in several cycles in order to generate the wanted amount of target DNA strands, doubling the amount of target DNA sequences in each run.

Quantitative PCR, or qPCR, is a type of PCR where the amount of target DNA can be measured in real-time. This is possible due to a fluorescent dye being added to the sample, a molecule that emits fluorescent light signal only when attached to double stranded DNA (dsDNA). This signal is measured after each cycle by the qPCR machine and it tells us after

which cycle the signal was big enough to be detected. The number of cycles needed to create a detectable signal is called a Ct-value. The smaller the Ct-value, the more DNA is present in the original sample. One problem with this approach is that the fluorescent molecule does not distinguish between target DNA and potential contaminants in the sample.

Digital droplet PCR (ddPCR) is another PCR method allowing for measurement of target DNA in a sample. This method is more accurate than the qPCR as the sample is divided into small water/oil-droplets. Each droplet will either contain a small amount of target DNA, or none at all. The fluorescence of each droplet is measured after the probe-based PCR amplification. A droplet will emit a fluorescent signal if a molecule of target DNA is present. The machine will then mark the droplet as positive. Droplets that does not give a signal will be marked as negative. The amount of target DNA is measured by calculating the ratio of positive droplets against the total number of droplets (Pinheiro et al., 2012).

1.2.3 Next-generation sequencing

Frederick Sanger is a scientist that has contributed a lot to the fields of molecular and microbial biology. He introduced a method for determining the sequence of nucleotides in DNA in 1977. The method involved the use of dideoxynucleotide triphosphates (ddNTPs) to terminate the replication of new strands at specific base pairs. Each of the four ddNTPs has its own separate synthesis, making strands of different lengths all ending with the same nucleotide in each of the four reactions. The final sequence is generated through the reading of a gel where the four ddNTPs reaction samples is run through different gel lanes. This creates band patterns showing the placement of the different on nucleotides on the sequence (Sanger et al., 1977).

Although Sanger sequencing is a good method for sequencing, it is time-consuming and is quite expensive. Cheaper and faster sequencing methods has since been developed, and we call these next-generation sequencing methods. One such method is illumina sequencing (McCombie et al., 2018).

Illumina sequencing can generate millions of sequences from different species in a single run and it is one of the leading next generation sequencing techniques. The illumina process uses sequencing by synthesis to generate the sequences.

The illumina process consist of three main steps: sample preparation, generation of clusters and sequencing. The sample preparation is the attachment of adapters to both ends of the targeted DNA fragments. The ends of these adapters are complementary to the oligonucleotides on the surface of the reaction chamber, also known as a flow cell. Oligonucleotides are fragments of genomic DNA, and there are two types of oligonucleotides arranged as high-density single-molecule arrays on the flow cell.

Clusters on the flow cell is generated through one of the adapters on the targeted DNA fragment attaching itself to its complimentary oligonucleotide on the surface of the flow cell. A complimentary strand is generated with the target DNA as a template before the original strand is cleaved of and washed away. The new DNA strands are then amplified through bridge amplification.

Bridge amplification is initiated when the adapters on the free ends of the fragments attaches to the other type of oligonucleotide on the flow cell, creating a bridge from one oligonucleotide to another. All the fragments are amplified in this bridge formation, and then denatured into two ssDNA strands at the end. This process is repeated simultaneously all over the flow cell until millions of clusters of new DNA strands are made. All the reverse strands are cleaved off at the end of this step, leaving only the forward strands to be sequenced.

The sequencing process is accomplished through the use of fluorescently labelled dNTPs. Each dNTP has its own unique fluorophore which generates a signal when it attaches to a DNA fragment and is excited by a laser. The fluorescing molecule blocks the 3’OH group, this ensures that only one nucleotide can attach to a strand at the time. This feature reduces the risk of misincorporation. A laser-induced excitation of the fluorophores after each incorporation-cycle makes it possible for a camera to detect which bases are where on the DNA strands. The last step in the sequencing cycle is the addition of a chemical for removal

of the fluorescent molecule on the attached dNTP. This allows the next dNTP to attach to the DNA fragment (Bentley et al., 2008). The whole process can be seen in figure 2.

Overall this is a very accurate sequencing method that generates a lot of sequences in a very short time compared to Sanger sequencing. The sequencing data from the illumina method can be used to analyse bacterial communities through operational taxonomic units.

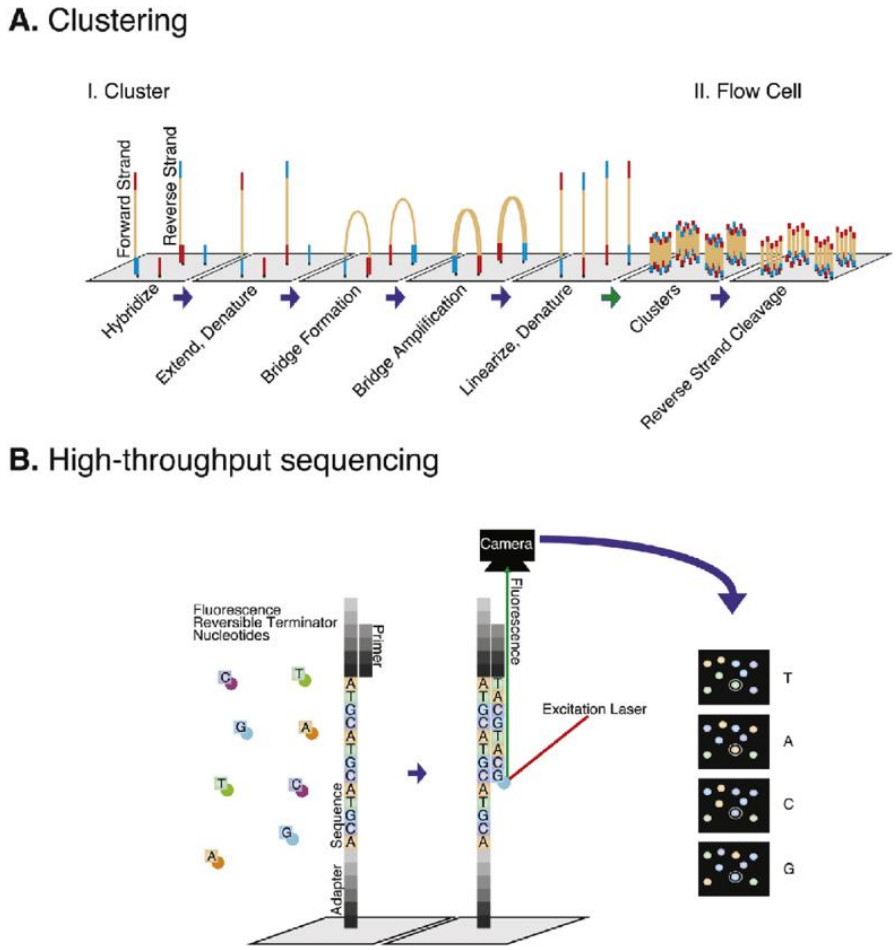

Figure 2: This is a schematic overview of the illumina process, from attachment to the flow cell, to the camera reads of the fluorescent dNTPs (https://www.researchgate.net/figure/Illumina-sequencing-and-data-processing-workflow-A-Denaturated-NGS-library-fragments_fig2_303984059).

1.3 Data analysis on 16S rRNA data

The emergence of next generation sequencing techniques has made it possible to generate massive amounts of data in a short time. Analysing that much data by hand would be time-consuming. More complex bioinformatical tools were developed to ease the interpretation of the generated data.

These bioinformatic tools are often constructed as software applications where the sequencing data from platforms like illumina can be processed. The software application to be used for this thesis is QIIME (Quantitative Insights In to Microbial Ecology).

QIIME is a bioinformatics pipeline where raw DNA sequencing data can be analysed. The pipeline is constructed to favour the use of multiple methods assigned to the different steps of the analysis. The process itself starts with a pre-processing step where the sequences are aligned according to their specific sample primers, and then the removal of the primers. The next step is clustering of the sequences to reveal the phylogenetic relations between them. The results from the clustering is then assigned into operational taxonomic units.

1.3.1 Operational taxonomic units

Operational taxonomic units, or OTU's, are markers that together constructs a phylogenetic tree. A phylogenetic tree aims to reconstruct the evolutionary relationships between species and is in this thesis the tree is based on the 16S rRNA-gene.

The 16S rRNA genes are hierarchical clustered into groups, or OTU's, based on how similar they are. A similarity threshold is usually set to 97% for the sequences to be put in a taxon together. Each OTU therefore represents a group of closely related taxa, and the different groups are assigned to their OTU's based on a data-reference group.

An OTU table is finally constructed when all the sequences have been assigned to their respective OTU's. This table shows what kind of microorganisms that were present in the different samples. This data can then be used to perform different kinds of statistical analysis, such as α- and β-diversity.

1.3.2 Alpha diversity

Alpha diversity is a statistical analysis of the number of different species present within a sample, also known as the biodiversity. Species richness can be calculated in a number of different ways, and the different calculations are called diversity indices. The diversity indices calculate the species richness as well as different aspects of the composition of the community (Whittaker, 1972).

Chao 1 is a diversity index that looks at the species richness in a community. It is based on the abundance of the different species present in the sample. Another way to calculate the biodiversity is the Shannon index. This index looks at the evenness in a sample/ community. It analyses the relative numbers of each species in the sample with regards to the other species.

1.3.3 Beta diversity

Beta diversity differs from alpha diversity in that it compares the diversities within different communities with each other. Like α-diversity, the β-diversity can be measured in several different ways. One way to statistically calculate β-diversity is the Bray-Curtis dissimilarity.

Bray-Curtis dissimilarity is a quantitative measurement of the dissimilarity between two different communities. It does so by looking at the abundance of the different species within each community. The number of shared species between communities is measured up against the total number of species within each community.

Another way to look at the β-diversity is the unifrac analysis. This analysis includes phylogenetic distances in the analysis. This means that the relative relatedness of the different species is considered when analysing the dissimilarity of species between communities.

1.4 Pro-PACT study

The samples that will be analysed in this thesis were originally collected for the Pro-PACT study, a sub study of the Prevention of Allergy Among Children in Trondheim (PACT). The PACT study was an intervention study focusing on specific interventions and their effect on childhood allergy. The interventions included increased consumption of n-3 polyunsaturated fatty acids, reducing tobacco exposure and reducing indoor dampness. The interventions were included in the recommended maternity care lifestyle counselling programme in Trondheim in 2002/2003. The Pro-PACT study found that administration of probiotic treatment to pregnant mothers resulted in 40% fewer cases of atopic dermatitis in children.

The Pro-PACT study were initiated during the intervention period of the PACT study, and the participating pregnant women were given either a probiotics or placebo treatment from the last trimester of the pregnancy and throughout the 3 first months after giving birth. All the children were breastfed during the first three months of life. A total of 278 mother/child-pairs were randomly selected into two groups, probiotics and a control group. The probiotic group consisted of 138 mother/child-pairs, and the placebo group consisted of 140 pairs.

The women in the probiotics group were given low-fat fermented milk (Biola®) containing the bacterial strands *Lactobacillus rhamnosus GG* (LGG), *Bifidobacterium animalis subsp. Lactis Bb-12* (Bb-12) and *Lactobacillus acidophilus La-5* (La-5). The placebo group were given skimmed fermented milk that were heat-treated to ensure that there were no probiotic bacteria present (Dotterud et al., 2010).

1.5 Aim of study

The development of the oral microbiota is important as this is the gateway to both the gut and, to some extent, the respiratory system. Maternal influence on the microbial communities of children is an interesting subject and should be investigated more closely.

The aim of this thesis was to **determine if probiotic intervention can alter the oral microbiota in mothers and their children**.

The sub goals while working on reaching the main aim is:

- Sequencing the 16S rRNA gene for the oral microbiota.
- Establish if there is an effect of probiotic treatment on the composition and development of the oral microbiota in children during the first 3 months of life.

Earlier Pro-PACT studies has shown that probiotic bacteria most is transferred from mother to child, although the transmission itself is not stated (Dotterud et al., 2015; Simpson et al., 2018).

2 Materials and methods

2.1 Study design and experimental setup

The samples used in this thesis were originally sampled for the Pro-PACT study. The saliva samples were collected from the mother-child pairs in two intervals after birth. The first interval was 10 days after birth, and the other were when the children had reached 3 months of age.

A total of 878 samples from the original study were used; 374 from the probiotic group and 504 from the placebo group. The reason for this somewhat skew distribution is that the samples were both randomized and blinded when they were collected for this thesis. 494 samples from the children were analysed together with 384 samples from the mothers, as table 1 shows.

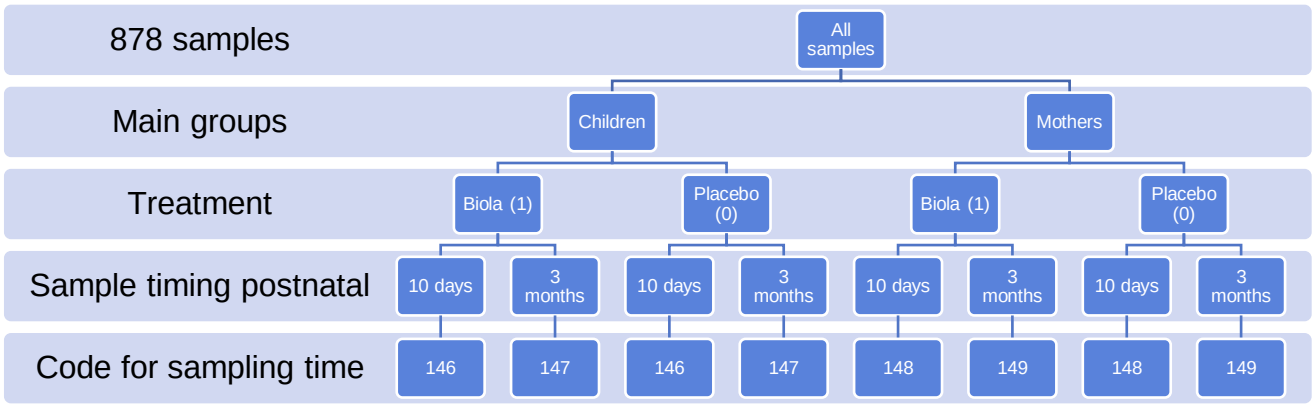

Figure 3: A schematic overview of the samples. The different treatments and sampling time with codes are illustrated here.

Table 1: A summary of the distribution of the samples analysed in this thesis. All the samples used were collected from the Pro-PACT study. A total of 494 samples where from the children, and 384 were from the mothers. The table shows the distribution of the samples within each main group (i.e. mothers and children), separated by the two different treatments and the timing of the sampling.

	Biola	**Placebo**
Children 10 days old	84	85
Children 3 months old	102	223
Mothers 10 days postportum	94	96
Mothers 3 months postpartum	94	100
		N=878

The saliva samples were prepared for next generation sequencing through extraction, three different types of PCR and verification steps in between these main steps. The last PCR and the sequencing itself were performed by laboratory engineer Inga Leena Angell at NMBU.

The workflow is illustrated in figure 4.

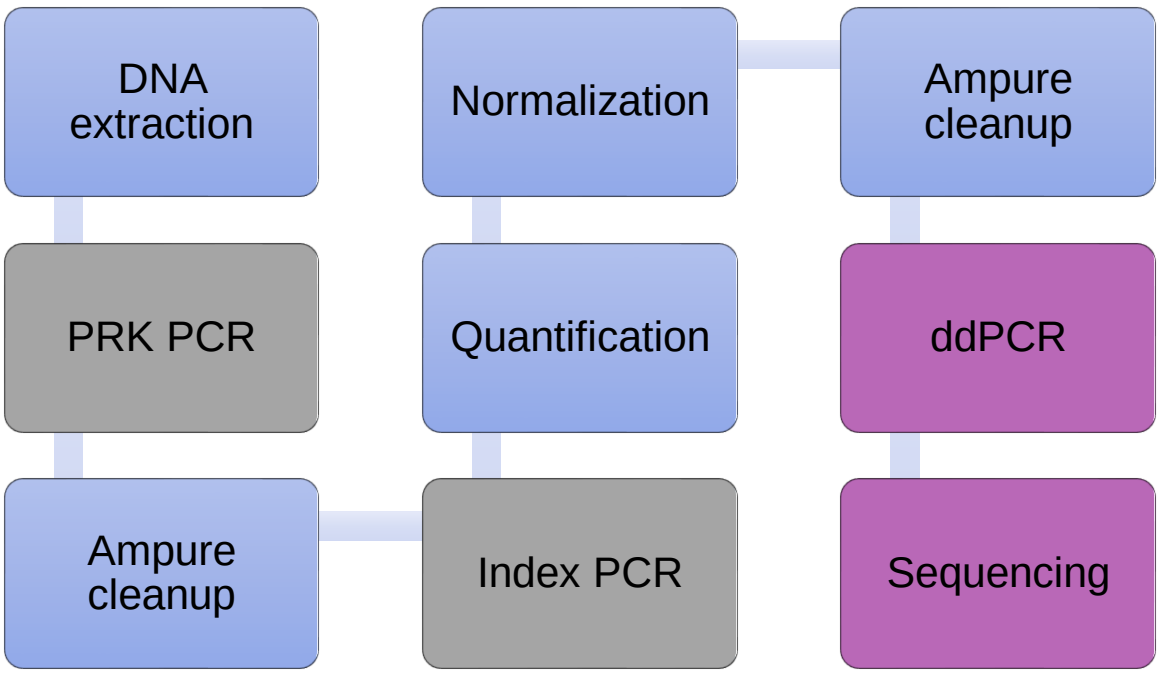

Figure 4: A schematic view of the workflow performed on the 878 samples from the Pro-PACT study. Grey boxes indicate the steps where the samples were verified by gel electrophoresis and Qubit measurements. The purple boxes represent the steps performed by laboratory engineer Inga Leena Angell at NMBU.

2.2 DNA extraction

The saliva samples were thawed on ice before the extraction process was initiated.
Stool transport and recovery buffer (S.T.A.R., Roche Diagnostics Corporation) were used to facilitate the extraction of the nucleic acid from the samples. This buffer has two important features: it minimizes the degradation of nucleic acids and binds nucleic acids to magnetic beads.

Both a mechanical and a chemical approach were used to isolate the DNA from the cells. The mechanical step involved bead-beating to breach the cell walls. The sample tubes for this step contained 200µl S.T.A.R. buffer and acid-washed glass beads in three different sizes (<106µm, 425-600nm and 2,5-3,5mm). Thawed cotton swabs with saliva-samples were put whole in the tubes before being processed for 2x40 sec at 1800rpm in FastPrep 96 with 5 minutes breaks between runs. The samples were centrifuged at 13000rpm for 10 minutes before 50µl of supernatant were taken out from each sample for chemical extraction.

The chemical isolation of the prokaryotic genomic DNA (gDNA) was achieved with the help of Thermo Scientific KingFisher Flex robot and with a MagMidi LGC kit. The samples were transferred to KingFisher 96 well plates with 50 µl lysis buffer and 5 µl proteinase in each well. The samples were incubated at 55°C for 10 minutes to ensure elimination of proteins in the solution before the isolation process was started.

The main feature of the chemical isolation is the use of paramagnetic mag particles. These beads bind to the negatively charged DNA through the formation of salt bridges. This feature makes it possible to clean the samples of all other contaminants, such as protein debris. A magnet in the robot keeps the bead-DNA complexes in one place, easing the washing process. The final step in this process is the addition of water to the samples. Water breaks the salt bridge between the beads and the DNA, releasing the gDNA into the solution while keeping the beads in place on the magnet.

2.3 Polymerase chain reactions

Both qualitative and quantitative PCR were used in this thesis. Qualitative PCR were used to ensure a big enough amount of target amplicons for further analysis. The quantitative PCR were used to calculate the amount of the product. PCR plates with 96 wells and a PCR plate centrifuge were used. All reagents were kept on ice during set-up, and the products were stored at -20°C after finishing each PCR procedure.

2.3.1 Prokaryote PCR (PRK PCR)

A PRK PCR was used to amplify the gDNA from the isolation-step. The two primers PRK341F and PRK806R were used. These primers target the V3-V4 area of the 16S rRNA gene. Each sample-reaction consisted of 5µl of extracted DNA, 1x HOT FIREPol® DNA polymerase, 0,2µM of both the forward primer (PRK341F) and the reverse primer (PRK806R). The sequence of the primers is shown in appendix A.
PCR-grade water was added until the sample had a final volume of 25µl. The thermal conditions of a run were: 95°C in 15 min (initialization), then 25 cycles of 95°C in 30 sec (denaturation), 55°C in 30 sec (annealing) and 72°C in 45 sec (elongation). The samples were then heated to 72°C in 7 minutes (final heating to ensure elongation) before cooling down to 4°C.
The final length of each amplicon was 466 bp.

All the samples were verified after the PRK PCR using Gel electrophoresis (2.5.1) and Qubit measurements (2.5.2).

2.3.2 Index PCR

Illumina sequencing was used to analyse the bacterial composition in each sample. A second PCR-reaction was initiated to add illumina-specific adapters to the fragments from step 2.3.1. Each sample had a different combination of forward and reverse primers. A total of 576 unique combinations were made by using 16 different forward primers and 36 different reverse primers.

The sample-reaction consisted of 1µl template DNA, 5xFIREPol® Master mix ready to load and 0,2µM of both forward and reverse primer. PCR-grade water was added until the sample had a final volume of 25µl.
The thermal conditions of each run were 95°C in 5 min (initialization), then 10 cycles of 95°C in 30sec (denaturation), 55°C in 1 minute (annealing) and 72°C in 45 sec (elongation). The samples were heated to 72°C in 7 minutes to ensure elongation at the end of the run, and finally cooled down to 4°C.
The final length of the amplicons was 594 bp.
The samples were finally tested on a 1% agarose gel, measured with Cambrex – FLX 800 CSE and cleaned with ampure beads.

2.3.3 Droplet digital PCR – quantitative PCR

The amount of DNA present in the pooled library were quantified using droplet digital PCR (ddPCR). A dilution series was used for the quantification. The reaction master mix consisted of 1x Super mix for EvaGreen, 0,2µM Illumina colony forward primer, 0,2 µM Illumina colony reverse primer and PCR-grade water.

BioRad QX200TM – Droplet generator were used to generate the droplets, following BioRad's instructions. The generated droplets were transferred to a PCR plate and sealed using BioRad PCR plate sealer.

The thermal conditions of each run were 95°C in 5 min, then 40 cycles of 95°C in 30sec (denaturation), 60°C for 30 seconds (annealing) and 72°C in 45 sec (elongation). The samples were cooled to 4°C for 5 minutes before being heated again to 90°C to deactivate the enzymes in the sample.

2.4 Ampure purification of PCR products

It is important to ensure that the samples are properly purified before sequencing can be performed. Contaminants left in the solution can interfere with both quantification and sequencing, resulting in inaccurate results. Samples with contaminants might also inhibit PCR-reactions later on in the process. Contaminants are often by-products of the PCR reaction itself, such as primer dimers and nucleotides and primers that have not been incorporated with the amplicons. The purification of the PCR products in this thesis was performed by using 0.1% Sera Mag Beads solution together with the Biomek 3000 robot.

10µl of sample were purified by using the automated protocol on a Biomek 3000. The samples were in a 96 well plate specifically made for the robot. The procedure started with the machine adding AMPure® XP beads to the samples in a 1:1 ratio. This ratio would exclude DNA fragments of length shorter than 200 bp. The solutions were incubated in 5 minutes in room temperature before the plate were put on a magnet.

The AMPure beads works in the same way as the beads used for the isolation step, they too are paramagnetic. So just as in step 2.2, the magnet keeps the bead-DNA complex in the wells while the supernatant is removed.
The supernatant was removed after the samples had been on the magnet for 2 minutes. The protocol then moves on to perform two wash steps were 100µl of 80% alcohol are added to each sample well. The sample is left to incubate for 30 seconds before the alcohol is removed again.
The sample is then left to air dry for 30 minutes, an additional step which ensures that all the alcohol is removed before moving on to the last step in the purification process.

Eluation is accomplished by adding 20µl of PCR-grade water to each sample and taking the plate off the magnet. As mentioned in step 2.2, this breaks the salt bridge between beads and DNA, releasing the DNA into the supernatant. The samples are then incubated for 2 minutes before being moved back to the magnet for another 5 minutes. Finally, the supernatant now containing the purified DNA can be extracted for further analysis.

2.5 Quality and quantity control of PCR products

Both quality and quantity of the PCR products were analysed after prokaryote PCR and index PCR. Only a small subset of the samples on each 96-well plate were analysed. A subset of these verifications can be seen in appendix B.

2.5.1 Gel electrophoresis

The PCR products were verified on agarose gels to ensure the presence of target DNA. The 1% agarose gels were made with 1x TAE running buffer and PeqGreen. 5µl of each sample to be tested was transferred to individual wells in the gel, together with a size marker 100bp DNA ladder (Solis BioDyne, Estonia). Both a positive and a negative control was tested together with the samples. The voltage was set to 80V and the gels run for 30 minutes. The fragments were visualized by using the Molecular Imager® Gel Doc™ XR Imaging (Bio-Rad laboratories, USA).

2.5.2 Qubit measurements

The DNA concentrations for the PCR products were measured with the help of a Qubit™ fluorometer using the Quant-iT™ assay. Working solutions were made according to the manufacturers recommendation and consisted of Quant-iT™ buffer and Quant-iT™ reagent in a 1:200 ratio.

The calibration of the fluorometer was performed by using two standards. The standard solution consisted of 10 µl of standard mixed with working solution to a total of 200 µl. For measurement of target DNA, 2 µl of each sample-DNA were mixed with working solution to a total volume of 200 µl.

The amount of target DNA present in a sample is measured by analysing the intensity of the fluorescent light that the sample emits.

The quantification of the cleaned index PCR products was measured on the Cambrex – FLX 800 CSE. The working solution were made by diluting 1:200 Qubit dsDNA HS reagent in Qubit dsDNA HS Buffer. 2 µl of sample DNA were mixed with working solution to a total volume of 72 µl. Each 96 well plate contained 4 wells with different standards. One well with only working solution and TE buffer, and three wells with 2 µl of standard and 70 µl of working solution.

2.6 Normalization and pooling of index PCR products

The standard curve and its equation from the quantification were used to normalize the DNA in the samples. The normalization is important because it ensures that the pooled samples are in equal concentrations. Equal concentrations are necessary to generate uniform cluster density on the chip. Two libraries were made, with the samples from one plate being present in both libraries as a control on the sequencing.

Both the normalization and the pooling were done using a Biomek 3000 robot, using a cut off value of 12. The pooled libraries were finally purified by hand with Ampure purification before being quantified by ddPCR.

2.7 Sequencing

The pooled libraries were diluted to 4 nM using Tris with pH 8.5 after the quantification. According to the manufacturer's instructions, a PhiX control were made. Both the PhiX control and the pooled libraries were denatured with NaOH before being diluted with a hybridization buffer. The final concentration of the libraries was 6 pM.

The control was added to both libraries so that each library contained 15% PhiX control. The libraries were finally added to MiSeq reagent cartridge for template loading to the MiSeq flow cell.

2.8 Data analysis

The 16S rRNA sequences from the illumina sequencing were processed through the QIIME pipeline. The pipeline puts together the reverse and forward reads and organizes the samples within each library. The next step is to use this information for OTU processing. The OTU's were clustered using a 97% similarity threshold. The SILVA database was then used to assign taxonomy to the OTU's.

The data was then processed in the last pipeline. This pipeline generates taxonomic tables and diversity analyses. Both alpha and beta diversity were calculated for the samples, and diversity indexes were produced.

For further statistical testing, the data had to be checked for normality with the Shapiro-Wilk normality test in R Studio. The Shapiro-Wilk test estimates the variance of a sample in two ways, and compares the two estimated values. These two values will be relatively equal if the sample is normally distributed, and the test will in such a case result in a p-value > 0.05.

Any significant difference between the sub-groups were investigated with the Friedman rank sum test. This is a non-parametric statistical test used to check for differences between groups when the data is not normally distributed.

3 Results

A total of 859 samples were sequenced, this included controls and duplicates. The reason for this number being lower than the initial number of samples (878), is that some of the samples were removed due to their DNA concentration being too low. The DNA concentration was measured in step 2.5.2 by the Cambrex -FLX 800 CSE, and samples were dropped from further analysis if the amount of DNA were too small based on the standard curve, a total of 133 samples.

The sequencing generated 22 681 212 sequences in total. The mean value was 26404 sequences per sample, ranging from 10 to 276127. The number of OTUs generated were 1295.

The samples were sequenced by Illumina Miseq in two batches.

3.1 Diversity in the samples

3.1.1 α-diversity

Qiime was used to calculate the α-diversity of the samples. The numbers generated through this calculation were used to produce rarefraction plots shown in figure 5, 6 and 7. The plots in these figures were made by investigating the total number of sequences and plotting it against the total number of observed species.

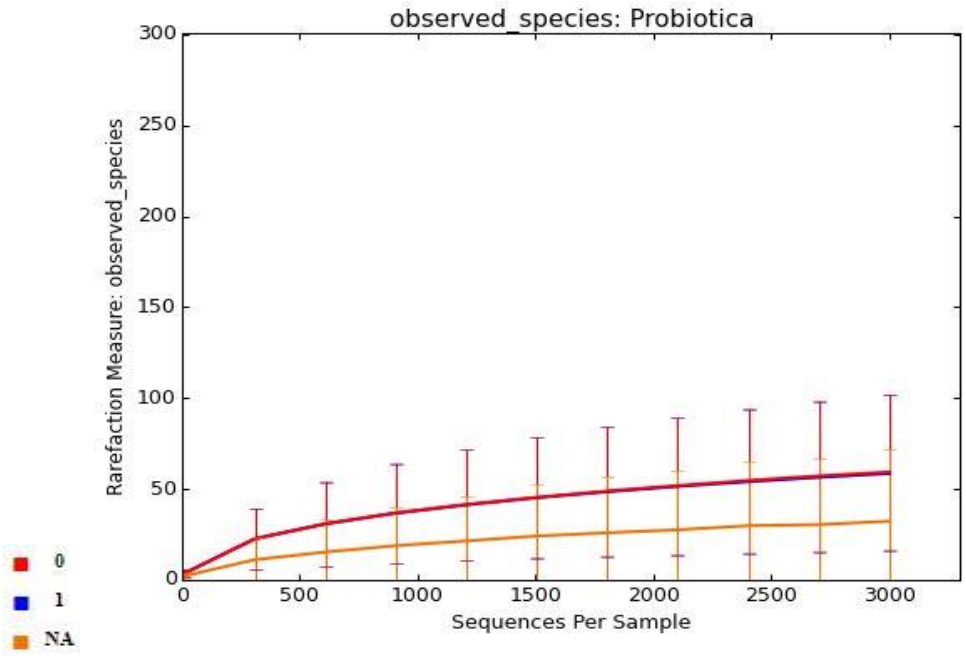

Figure 5: This rarefraction curve shows observed species in response to sequences per sample between the probiotic group (1) and the placebo group (0). NA is the positive and negative control-samples.

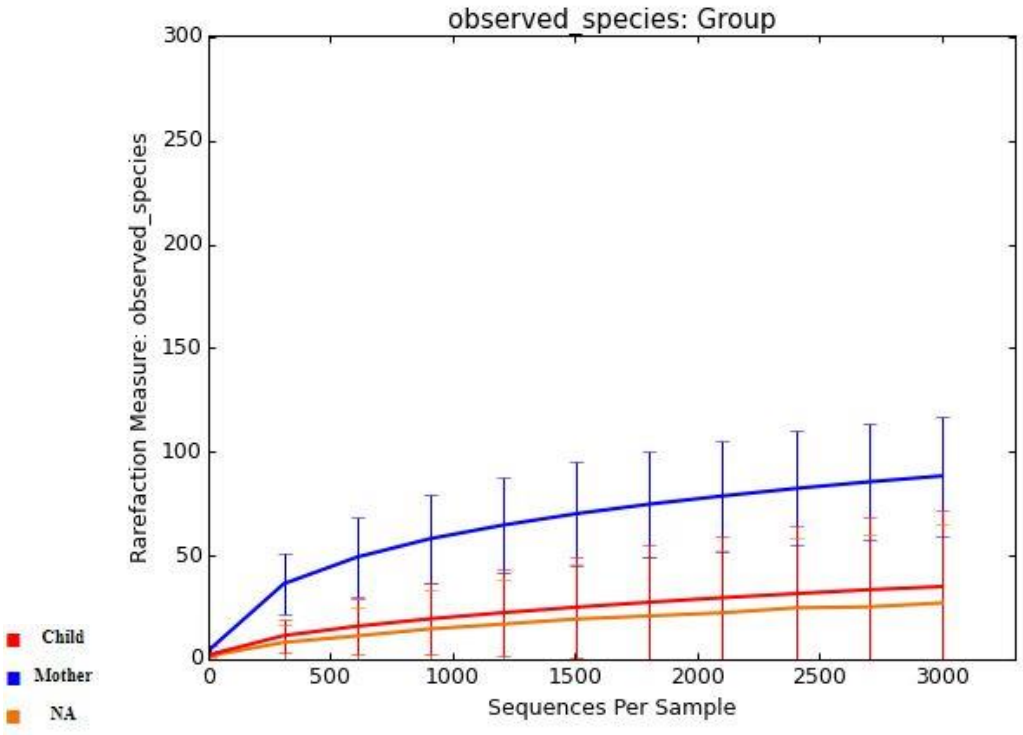

Figure 6: This rarefraction curve shows observed species in response to sequences per sample between the mothers and their children. The negative and positive control samples are here marked as NA.

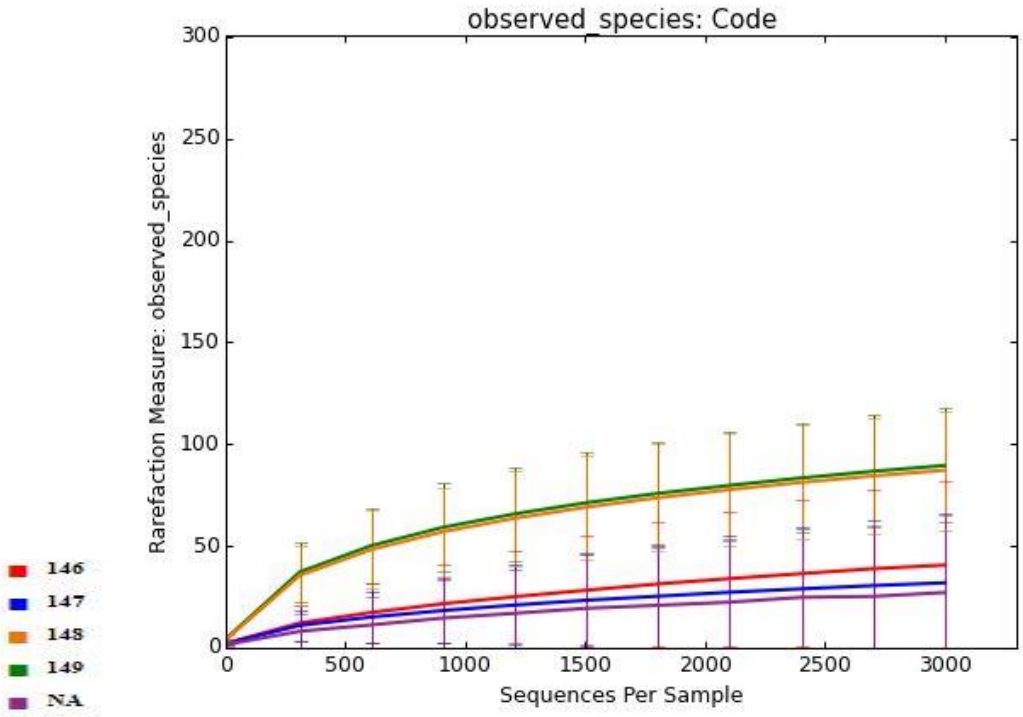

Figure 7: This rarefraction curve shows observed species in response to sequences per sample with regards to both sample timing and mothers/children. Codes 146 and 147 is samples from children at 10 days and 3 months of age. The corresponding samples from mothers has the codes 148 (10 days postpartum) and 149 (3 months postpartum). NA is the positive and negative control-samples.

Probiotic consumption does not seem to influence species richness based on the rarefraction curve in figure 5. The curves representing the two treatments looks identical, indicating that there is no difference between them.

The rarefraction plot based on the two main groups, mothers and children, reveals that the species richness differs between these two groups. As seen in figure 6, the mothers participating in the trial has a higher species richness than the children in general. This is more specified in figure 7.

Figure 7 displays the differences in species richness between the sub-groups with regards to sample-timing. The blue line represents the children at the age of 3 months, and the red line is the samples from the children 10 days after birth. The fact that the blue line is below the red line indicates that the bacterial diversity has decreased from the children were 10 months to they reach 3 months of age.

3.1.2 β-diversity

The β-diversity, or the diversity between the different sub-groups were investigated through principal coordinate analysis plots (PCoA). The data used for the beta diversity came from the Qiime Bray- Curtis dissimilarity analysis and the weighted unifrac analysis. The emperor package was used to create the PCoA-plots.

The PCoA's in figure 8 show no specific clustering an any of the group. The beta-diversity seems to be the same in both the group with probiotic consumption and the placebo group. This indicates that the consumption of probiotics by the mothers might not influence neither mothers nor children's oral microbiota.

The PCoA for the weighted unifrac distance explains more of the variance in the model than the PCoA for the Bray-Curtis dissimilarity (74,23% vs 52,72). This indicates that the relative relatedness of the species shared between samples are important to consider while analysing the beta-diversity.

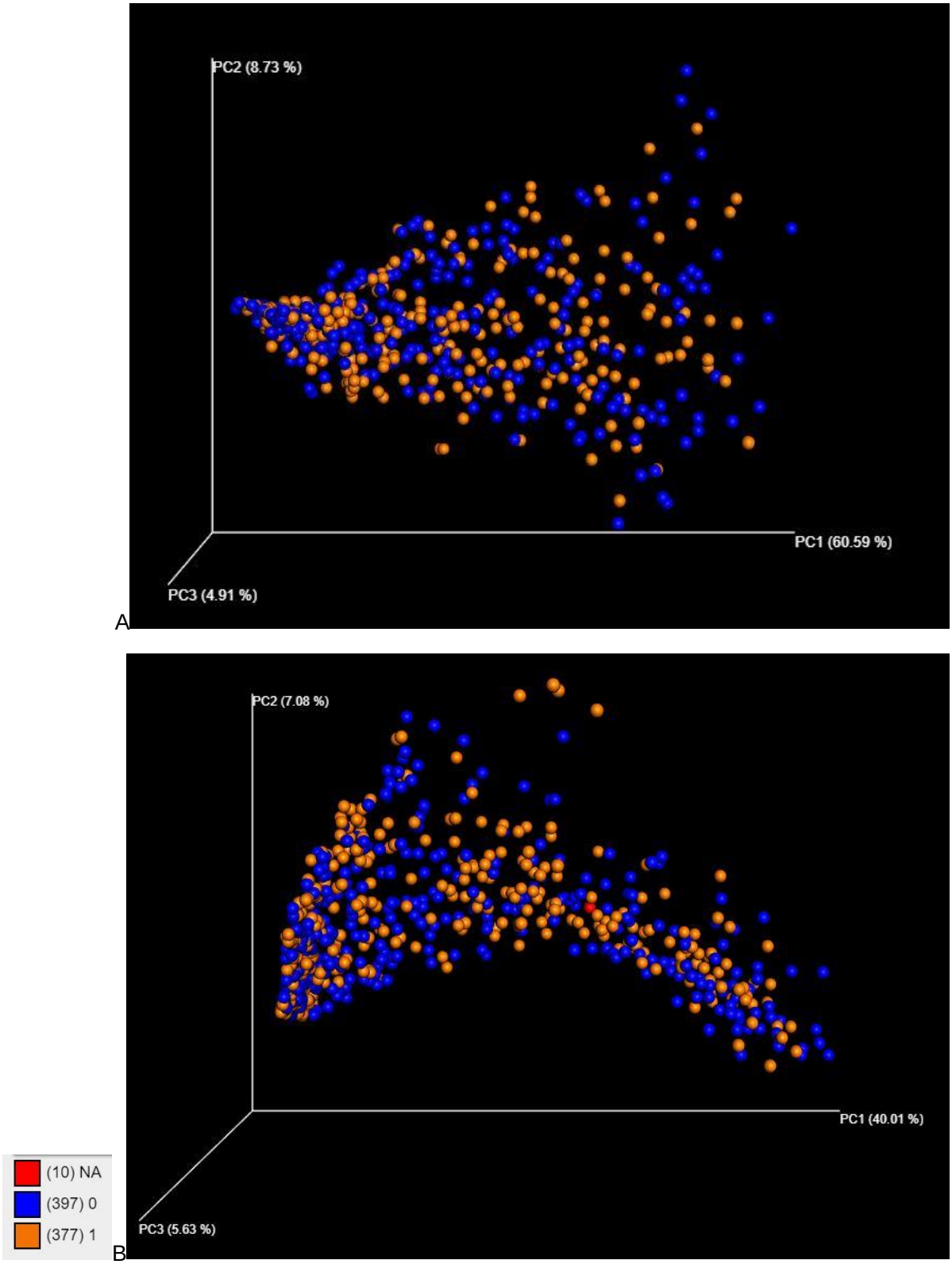

Figure 8: The PCoA for the weighted unifrac distance (8A) and for the Bray-Curtis Dissimilarity (8B) for the placebo and the probiotics group. The orange dots represent group 1, the group receiving biola. 0 is the placebo group, and NA are the negative and positive control samples. The numbers in the parenthesis are the number of samples in each group.

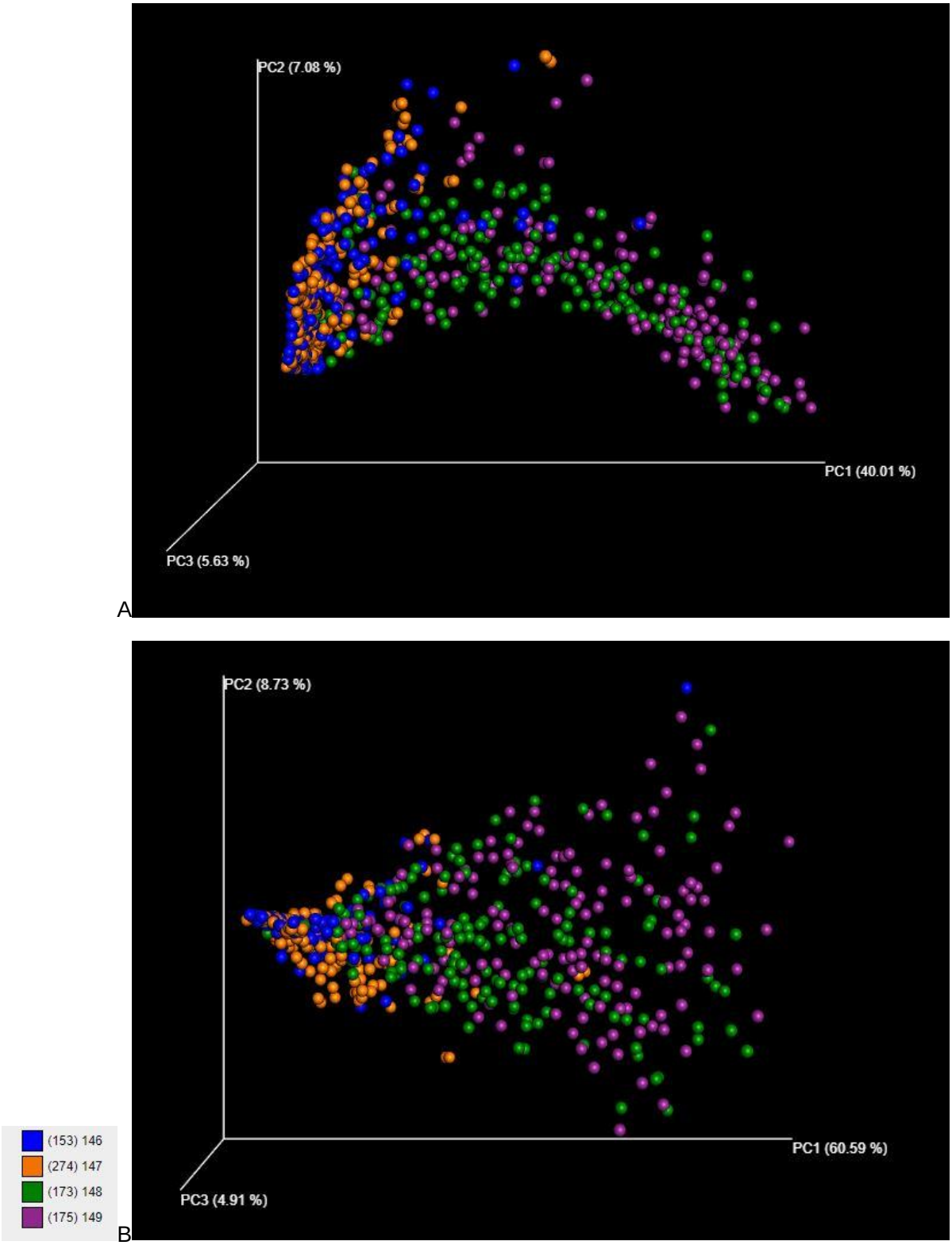

Figure 9: The PCoA for the beta-diversity for the mother- and children-samples. Blue dots represent code 146, samples from children 10 days after birth. Orange dots is from children aged 3 months, group 147. 148 is from mothers 10 days after giving birth, these dots are green in colour. The purple dots represent code 149, mothers 3 months after giving birth. The Bray-Curtis Dissimilarity is shown in 9A, while the weighted unifrac is shown in 9B. The numbers in the parenthesis is the number of samples in each group.

The beta-diversity indexes based on the sampling-time of the sub-groups shows apparent clustering of some of the sample-groups. Although the groups have some level of clustering with each other, there are also evident clustering that distinguishes the children from the mothers.

The PCoA for the Bray-Curtis dissimilarity shows clustering of the mother samples. The mother-samples collected 3 months postpartum indicates to be a bit more similar to the children samples than the samples collected 10 days postpartum. The beta-diversity in the children's oral microbiota has similarities within the main group. But the samples collected 10 days after birth are also more related the mother's beta-diversity than the samples collected 3 months after birth.

The PCoA for the weighted unifrac distance shows a denser clustering of the children-samples while the mother-samples have a lighter density. This indicates that the children have more similar beta-diversities with the other children, and that the diversity within the mother main-group are more divergent. The density of the clustering makes it hard to see whether or not any of the mother-sample are more similar to the children-samples, and the same goes for the children samples. But, as seen in figure 11, there are some outliers from both the children groups that are more similar to the mothers than they are to the other children. The PCoA for the weighted unifrac distance is better at explaining the variance in the model in this analysis as well.

3.1.3 Diversity at genus level

The data generated by Qiime was processed in excel to identify the 10 most dominant OTU's in the community as a whole. Then the dataset was divided into sub-groups with regards to treatment (biola/placebo) and sample timing (10 days/ 3 months after birth). Bar charts illustrating the percentage of the different genera within each sub-group were made.

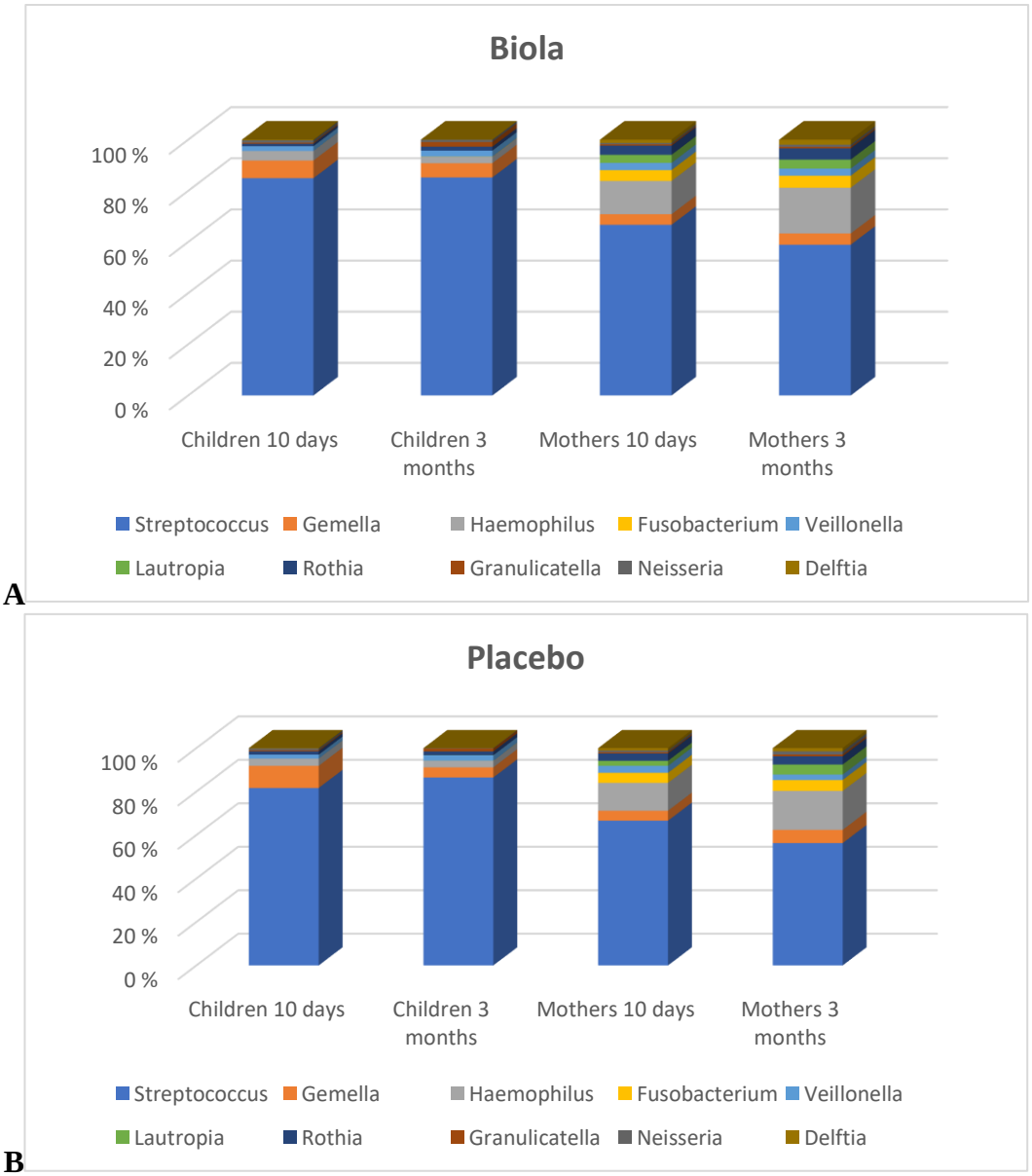

Figure 10: Bar charts illustrating the distribution of the 10 most dominant taxa within each sub-group by percentage. Figure 2A and 2B shows the distribution within the biola/placebo-groups, as well as the distribution within the children/mothers-groups. The different colours in the bars represent one genus each and are named beneath the bars.

The bar charts based on the sub-groups show no major differences between the children samples collected at 10 days of age and the samples at 3 months of age. Figure 10 shows that the most abundant bacterial genera in all the children are *Streptococcus*, followed by *Gemella* and *Haemophilus*. The quantity of *Gemella* is almost doubled in the placebo-group as opposed to the biola-group in the samples collected at 10 days of age (1935 fragments vs 1085).

The biggest difference in the 3-month-old children is the genus *Neisseria*. The children in the biola group has almost 6 times as many fragments of this genus as the children from the mothers administrated placebo treatment.

The bacterial distribution seems to have a slight change in the mothers from 10 days postpartum to 3 months postpartum as seen in figure 10. This change is seen in both the mothers receiving biola, and the mothers receiving placebo treatment. The quantity of the genus *Streptococcus* has decreased while the amount of *Haemophilus* has increased in the mothers during these 3 months. Just as in the children, the genus *Gemella* is slightly more abundant in the mothers who got the placebo treatment than in the mother who received biola treatment (1012 vs 572 fragments in the samples collected 3 months postpartum).

3.2 Friedman rank sum test

R Studio was used to further explore if there were any significant differences between the different sub-groups. Two different tests for normality were performed on the dataset in order to decide which statistical test that were appropriate for the analysis. A histogram for each sub-group indicated visually that no group were normally distributed, the plot for some of the sub-groups can be found in appendix E.

A Shapiro test was then performed to verify the visual indications. All the sub-groups got a p-value smaller than 0.05, implying that the distribution of the data is significantly different from normal distribution. The Friedman rank sum test was therefore chosen to analyse the dataset. This is a non-parametric test used as an alternative to one-way ANOVA when the normality assumption is not met.

The Friedman rank sum test was used to further investigate any statistically significant differences between the sub-groups. The test produced p-values based on the 20 most abundant species in the community as a whole. The commands used in the R analysis can be found in the script in Appendix C.

Table 2: The p-values from the Friedman rank sum test. Groups giving a p-value less than 0.05 when compared is significantly different from each other. The significant numbers are marked in blue in the table, and the explanation for the codes can be found in figure 3.

p-values between the different sub-groups							
Groups	1460	1461	1470	1471	1480	1481	1490
1461	0.07364						
1470	0.3711	0.07364					
1471	0.6547	0.6547	1				
1480	0.07364	0.07364	1	0.1797			
1481	0.1797	0.3711	1	0.3711	0.001745		
1490	0.00729	0.00729	0.6547	0.1797	0.07364	0.00729	
1491	0.3711	0.07364	1	0.6547	0.07364	0.6547	0.001745

There are no obvious differences between the children-groups as all the p-values are higher than 0.05. Both the children groups sampled at 10 days postpartum are significantly different from the 3 month-samples from mothers receiving placebo treatment.

The biggest differences can be found within the mother main group. There are significant differences between the mothers receiving biola and those that did not both in the samples collected 10 days postpartum and the samples obtained at 3 months postpartum. There is also a significant difference in the bacterial distribution between the mothers administrated biola at 10 days after birth and the mothers receiving placebo treatment and their samples collected 3 months postpartum.

3.3 The *Lactobacillus* and *Bifidobacterium* genera in the samples

The treatment with biola have the possibility to increase the amount of the three bacterial strands *Lactobacillus rhamnosus GG* (LGG), *Bifidobacterium animalis subsp. Lactis Bb-12* (Bb-12) and *Lactobacillus acidophilus La-5* (La-5) in the oral cavity. To investigate if the treatment could change the abundance of these bacteria in the oral cavity, the different OTU representing the genera *Lactobacillus* and *Bifidobacterium* were identified in the Qiime dataset and processed in excel. Only *Bifidobacterium animalis* were identified on species level. The data was processed with regards to the number of sequenced fragments in each sub-group.

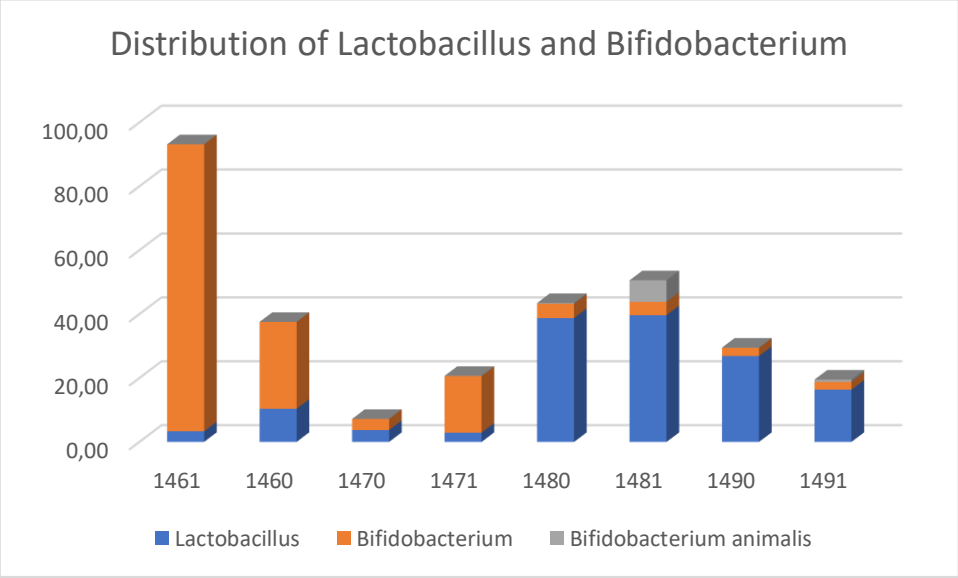

Figure 11: The relative abundance of the different bacteria from the genera Lactobacillus and Bifidobacterium in the different samples. The abundance is measured in sequenced fragments within each sub-group. The different colours represent the different genera and are described beneath the bars. Explanation for the sample codes can be found in figure 3.

The abundance of the different bacteria belonging to the genera *Lactobacillus* and *Bifidobacterium* differ a lot between the different sub-groups, as shown in figure 11. The most noticeable difference is in the abundance of the two genera in the two main groups, mother and children. The mothers have more of the genus Lactobacillus than the genus Bifidobacterium, while the opposite distribution is found in their children.

The abundance of these two genera is biggest in the children at 10 days of age from mothers receiving biola. The smallest abundance was found in the saliva of children at three months of

age and whose mothers receiving placebo treatment. The abundance of the probiotic bacteria seems to vary less between the different mothers than between their children. The only resemblance to be found between the samples is that the relative abundance of these two genera decreases with time.

3.3 The control-samples

A total of 94 samples were sequenced on both runs as a control. *Streptococcus* was the most abundant genus, and an analysis on this genus with regards to the duplicate samples were performed in excel. The number of sequenced fragments of *Streptococcus* were looked at in 10 of the samples that were sequenced in both runs.

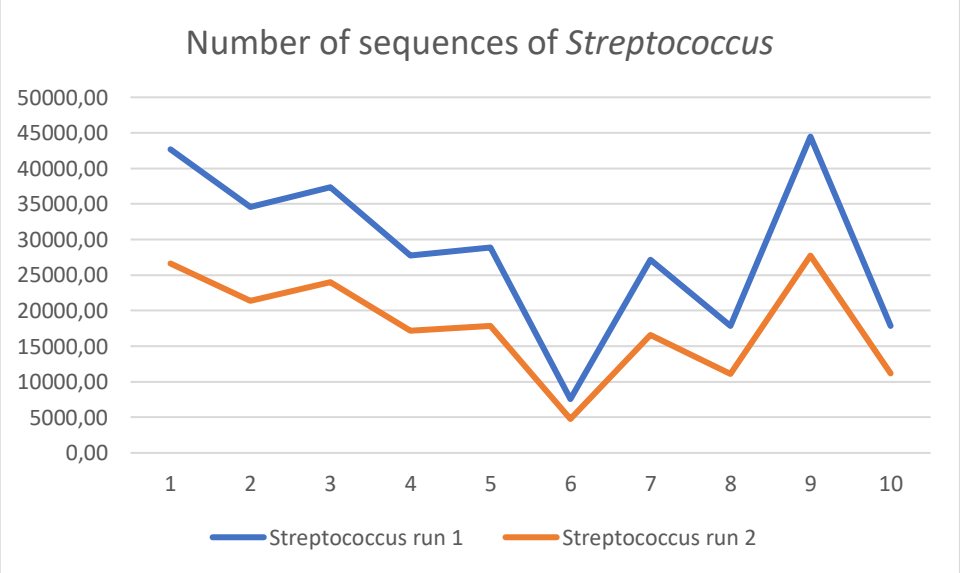

Figure 12: A graph illustrating the number of sequenced fragments of Streptococcus in 10 samples. The 10 samples were sequenced in two different runs.

The blue graph illustrating run 1 indicates that this run yielded more sequenced fragments than run 2 did. This seems to be consistent in all the 10 duplicate samples. The difference in number of fragments sequenced varies between the samples. The biggest difference is found in sample 9, a total of 16731 fragments in difference. Only 2796 fragments differ between the two runs on sample 6 as illustrated in figure 12.

4 Discussion

The most solid finding during this study is that the probiotic treatment of the mothers does seem to change their oral microbiota to some extent. No major probiotic influence was found in the children's microbiota, but smaller changes were indicated. No previous studies have investigated the maternal consumption of probiotics and its effect on their children during pregnancy and in the first months after birth. And no earlier studies could be found that investigated the long-term influence of probiotic treatment on the oral microbiota.

4.1 Probiotics and its effect on oral microbiota

The maternal probiotic treatments did not seem to have any major effect on the composition of the oral microbiota in the children. The bar charts in figure 10 reveals that the only distinction is in the genus *Gemella* in the 10 days old children. The children in the biola group has a mean of 1935 fragments, near double the amount of the children in the placebo group who has a mean of 1085 fragments. These findings are also reflected in the mothers, the women in the placebo group have more fragments of the *Gemella* genus than the women treated with probiotics. This might indicate that the probiotic treatment has an influence on the abundance of this species of bacteria, and that this influence seems to be transferred to their children.

The biggest difference in the 3-month-old children is the genus *Neisseria*. This genus is in a 6:1 ratio between the biola- and the placebo group. These differences cannot be found in the mothers, so the correlation to the biola treatment is not certain. Further research is needed in order to verify if these differences is a result of the treatment with biola, or just a part of natural variation. Previous studies have stated that *Neisseria* is one of the most frequent genera in children at 5 months of age (Dzidic et al., 2018). A possible explanation is that the probiotic consumption by their mothers might influence their children's oral microbiota to mature faster than it normally would.

The distribution of *Lactobacillus* and *Bifidobacterium* is quite interesting. It shows that children in general have more bacteria from the genus *Bifidobacterium* than from *Lactobacillus* in their oral cavity, while the opposite is found in the mothers. It does also seem like the children in the biola group have more fragments from these two genera than the children in the placebo group. That might suggest that the bacteria from the biola treatment does transfer to the children at some point. There are no apparent differences between the mothers with regards to treatment, but there seem to be a decrease in these two genera from 10 days postpartum to 3 months postpartum.

The distribution of *Lactobacillus* and *Bifidobacterium* in figure 11 is quite interesting. It indicates that children in general have more bacteria from the genus *Bifidobacterium* than from *Lactobacillus* in their oral cavity, while the opposite is found in the mothers. It does also seem like the children in the biola group has more fragments from these two genera than the children in the placebo group. That might suggest that the bacteria from the biola treatment does transfer to the children at some point. There are no major differences between the mothers with regards to treatment, but there seem to be a decrease in these two genera from 10 days postpartum to 3 months postpartum. Mothers that consumed biola also had more of *Bifidobacterium animalis* in their oral microbiota 10 days after giving birth compared to the mothers that received placebo treatment.

All the children participating in this study were breastfed, and the microbiota of the human milk contains *Bifidobacterium* (Fitzstevens et al., 2017). A possible explanation for the abundance of this genus in the children's oral microbiota compared to their mothers, is that they receive these bacteria through being breastfed.

The Friedman rank sum test revealed significant differences between mothers who received probiotic treatment and those that got the placebo treatment. This indicates that probiotic treatment does indeed alter the microbial composition in the oral microbiota of the treatment receivers. This did not show up in the β-diversity plots as they were not specified towards both the different treatment and the sub-groups in the same plot.

4.2 Diversity of oral microbiota in connection to age

Diversity of the oral microbial communities in children was in this thesis found to be lower than the diversity in their mothers. This was an expected result as other studies have shown an increase in the complexity of the oral microbiota over time (Tanner et al., 2002). The oral microbiota is in constant development and the maternal influence from birth decreases with age.

A rather surprising finding in this thesis is the decrease in diversity in children with age, as seen in figure 7. The amount of observed species in the children seems to drop between the ages of 10 days and 3 months, quite the opposite of the expected result. A theory for this result can be found by looking at a normal baby's environmental influences and what we know about the development of the oral microbiome. The baby's oral microbiome is highly influenced by the mother's microbiome after the birth. Many of the bacterial species present at this time is not able to colonize the baby at this point, but might be present in the oral cavity at 10 days of age (Sampaio-Maia, Benedita & Monteiro-Silva, Filipa, 2014). The bacterial species that are not able to colonize are most likely gone when the baby reaches 3 months of age, explaining the difference in the diversity between the two sampling times.
In addition to this, two of the most important environmental factors have most likely not influenced the child yet; eating solid food and teething. Teething usually does not occur until the child is 6 months old, and solid food is normally introduced at 4-5 months of age. Both these factors can change the oral microbiota drastically (Dzidic et al., 2018; Verma et al., 2018). The diversity might therefore not increase before the baby encounters one of these major environmental factors.

The 10 most abundant genera in the samples investigated in this thesis belongs to the phyla *Firmicutes, Proteobacteria, Fusobacteria* and *Actinobacteria*. This is just as anticipated from findings in studies investigating the composition of a healthy oral microbiota. The 6 different phyla that comprises 96% of the total oral bacteria in humans are *Firmicutes, Bacteriodetes, Proteobacteria, Actinobacteria, Spirochaetes* and *Fusobacteria* (Dewhirst et al., 2010).

The genus *Streptococcus* were in this study found to be the most abundant in all the groups, but it is present in bigger amounts in the children than in the mothers. Species from the *Streptococcus* genus takes part in the initial colonization of the oral cavity. Their bi-products makes the oral cavity habitable for other bacterial species(Sampaio-Maia, B. & Monteiro-Silva, F., 2014). The abundance of *Streptococcus* in young children's oral cavity is therefore quite essential to the development of the oral microbiota. The decrease of *Streptococcus* with age is natural as the oral microbiota evolves and new bacterial species becomes a part of it.

Another important aspect with the genus *Streptococcus* is that many of these species produce immunoglobulin A1 protease. This is a trait which is believed to contribute to the survival of the species in an IgA-rich environment, which the oral cavity is during breast feeding (Cole et al., 1994). This is an additional explanation for the difference in the abundance of this genus between children and their mothers.

The beta-diversity shows no apparent clustering between mothers and children. This also backs up previous findings suggesting that the oral microbiota develops over many years, and does not become similar to adults until after puberty (Cephas et al., 2011).

4.3 Technical challenges

The samples used in this thesis were collected using cotton swabs. As different bacteria colonize different areas of the oral cavity, the use of just one swab might make it difficult to get an optimal representation of the whole bacterial community. The type of swab used has an influence on the results as well. Studies have shown that bacterial uptake and release is significantly dependent on the sampling conditions and the type of swab(Warnke et al., 2014).

The difference in the number of sequenced fragments between the two runs (figure 12) is an indication that there are deviations in some of the steps in the workflow. The deviations might be due to problems during amplification of the samples, most likely during the Illumina process.

4.4 Further research

To address the first technical challenge, a study to investigate the different sampling methods for oral microbiota could be performed. There are several different ways to collect such samples, and these should be compared in order to find the method giving the best representation of the oral microbiota. It would be important to consider sampling of babies' oral microbiota in such a study as they are harder to collect the samples from than adults.

The mother's vaginal microbiota was also sampled during the Pro-PACT study. It would be very interesting to compare these results to the oral microbiota of the children. This comparison could possibly show some interesting findings when investigated while looking at the biola- and placebo groups.

It might also be interesting to conduct a new study on these same samples with the specific mother-child pairs in mind. The samples were randomized and blinded when they were picked out for this thesis, so mother-child pairs were not considered.

5 Conclusion

Maternal probiotic consumption does not seem to have any apparent effect on the children's oral microbiota. Although some small differences can be found when comparing children whose mothers received biola to those that did not, further research should be conducted in order to verify these findings. There is no significant difference between the two groups of children at 3 months of age either, suggesting that the probiotic consumption in mothers does not alter the development of the oral microbiota in their children. The initial sampling method used on the samples might not have been able to give an optimal representation of the bacterial community present in the oral cavity.

6 References

Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., et al. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456: 53. doi: 10.1038/nature07517

https://www.nature.com/articles/nature07517#supplementary-information.

Bianconi, E., Piovesan, A., Facchin, F., Beraudi, A., Casadei, R., Frabetti, F., Vitale, L., Pelleri, M. C., Tassani, S., Piva, F., et al. (2013). An estimation of the number of cells in the human body. *Annals of Human Biology*, 40 (6): 463-471. doi: 10.3109/03014460.2013.807878.

Cephas, K. D., Kim, J., Mathai, R. A., Barry, K. A., Dowd, S. E., Meline, B. S. & Swanson, K. S. (2011). Comparative analysis of salivary bacterial microbiome diversity in edentulous infants and their mothers or primary care givers using pyrosequencing. *PloS one*, 6 (8): e23503-e23503. doi: 10.1371/journal.pone.0023503.

Cole, M. F., Evans, M., Fitzsimmons, S., Johnson, J., Pearce, C., Sheridan, M. J., Wientzen, R. & Bowden, G. (1994). Pioneer oral streptococci produce immunoglobulin A1 protease. *Infection and immunity*, 62 (6): 2165-2168.

Dewhirst, F. E., Chen, T., Izard, J., Paster, B. J., Tanner, A. C. R., Yu, W.-H., Lakshmanan, A. & Wade, W. G. (2010). The Human Oral Microbiome. *Journal of Bacteriology*, 192 (19): 5002-5017. doi: 10.1128/jb.00542-10.

Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N. & Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A*, 107 (26): 11971-5. doi: 10.1073/pnas.1002601107.

Dotterud, C. K., Storro, O., Johnsen, R. & Oien, T. (2010). Probiotics in pregnant women to prevent allergic disease: a randomized, double-blind trial. *Br J Dermatol*, 163 (3): 616-23. doi: 10.1111/j.1365-2133.2010.09889.x.

Dotterud, C. K., Avershina, E., Sekelja, M., Simpson, M. R., Rudi, K., Storro, O., Johnsen, R. & Oien, T. (2015). Does Maternal Perinatal Probiotic Supplementation Alter the Intestinal Microbiota of Mother and Child? *J Pediatr Gastroenterol Nutr*, 61 (2): 200-7. doi: 10.1097/mpg.0000000000000781.

Dzidic, M., Collado, M. C., Abrahamsson, T., Artacho, A., Stensson, M., Jenmalm, M. C. & Mira, A. (2018). Oral microbiome development during childhood: an ecological succession influenced by postnatal factors and associated with tooth decay. *The ISME Journal*, 12 (9): 2292-2306. doi: 10.1038/s41396-018-0204-z.

Escapa, I. F., Chen, T., Huang, Y., Gajare, P., Dewhirst, F. E. & Lemon, K. P. (2018). New Insights into Human Nostril Microbiome from the Expanded Human Oral Microbiome Database (eHOMD): a Resource for the Microbiome of the Human Aerodigestive Tract. *mSystems*, 3 (6): e00187-18. doi: 10.1128/mSystems.00187-18.

Fitzstevens, J. L., Smith, K. C., Hagadorn, J. I., Caimano, M. J., Matson, A. P. & Brownell, E. A. (2017). Systematic Review of the Human Milk Microbiota. *Nutrition in Clinical Practice*, 32 (3): 354-364. doi: 10.1177/0884533616670150.

Gomez, A. & Nelson, K. E. (2017). The Oral Microbiome of Children: Development, Disease, and Implications Beyond Oral Health. *Microbial ecology*, 73 (2): 492-503. doi: 10.1007/s00248-016-0854-1.

https://www.researchgate.net/figure/Illumina-sequencing-and-data-processing-workflow-A-Denaturated-NGS-library-fragments_fig2_303984059, N. g. s. t. a. g. d. a. P. f. r. r.-S. F. o. R. A. f.

Janda, J. M. & Abbott, S. L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol*, 45 (9): 2761-4. doi: 10.1128/jcm.01228-07.

McCombie, W. R., McPherson, J. D. & Mardis, E. R. (2018). Next-Generation Sequencing Technologies. *Cold Spring Harbor Perspectives in Medicine*. doi: 10.1101/cshperspect.a036798.

Pinheiro, L. B., Coleman, V. A., Hindson, C. M., Herrmann, J., Hindson, B. J., Bhat, S. & Emslie, K. R. (2012). Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal Chem*, 84 (2): 1003-11. doi: 10.1021/ac202578x.

Sampaio-Maia, B. & Monteiro-Silva, F. (2014). Acquisition and maturation of oral microbiome throughout childhood: An update. *Dent Res J (Isfahan)*, 11 (3): 291-301.

Sampaio-Maia, B. & Monteiro-Silva, F. (2014). Acquisition and maturation of oral microbiome throughout childhood: An update. *Dental research journal*, 11 (3): 291-301.
Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America*, 74 (12): 5463-5467.
Sender, R., Fuchs, S. & Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS biology*, 14 (8): e1002533-e1002533. doi: 10.1371/journal.pbio.1002533.
Simpson, M. R., Avershina, E., Storrø, O., Johnsen, R., Rudi, K. & Øien, T. (2018). Breastfeeding-associated microbiota in human milk following supplementation with Lactobacillus rhamnosus GG, Lactobacillus acidophilus La-5, and Bifidobacterium animalis ssp. lactis Bb-12. *Journal of Dairy Science*, 101 (2): 889-899. doi: 10.3168/jds.2017-13411.
Tanner, A. C., Milgrom, P. M., Kent, R., Jr., Mokeem, S. A., Page, R. C., Liao, S. I., Riedy, C. A. & Bruss, J. B. (2002). Similarity of the oral microbiota of pre-school children with that of their caregivers in a population-based study. *Oral Microbiol Immunol*, 17 (6): 379-87.
Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R. & Gordon, J. I. (2007). The human microbiome project. *Nature*, 449 (7164): 804-810. doi: 10.1038/nature06244.
Verma, D., Garg, P. & Dubey, A. (2018). *Insights into the human oral microbiome*, vol. 200.
Warnke, P., Warning, L. & Podbielski, A. (2014). Some are more equal--a comparative study on swab uptake and release of bacterial suspensions. *PLoS One*, 9 (7): e102215. doi: 10.1371/journal.pone.0102215.
Whittaker, R. H. (1972). Evolution and Measurement of Species Diversity. *Taxon*, 21 (2/3): 213-251. doi: 10.2307/1218190.
Willey, J. M., Sherwood, L. M. & Woolverton, C. J. (2014). *Prescott's Microbiology*. 9 ed.: McGraw-Hill Education.

Appendix

Appendix A: Polymerase chain reactions primers:

PRK PCR Primers (5'-3'):

Forward: CCTACGGGRBGCASCAG

Reverse: GGACTACYVGGGTATCTAAT

Illumina primers:

Forward primers 5'-3':

1. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctagtcaa CCTACGGGRBGCASCAG
2. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctagttcc CCTACGGGRBGCASCAG
3. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctatgtca CCTACGGGRBGCASCAG
4. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctccgtcc CCTACGGGRBGCASCAG
5. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctgtagag CCTACGGGRBGCASCAG
6. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctgtccgc CCTACGGGRBGCASCAG
7. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctgtgaaa CCTACGGGRBGCASCAG
8. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctgtggcc CCTACGGGRBGCASCAG
9. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctgtttcg CCTACGGGRBGCASCAG
10. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctcgtacg CCTACGGGRBGCASCAG
11. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctgagtgg CCTACGGGRBGCASCAG
12. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctggtagc CCTACGGGRBGCASCAG
13. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctactgat CCTACGGGRBGCASCAG
14. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctatgagc CCTACGGGRBGCASCAG
15. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctattcct CCTACGGGRBGCASCAG
16. aatgatacggcgaccaccgagatct acactctttccctacacgacgctcttccgatctcaaaag CCTACGGGRBGCASCAG

Reverse primers 5'-3':

1. caagcagaagacggcatacgagatcgtgat gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

2. caagcagaagacggcatacgagatacatcg gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

3. caagcagaagacggcatacgagatgcctaa gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

4. caagcagaagacggcatacgagattggtca gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

5. caagcagaagacggcatacgagatcactct gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

6. caagcagaagacggcatacgagatattggc gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

7. caagcagaagacggcatacgagatgatctg gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

8. caagcagaagacggcatacgagattcaagt gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

9. caagcagaagacggcatacgagatctgatc gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

10. caagcagaagacggcatacgagataagcta gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

11. caagcagaagacggcatacgagatgtagcc gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

12. caagcagaagacggcatacgagattacaag gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

13. caagcagaagacggcatacgagatttgact gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

14. caagcagaagacggcatacgagatggaact gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

15. caagcagaagacggcatacgagattgacat gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

16. caagcagaagacggcatacgagatggacgg gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

17. caagcagaagacggcatacgagatctctac gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

18. caagcagaagacggcatacgagatgcggac gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

19. caagcagaagacggcatacgagattttcac gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

20. caagcagaagacggcatacgagatggccac gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

21. caagcagaagacggcatacgagatcgaaac gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

22. caagcagaagacggcatacgagatcgtacg gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

23. caagcagaagacggcatacgagatccactc gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

24. caagcagaagacggcatacgagatgctacc gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

25. caagcagaagacggcatacgagatatcagt gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

26. caagcagaagacggcatacgagatgctcat gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

27. caagcagaagacggcatacgagataggaat gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

28. caagcagaagacggcatacgagatcttttg gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

29. caagcagaagacggcatacgagattagttg gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

30. caagcagaagacggcatacgagatccggtg gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

31. caagcagaagacggcatacgagatatcgtg gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

32. caagcagaagacggcatacgagattgagtg gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

33. caagcagaagacggcatacgagatcgcctg gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

34. caagcagaagacggcatacgagatgccatg gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

35. caagcagaagacggcatacgagataaaatg gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

36. caagcagaagacggcatacgagattgttgg gtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

Appendix B: Gel electrophoresis:

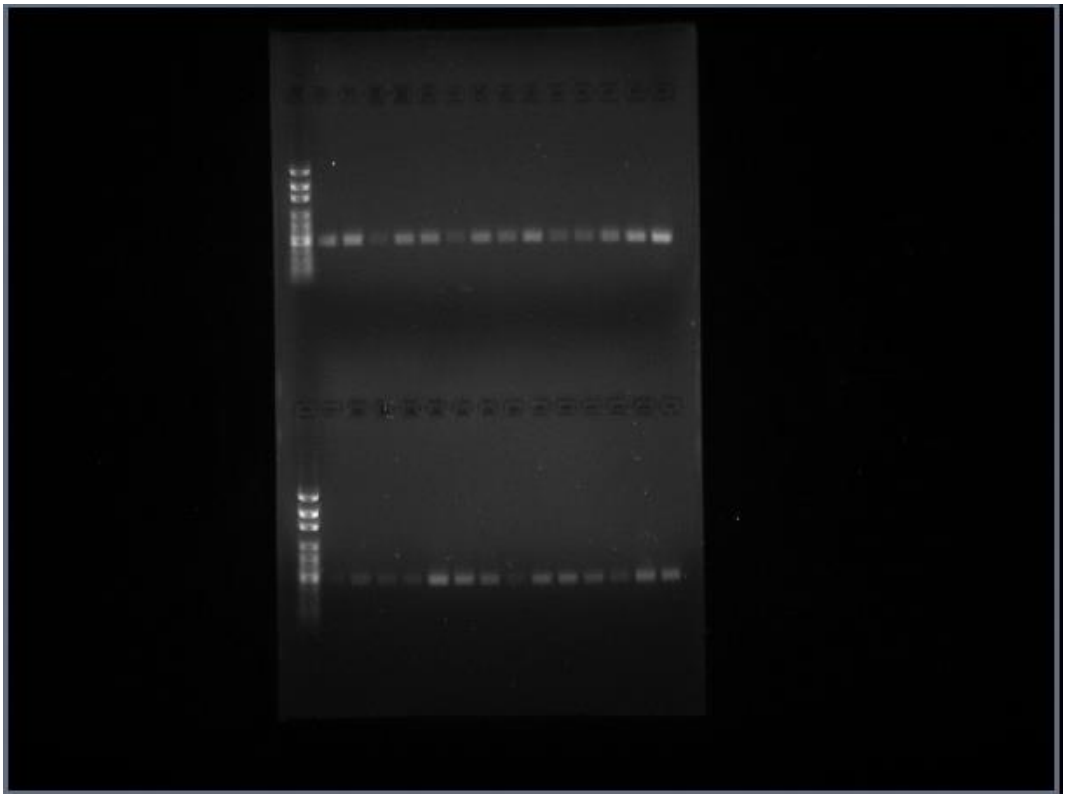

Figure A1: Picture of a gel with samples after a PRK PCR-process.

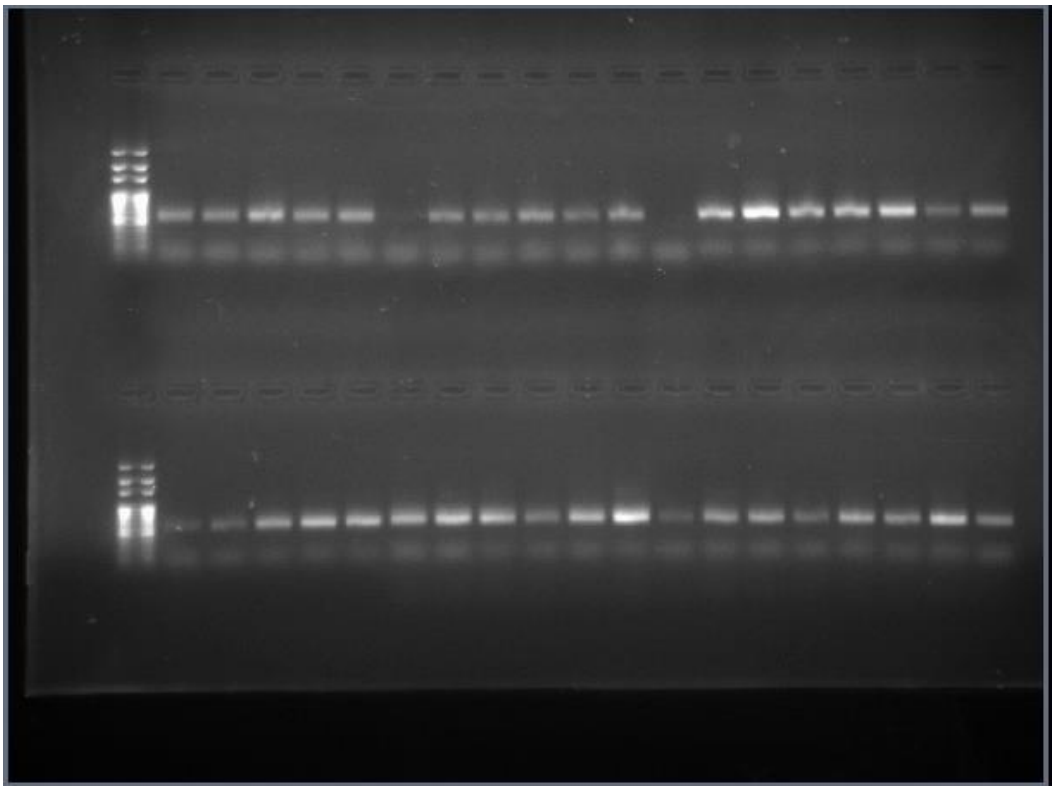

Figure A2: This picture is of a gel with samples after an index PCR.

Appendic C: R Studio-script:

```
#Script:

source("https://bioconductor.org/biocLite.R")
biocLite("phyloseq")
>library(ape)
>library(dplyr)
>library(ggplot2)
>library(gplots)
>library(lme4)
>library(phangorn)
>library(phyloseq)
>library(plotly)
>library(tidyr)
>library(vegan)
>library(VennDiagram)
#Loading the dataset with the 10 most abundandt OTU's:
> library(readxl)
> OTU10 <- read_excel("OTU_table.xlsx")
> View(OTU_table)
#Setting the first row of the dataset as the row name:
>OTU <- data.frame(OTU10, row.names = 1)
#Performing two-sided t-test:
> t.test(OTU$X1461, OTU$X1460, alternative = "two.sided", mu = 0.0, conf.level = .95, paired = TRUE)

Paired t-test

data:  OTU$X1461 and OTU$X1460
t = -1.1546, df = 9, p-value = 0.278
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
  -946.0239  306.6519
sample estimates:
  mean of the differences
-319.686

#Loading the dataset with the 20 most abundandt OTU's:
> library(readxl)
> OTU20 <- read_excel("C:/Users/SirilMaleneIsaksen/OneDrive - Lyfstone AS/Studier/Microbiota/Oral_microbiota/OTU20.xlsx")
#Setting the first row of the dataset as the row name:
>OTU2 <- data.frame(OTU20, row.names = 1)
#Performing two-sided t-test:
> t.test(OTU2$X1460, OTU2$X1461, alternative = "two.sided", mu = 0.0, conf.level = .95, paired = TRUE)
```

Figure A3: R Studio script showing what commands were performed to analyse the data.

```
#Checking the dataframe:
> dim(OTU)
[1] 20  8

#Checking the normal distribution visually:
hist(OTU$X1460, col = "blue")

#Checking the normal distribution with the Shapiro-Wilk test:
shapiro.test(OTU$1460)

Shapiro-Wilk normality test

data:  OTU$X1460
W = 0.30782, p-value = 8.728e-09

shapiro.test(OTU$X1460)

#Installing packages needed to perform Friedman test:
>install.packages("reshape2")
> library(reshape2)
>install.packages("yaml")
> library(yaml)
>install.packages("statsr")
>library(statsr)

#Performing the Friedman rank sum test:
>T1 <- na.omit(with(OTU, cbind(X1460, X1461)))
>cat("\nMedians:\n")

Medians:
  > print(apply(T1, 2, median))
X1460    X1461
77.97647 88.32143
> friedman.test(T1)

Friedman rank sum test

data:  T1
Friedman chi-squared = 3.2, df = 1, p-value = 0.07364
```

Figure A4: R Studio-script showing the commands used to perform both normality tests and the Friedman rank sum test.

Appendix D: Standard curve

This is the standard curve made from the quantification. The equation was used to calculate the amount of DNA present in the samples.

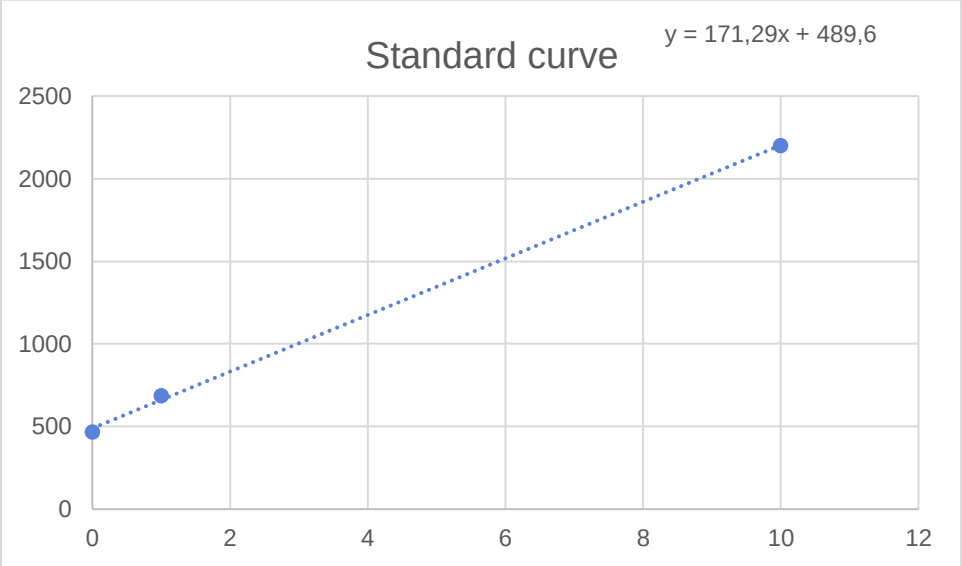

Figure A5: The standard curve made for calculation of the DNA in the samples.

Appendix E: Histogram from normality test

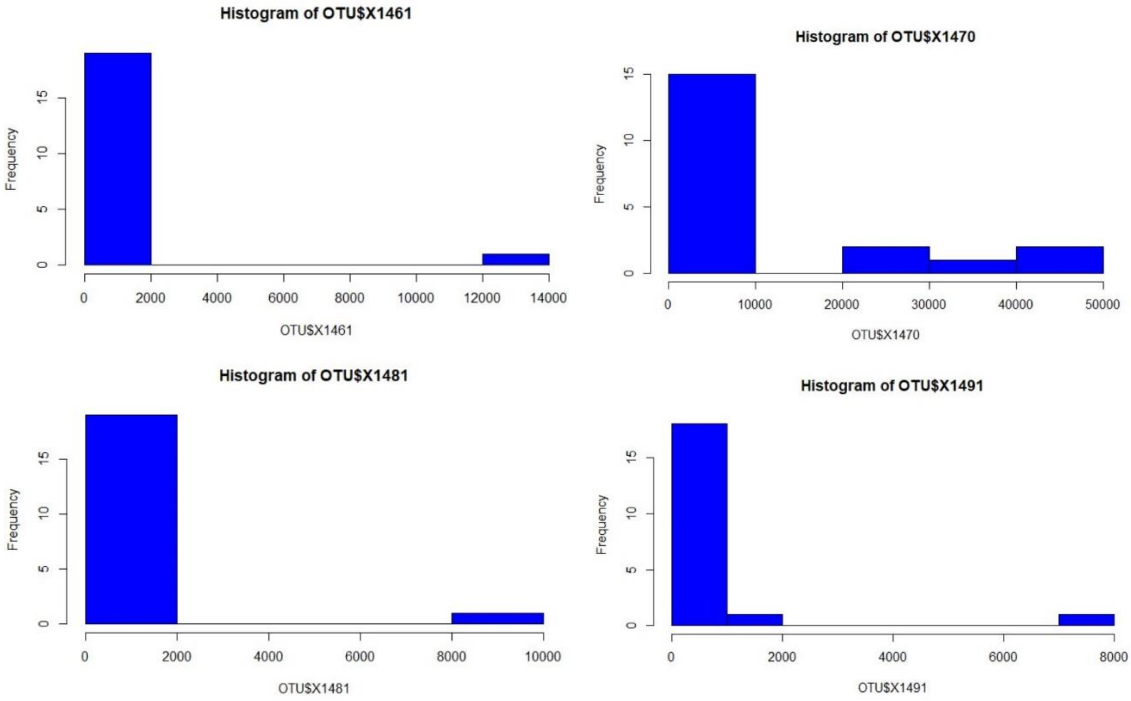

Figure A6: The generated histograms from the normality check for 4 of the sub-groups.

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

Postboks 5003
NO-1432 Ås
Norway