Gut Microbiota and Short Chain Fatty Acids (SCFAs) Composition the First Year of Life

Morten Nilsen

M-Biotechnology
Faculty of Biotechnology, Chemistry and Food Science

Acknowledgments:

This thesis was performed at the Faculty of Chemistry, Biotechnology and Food Sciences, at the Norwegian University of Life Sciences, under the supervision of Professor Knut Rudi.

First, I would like to thank my supervisor, Knut Rudi for introducing me to this project. Knut Rudi has helped me throughout the whole project with his positive attitude, his clever ideas, solutions to problems, answering and correcting all questions that were put forth, and for that, he has my deepest gratitude.

I would also like to give my thanks to Laboratory Engineer Inga Leena Angell for all the help in the laboratory, for answering all questions, and for all her hours troubleshooting the gas chromatograph, which made this thesis possible.

I would also like to thank the Master students, Daniel Lycke Kristiansen, Inger Andrea Goa, Katrine Amlie, Amanda Trueman Morken and Siril Malene Isaksen for keeping me company in the laboratory. The same goes for the rest of the MiDiv group whose always been available to help and kept up the good spirits throughout the year.

I would like to give a special thanks to Karin C. Lødrup, and the rest of the scientists included in the PreventADALL project for the samples, all their information on these, and for making me feel welcome in the project.

Last, but not least, I would like to thank my family and girlfriend Ida, C. Sandli for all their support and help throughout my studies, and for always believing in me.

Ås, 2018

_____

Morten Nilsen

Abbreviations:

SCFA	Short Chain Fatty Acid
Meconium	First feces from a child
OTU	Operational Taxonomic Unit
GC	Gas Chromatograph
HPLC	High Performance Liquid Chromatography
NMR	Nuclear Magnetic Resonance
CE	Capillary Electrophoresis
FID	Flame Ionization Detector
TCD	Thermal Conductivity Detector
ECD	Electron Capture Detector
DNA	Deoxyribonucleic acid
dsDNA	Double Stranded Deoxyribonucleic acid
ssDNA	Single Stranded Deoxyribonucleic acid
PCR	Polymerase Chain Reaction
ddPCR	Droplet Digital Polymerase Chain Reaction
NGS	Next Generation Sequencing
PPi	Pyrophosphate
ATP	Adenosine triphosphate
SBS	Sequencing by synthesis
GI	Gastrointestinal
rRNA	Ribosomal ribonucleic acid

List of Figures:

List of Tables:

Abstract:

The mutualistic relationship between humans and our gut microbiota is essential in immune development. The microbial colonization of the human GI tract starts prenatal and continues with facultative and strict anaerobes as we age, due to gut environmental changes and diet. The bacteria in the gut microbiota utilize indigestible sugars and synthesize health-beneficial substrates, such as the short chain fatty acids (SCFAs). SCFAs are organic acids which act as a link between the gut microbiota and the maturation of our immune system. To investigate SCFAs' effect on our health, analysis of aging children might be key to prevent immunological disorders developing early life. Therefore, the aim of this study was to determine the microbial composition, short chain fatty acids composition and how these correlates in children their first year of life.

Fecal samples from 100 mothers and their children were received from the cohort study Prevent Atopic Dermatitis and Allergies (PreventADALL). The longitudinal sampling of children represented newborns, 3 months, 6 months and 12 months of age. The gut microbiota composition was determined by sequencing, while SCFAs composition was detected by gas chromatography.

The majority of the microbial composition as the children aged corresponded with earlier studies. The SCFAs composition changed significantly as the child aged in correlation with the microbial shifts observed. Both microbial and SCFAs composition of the children increased in similarity to their mothers as they aged. *Clostridium* represented 66% of the gut microbiota at the age of 12-months with a positive correlation to butyrate. The increased proportion of propionate in the same age group was positively correlated to *Bacteroidales*. These positive correlations strengthen recent knowledge of these bacteria being important SCFA producers in the gut microbiota. Between 6 and 12 months, the shift from an infant- to an adult-like gut microbiota and SCFAs composition might be initiated and influenced by weaning and introduction to solid foods. In conclusion, this work lays the foundation for further research investigating the immunological effects of SCFAs.

Sammendrag:

Det mutualistiske forholdet mellom mennesker og vår tarmmikrobiota er essensielt for immunutvikling. Den mikrobielle koloniseringen av menneskets mage-tarmkanal starter prenatal og fortsetter med fakultative og strikt anaerober som et resultat av forandringer i tarmmiljøet og diett. Bakteriene i tarmmikrobiotaen bryter ned ufordøyde sukkermolekyler og syntetiserer helsefremmende substrater, som kortkjedete fettsyrer. Kortkjedete fettsyrer er organiske syrer som virker som et bindeledd mellom tarmmikrobiotaen og utviklingen av vårt eget immunforsvar. For å videre undersøke kortkjedete fettsyrers effekt på vår helse, kan analyser av barn under oppvekst være en nøkkel for å forhindre utvikling av autoimmune sykdommer tidlig i livet. Målet for denne studien er å analysere den mikrobielle sammensetningen, kortkjedete fettsyre sammensetningen og hvordan disse korrelerer i barn under oppvekst.

Avføringsprøver fra 100 mødre og deres barn ble mottatt fra kohort studien Prevent Atopic Dermatitis and Allergies (PreventADALL). Den langsgående prøvetakingen av barn representerte nyfødte, 3 måneder, 6 måneder og 12 måneder gamle. Tarmmikrobiota sammensetningen ble fastslått ved hjelp av sekvensering, mens den kortkjedete fettsyresammensetningen ble funnet ved hjelp av gass kromatografi.

Majoriteten av den mikrobielle sammensetningen under oppvekst korresponderte med tidligere studier. Den kortkjedete fettsyre profilen viste signifikante forandringer i korrelasjon med de mikrobielle skiftene observert. Både mikrobiell og kortkjedet fettsyre sammensetning av barna økte i likhet til deres mødre under oppveksten. *Clostridium* representerte 66% av tarmmikrobiotaen til barna når de var 12-måneder gamle og var positivt korrelert til butyrat. Den økte proporsjonen av propionat i 12 måneders gruppen var positivt korrelert til *Bacteroidales*. Disse positive korrelasjonene styrker nylig kunnskap om at disse bakteriene er viktige kortkjedete fettsyre produsenter i tarmmikrobiotaen. Mellom 6 og 12 måneder kan forandringen fra barne- til voksenliknende mikrobiota og kortkjedete fettsyre sammensetning bli initiert og påvirket av brystmelk avvenning og introduksjon til fast føde. Dette arbeidet legger til grunne for videre forskning for å bedre forstå de kortkjedete fettsyrenes immunologiske effekt under barnets oppvekst.

Table of Contents:

1. Introduction:

1.1 The Human Gut Microbiota

The human gut microbiota is located in the gastrointestinal tract (GI) and comprises a large diversity of microorganisms, reaching up to 160 bacterial species (Rodríguez, J. M. et al. 2015). Most of the GI tract contains a sparse number of bacteria, with the exception of the colon, where most of the bacteria reside. The gut microbiota acts as a mutualistic relationship between bacteria and host and is of great importance for human health (Rodríguez, J. M. et al. 2015). This relationship affects us from the beginning of our lives, and throughout it. Our gut microbiota is constantly changing as we age, but several factors may disrupt the normal colonization pattern of the bacteria, which might affect the host's health throughout their lives.

The absorption of nutrients takes place in our GI tract, in proximity to the gut microbiota. The gut microbiota impacts our metabolism by degrading and producing substrates from indigestible sugars such as cellulose and xylans found in plants (LeBlanc, J. G. et al. 2017; Anand, P. A. A. et al. 2010). During the fermentation process, the bacteria produce different gases (CO_2, H_2, and CH_4), vitamins and short chain fatty acids (SCFAs) (LeBlanc, J. G. et al. 2017). SCFAs are readily utilized and help mature our immune system.

Several large projects, such as the Human Microbiome Project (HMP) have analyzed the gut microbiota to find a general bacterial composition (Turnbaugh, P. J. et al. 2007). As a result of the HMP, projects to determine the gut microbiota of individuals with different diseases soon followed. By analyzing sick subjects, such as individuals with type 2-diabetes, they found the gut microbiota to differ from the general gut microbiota composition (Woting, A. B. et al. 2016). This difference in healthy and sick individuals implies that a change in our gut microbiota can either be correlated with a disease or be the causative agent for the disease itself, although the exact relationship is yet to be determined (Woting, A. B. et al. 2016). To better understand these mechanisms, we need to understand the range of factors that influence the gut microbiota colonization.

The adult gut microbiota consists of facultative and strict anaerobic bacteria. This raised two hypotheses on how and when the development of the bacteria takes place (D'Argenio V & Salvatore, F. 2015). The first is based on the bacteria being transferred vertically from mothers to their children, while the other hypothesis states that the bacteria are acquired from the environment (Avershina, E. L. et al. 2016).

The human gut microbiota is dynamic due to the continual pressure from the environment as we age. External factors, diets, and antibiotics lay the foundation for the microbial shifts we experience (Rodríguez, J. M. et al. 2015). The main factor for the early microbial shift derives from internal changes in the gut environment, as a result of bacterial colonization and diet (Koenig, J. E. et al. 2011).

1.1.1 Human Gut Microbiota Development Early Life

The microbial colonization of the GI tract has been found to start prenatally. Studies supporting a prenatal colonization of the gut microbiota found bacteria in placentas (Satokari R. et al., 2008), umbilical cords (Jiménez, E. et al., 2005), amniotic fluid (DiGiulio D. B., et al. 2008) and in meconium (Jiménez, E. et al., 2008). In these studies, infants showed to harbor a complex community of bacteria. The gut microbiota consisted mostly of *Bacilli* and *Firmicutes*, with low proportions of *Enterococcus* and *Escherichia* (Moles, L. et al. 2013; Funkhouser, L. J. & Bordenstein, S. R. 2013). As the children aged to 3 weeks, a bacterial shift was discovered, where the *Proteobacteria* dominated (Moles, L. et al. 2013).

Located in the *Proteobacteria* phylum, we find *Enterobacteriaceae*, which are facultative anaerobic bacteria (Bøyre, K. 2014). The *Enterobacteriaceae* starts a selection process for strict and facultative anaerobic bacteria by depleting the gut of its oxygen, turning it into an anaerobic environment (Matamoros, S. et al. 2013). This environmental change favors the development of bacteria that are considered to be health-beneficial, such as *Bacteroides*, *Clostridium*, and *Bifidobacterium.* These bacteria rapidly colonize the gut and remain the dominant bacteria until breastmilk weaning commences (Matamoros, S. et al. 2013; Bäckhed, F. et al. 2015).

The increase of *Bifidobacterium* is due to another selection process connected to the infant's diet. The breastmilk acts as both a pro- and prebiotic for *Bifidobacterium*. The breastmilk harbors the bacteria and possesses complex sugars that *Bifidobacterium* can degrade (Avershina, E. L. et al. 2016; Grönlund, M. M. et al. 2017). As a result, the bacteria thrive in the gut during this time, but decreases as the amount of breastmilk is reduced (Bäckhed, F. et al. 2015). After weaning has begun and the children have been introduced to solid foods, *Bifidobacterium* decrease, and the gut microbiota starts to increase in similarity to their mothers (Bäckhed, F. et al. 2015). However, if the infants are formula fed, the child's gut microbiota is dominated by the *Enterobacteriaceae* family, as seen in Figure 1.1 (Matamoros, S. et al. 2013).

When analyzing sequencing data from bacteria, Operational Taxonomic Units (OTUs) are used. OTUs are based on similarities between DNA strands of bacteria. If two bacteria possess a 16S rRNA gene that is 97% or more equal, they are said to be the same OTU. If the gene similarity is less than 97%, they are considered two different OTUs, which represents two bacterial species. OTU prevalence has been shown to have a larger dissimilarity between mother and children up to the age of 1 year compared to mother and 2-year olds (Avershina, E. L. et al. 2016). The microbial increase in diversity and evenness of the gut microbiota indicate that the children reach an adult-like gut microbiota at the age of 2-3 years (Rodríguez, J. M. et al. 2015).

The human gut microbiota is less resilient early in life, and the colonization patterns of bacteria are easily disrupted (Lozupone, C. A. et al. 2012). Several factors influence this, such as antibiotic treatments early life, mode of delivery and diet (Rodríguez, J. M. et al. 2015; Matamoros, S. et al. 2013).

Mode of delivery affects which bacteria colonize the infant based on the surrounding environment. If the child is delivered vaginally, the gut microbiota consists of bacteria derived from the mother's vaginal flora (Dominguez-Bello M. G. et al. 2010). However, if the child is delivered with Cesarean section, an increase of *Staphylococcus*, *Corynebacterium, Propionibacterium* and other environmental bacteria dominate the gut microbiota (Dominguez-Bello M. G. et al. 2010).

A disruption of the gut microbiota colonization may affect the number of substrates produced from the fermentation process. As some of those substrates are important for the healthy maturation of the immune system, such a disruption may result in an improperly functioning immune system (Corrêa-Oliveira R. et al. 2016).
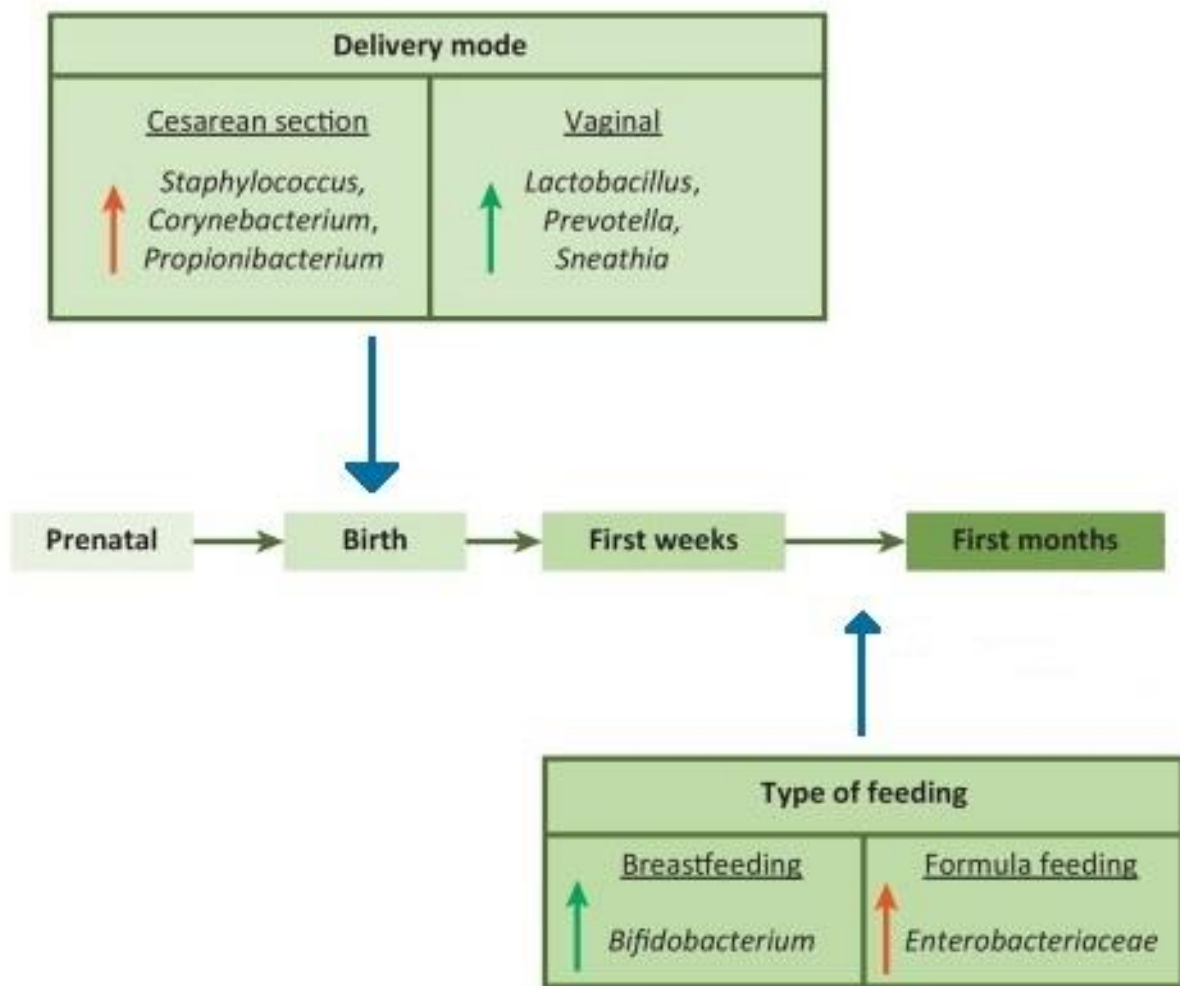
Figure 1.1. Illustration of two factors influencing gut microbiota colonization during the first months of life. The illustration shows the increase of bacteria by diet and delivery mode during the early months of life. Increase shown in red is an unwanted increase, while the green is a wanted increase in bacteria in the gut microbiota. The Figure is taken and edited from Matamoros, S. et al. (2013) Figure 1.

1.2 Short Chain Fatty Acids

The human gut microbiota has been compared to a metabolic organ because of its capabilities of breaking down indigestible food components and synthesizing other substrates, such as SCFAs (Xu, J. & Gordon I. J. 2003). Indigestible carbohydrates consisting of β(1→4) linkage bonds cannot be degraded by human enzymes themselves, and this is where the gut microbiota's mutualistic relationship to the host arise (Mathews, C. K. et al. 2013; Anand, P. A. A. et al. 2010). The bacteria degrade the indigestible carbohydrates by saccharolytic fermentation, resulting in beneficial substrates and SCFAs as the end products (Roy, C. C. et al. 2006; Morrison, D. J. & Preston, T. 2016).

Short chain fatty acids (SCFAs), or volatile fatty acids (VFAs), are organic acids consisting of one to six carbons (den Besten, G. et al. 2013). The three most common and dominant SCFAs produced by the gut microbiota are; acetate (C2), propionate (C3), and butyrate (C4) (den Besten, G. et al. 2013). The SCFA composition in adults has been found to have a ratio of 60:20:20, respectively (Binder, H. J. 2010).

1.2.1 Short Chain Fatty Acid Production in Humans

In the adult gut microbiota, the dominant bacterial phyla *Bacteroidetes* and *Firmicutes* are well-known SCFA producers (Woting, A. B. et al. 2016; den Besten, G. et al. 2013). The *Bacteroidetes* phylum mainly produces acetate and propionate (Macfarlane S, Macfarlane G. T. 2007). *Firmicutes* contains butyrate producers such as *Clostridium*, in addition to some propionate producers (Macfarlane S, Macfarlane G. T. 2007).

The production of butyrate is divided into two main pathways; the butyryl-CoA pathway and butyryl-CoA:acetate CoA-transferase pathway (Flint, H. J. et al. 2014). The most utilized pathway of the two is the butyryl-CoA:acetate CoA-transferase pathway (Louis, P. et al. 2010).

There are three major propionate formation pathways; The succinate pathway, acrylate pathway and the propanediol pathway (Reichardt, N. et al. 2014). The succinate pathway is the one used by both *Bacteroidetes* and *Firmicutes* (Flint, H. J. et al 2014).

For acetate there are two main production pathways in the gut microbiota; the Wood-Ljungdahl pathway (Miller, T. L. & Wolin, M. J. 1996), and the result of carbohydrate fermentation (Ríos-Covián, D. et al. 2016).

1.2.2 SCFAs' Effect in Humans

The gut microbiota's production of SCFA metabolites acts as a link between the microbiota and our immune system (Corrêa-Oliveira R. et al. 2016). The several SCFAs produced by our gut microbiota is utilized differently in the host.

Butyrate is partially used as an energy source for the epithelial colonic cells known as colonocytes (Corrêa-Oliveira R. et al. 2016). It has been demonstrated that 70% of the energy obtained by the colonocytes derives from butyrate, and studies have shown that it is the preferred energy source when affinities to butyrate were compared with glutamine, glucose, acetate, and propionate (Roy, C. C. et al. 2006). This makes butyrate an important SCFA for humans (LeBlanc, J. G. et al. 2017). In addition to being an energy source for the colonocytes, butyrate has been found to possess the properties of inhibiting colon carcinoma cell growth, produce growth arrest, help with cell apoptosis and differentiation, and it might play an important role in preventing colon cancer (Velázquez M. et al. 2000; Roy C. C. et al. 2006).

Acetate mainly enters the peripheral circulation, where it is metabolized (Wong, J. M. W. et al. 2006). The acetate may inhibit and suppress accumulation of body fat and liver lipids, and if it is absorbed in the colon, an increase of cholesterol synthesis follows (Kondo, T. et al. 2009; Wong, J. M. W. et al. 2006).

The majority of propionate is taken up by the liver, where it is metabolized with residual butyrate and used in the gluconeogenesis (Wong, J. M. W. et al. 2010). Propionate has been shown to lower fatty acid content and might improve tissue insulin sensitivity, which is considered beneficial for obese individuals, or individuals with diabetes type 2 (Al-Lahham, S. H. et al. 2010). The propionate is also a gluconeogenerator, and counteracts the cholesterol synthesis from acetate, resulting in decreased cholesterol (Wong, J. M. W. et al. 2006). Substrates capable of decreasing the acetate:propionate ratio may decrease the chances of cardiovascular diseases (Wong, J. M. W. et al. 2006).

SCFAs have been found to help mature our immune system and have a positive impact on its development (Corrêa-Oliveira R. et al. 2016). An example of the mutualistic relationship between the gut microbiota and host is that the major SCFAs help increase the expression of antimicrobial peptides, which is secreted to an external surface by the epithelial cells (Corrêa-Oliveira R. et al. 2016). The bacteria which normally inhabit our gut microbiota have been shown to possess a higher tolerance for antimicrobial peptides, giving them an advantage of survival compared to the pathogenic microorganisms trying to colonize (Cullen, T. W. et al. 2015).

SCFAs have also been found to modulate the production of mediators such as IL-18, which is a proinflammatory cytokine that repairs and maintains epithelial cell integrity (Dinarello, C. A. 1999; Corrêa-Oliveira R. et al. 2016). Immune cell differentiation, recruitment, and activation are all affected by the SCFAs, as illustrated in Figure 1.2 (Corrêa-Oliveira R. et al. 2016).
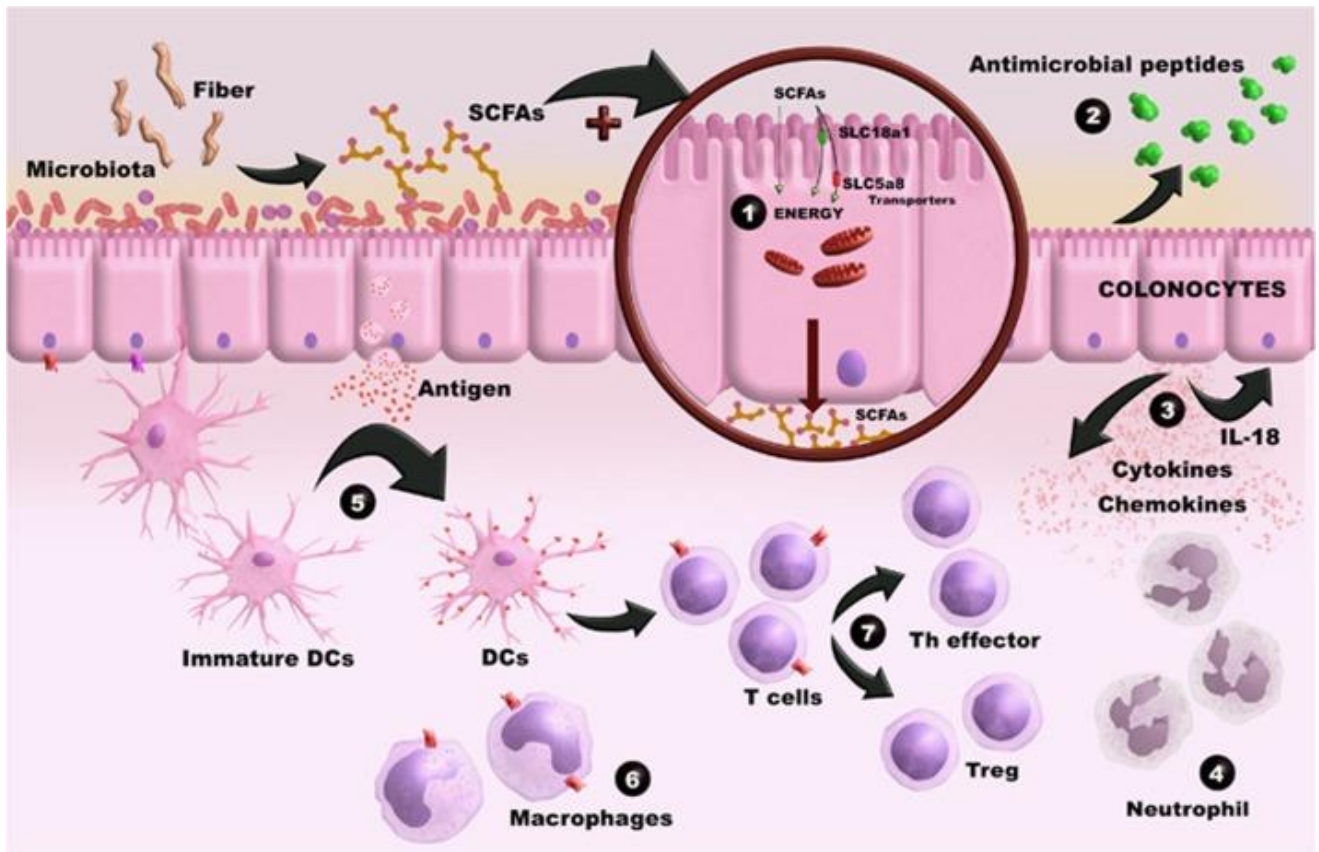
Figure 1.2. Illustration on how SCFA affects our immune system. The illustration shows how SCFAs produced by microbiota gives the colonocytes their energy, the production of antimicrobial peptides, recruitment, activation and differentiation of our immune cells. The illustration is made and taken from Corrêa-Oliveira R. et al. (2006) Figure 2.

To analyze the composition and production of SCFAs in humans, fecal samples are often used as they are easily accessible. The most used techniques for this are high performance liquid chromatography (HPLC), nuclear magnetic resonance (NMR), capillary electrophoresis (CE) and gas chromatography (GC), the latter being the most commonly used technique (Primec, M. et al. 2017).

1.3 Gas Chromatograph for Short Chain Fatty Acids Analysis

The principle of gas chromatography is to separate several compounds into individual ones by using a mobile and stationary phase (Grob, R.L & Barry, E. F, 2004; Linde A. G., GC, 2018). By separating the mixture into individual compounds, they can be quantitatively or qualitatively measured (Linde, A. G., GC, 2018).

The prepared sample is injected into the gas chromatograph and evaporates in the injector. The components of the mixture will then be separated based on physiochemical properties through a column, before the compounds reach a detector that detects their qualitative and quantitative properties (Linde, A. G., GC, 2018). This is then analyzed by a computer.

The compounds are separated based on their molecular weight, melting point, column material, column temperature and carrier gas (Linde, A. G., GC, 2018). All these factors ensure that different compounds reach the detector with different retention times, which is then used to identify the compounds in the mixture. Low molecular compounds reach the detector first, such as acetate (C2), followed by propionate (C3) (Linde, A. G., FID, 2018).

There are several individual parts that might affect the result of a GC run. The liner serves an important function, where the sample which is injected pass from a liquid phase to a gaseous phase, and further evaporates into the GC column. Several types of liners can be used, and the liner should be chosen based on the sample material (Morgan, P. 2012). After evaporation, the sample is carried through the GC column by a carrier gas (Linde, A. G. GC 2018).

Several gases can be used as a carrier gas if they meet the requirement of being inert, meaning they don't react with the molecules in the sample nor react with the stationary phase (the column) (Linde, A. G., GC. 2018). The most commonly used carrier gases are hydrogen, helium, and nitrogen, which all have different properties and react differently when used. The samples are then introduced to the column which is covered by a liquid or film that affects the gas compounds. The two most used columns are packed and capillary column (Linde, A. G., GC. 2018).

When the molecules reach the end of the column, they meet the detector which analyzes the individual compounds. The detectors used are often Flame Ionization Detector (FID), Thermal Conductivity Detector (TCD) or Electron Capture Detector (ECD) (Linde, A. G., GC. 2018).

The FID consists of a hydrogen/air flame with a collector plate, which introduces a flame to the passing molecules which then ionizes and can be detected (Linde, A. G., FID. 2018; Cambustion, 2018). The FID is often used to detect hydrocarbons (Linde, A. G., FID. 2018).

1.4 DNA Extraction, Quantification and Sequencing

1.4.1 DNA Extraction

DNA extraction is crucial for all the techniques following using DNA. If DNA extraction is not sufficient, the samples will not be representative. To obtain the DNA from the bacterial cells, the cell walls of the bacteria needs to be disrupted. There are three methods to disrupt the cell-walls: chemical, mechanical and enzymatic. These can either be used by themselves, or in a combination, depending on the bacteria and sample of interest. The mechanical lysis is often obtained by using glass beads with the sample material, and when shook sufficiently the glass beads disrupts the cell walls. Chemical lysis is obtained by using chemicals that react and disrupts the cell walls, which is also the case with enzymatic disruption, but by using enzymes instead of chemicals. The disruption is important to obtain DNA from the samples, which in turn is used for sequencing. To be able to sequence, the DNA needs to be amplified, which is done by polymerase chain reactions (PCR).

1.4.2 Polymerase Chain Reactions

Polymerase chain reactions (PCR) are used to study diverse and complex microbial communities, such as the gut microbiota. By using PCR based methods, we can avoid the problematic issues of cultivating the bacteria. The PCR has become a key technique in microbiology and other scientific fields. PCR can be divided into two main techniques: quantitative PCR and qualitative PCR.

Qualitative PCR

The PCR principle is separated into three sections. The first is denaturation of the double stranded DNA (dsDNA). This is done by heating up the samples, 95°C is usually sufficient. The dsDNA denatures at different temperatures based on its GC content, a high GC content makes the DNA more heat stable. The denaturation of the DNA ensures that primers will be able to attach themselves to their specific sites on the DNA (Schochetman, G. et al. 1988). The second step is the annealing of primers onto to the DNA strands. The primers are single stranded oligonucleotides and complementary to the strand of the original DNA (Schochetman, G. et al. 1988). The third and last step is the primer extension. The primer extension is done by a DNA polymerase that attaches nucleotides behind the primer site, making a replica of the chosen DNA region (Schochetman, G. et al. 1988).

These 3 steps will then be repeated a given number of times to amplify the specific region from the DNA strands. A single DNA strand can be amplified to 1,000,000 DNA fragments after 30 cycles (Schochetman, G. et al. 1988). One of the challenges of the PCR-based method was finding a suitable DNA polymerase that could perform primer extension at high temperatures. Today, the most used DNA polymerase derives from *Termophilus aquaticus,* a bacterium that lives close to hot springs, and is adapted to high temperatures, making its DNA polymerase heat-stable (Schochetman, G. et al. 1988). The qualitative PCR is an end-point analysis and needs a second step of verification. This step is often agarose gel electrophoresis to review the fragment of interest.

Quantitative PCR

The principle for quantitative PCR is the same as qualitative PCR. The difference between the two is that the quantitative PCR analyzes real-time. To measure the DNA fragments in real-time, different fluorophores are being used. The fluorophores used are often non-specific and bind to all dsDNA in a sample. Examples of fluorophores used are SYBR Green I, which is a non-specific binding fluorophore and LUX which is a fluorophore-labeled primer. For sequence-specific primers; Scorpions can be used (Busing, S.A. 2005). The quantitative PCR detects the fluorophore when it reaches a set threshold value. By taking account of the PCR efficiency and the number of cycles processed before the threshold value was reached, the original amount of DNA fragments can be calculated.

Droplet Digital PCR (ddPCR)

Droplet digital PCR (ddPCR) is a PCR-based method using end-point analysis to calculate absolute concentrations of a given sample. In ddPCR, the target molecules are distributed across droplets, where some have the template and other droplets do not (Hindson, B. J. et al. 2011). By amplifying the template using qualitative PCR with fluorophores, the absolute concentration of the sample can be calculated (Hindson, B. J. et al. 2011).

The original sample is mixed with droplet generating oil before being put onto a droplet generator which vacuums both the sample and the oil through a single tube (Hindson, B.J. et al. 2011). Because the oil is not water soluble, the sample will disperse, creating droplets.

The DNA template in the droplets will then be amplified by qualitative PCR with fluorophores. After amplification, the droplets containing DNA template binds the fluorophore (Hindson, B. J. et al. 2011). The sample can then be analyzed using a ddPCR machine, which detects droplets containing fluorophores. This is a digital end-point analysis, and the detector either finds a droplet containing the fragment or not, based on fluorophore readings. After determining the number of droplets containing fragments, a Poisson model is used to determine if the droplets contain one or more fragments and calculate the absolute concentration of the fragments in the samples (Hindson, B. J. et al. 2011).

1.4.3 Sequencing Technologies

Sequencing technology is crucial tools for analyzing complex microbial communities, and several techniques have been designed. Even though there have been advancements in the technology, some of the oldest techniques are still being used. The real breakthrough in sequencing technology started in 1977, when Carl Woese & Fox suggested the use of ribosomal RNA genes as molecular markers for bacterial classification. This later developed into using 16S rRNA genes, referred to as 16S (Escobar-Zepeda, A. et al. 2015). This unit has conserved regions and 9 variable regions. The conserved regions make amplification of 16S from different species possible, while the variable regions distinguish the different species. This idea combined with PCR technology and Sanger sequencing revolutionized the field.

Sanger sequencing was one of the first methods that could sequence DNA fragments and was invented by Sanger in 1977 (Sanger, F. et al. 1977). The Sanger sequencing principle is using 2'3'-dideoxynucleotides (ddNTP) to stop elongation. By using a combination of all four nucleotides as ddNTP, the elongation stops at different incorporations. By using restriction enzymes and agarose gel, the nucleotides could be read off the gel based on their length and cut-offs (Sanger, F. et al. 1977). This technique is still being used and can sequence up to 96 sequences per run with an average base pair length of 650 (Escobar-Zepeda, A. et al. 2015).

Sanger sequencing became more popular than the first sequencing method from Maxam A. M. and Gilbert, W. Their technique used chemicals to break terminally labeled molecules partially at each base repetition (Maxam A.M. & Gilbert, W. 1977). The length of the labeled fragment would then identify the position of that base (Maxam A.M. & Gilbert, W. 1977). This technique could at minimum sequence 100 bases by analyzing the fragments on a polyacrylamide gel to read off patterns (Maxam A. M. & Gilbert, W. 1977). After Sanger sequencing, a race to create a better sequencing technology started, and the first company to make a valid sequencer was Roche, which is referred to as Next Generation Sequencing (NGS) or 2nd generation sequencing (Liu, L. et al. 2012).

The most popular sequencing technologies in the 2nd generation sequencing were made by Roche and Illumina. The Roche 454 uses a technique referred to as pyrosequencing. Pyrosequencing uses a detection of light emission based on pyrophosphate (Liu, L. et al. 2012). When a nucleotide is incorporated to the DNA strand, pyrophosphate (PPi) is generated. The PPi released is equal to the number of incorporated nucleotides and generates ATP with the help of enzymes (Liu, L. et al. 2012). The ATP is used to oxidize luciferin to oxyluciferin, which emits light. A higher amount of ATP will oxidize more luciferin, resulting in more emitted light (Liu, L. et al. 2012). The detector is then able to detect the number of nucleotides, and which nucleotide is incorporated (Liu, L. et al. 2012). After Roche, the next big contender is Illumina.

Illumina's technology is based on sequencing by synthesis (SBS). Illumina uses a flowcell consisting of 2 different oligos (Illumina, Inc. 2017). By using two different adapters, and attaching them to the fragment of interest, the fragments will be able to attach themselves to the oligos (Illumina, Inc. 2017). The adapters contain a primer binding site, barcode and the complementary nucleotides to the oligos on the flowcell (Illumina, Inc. 2017). The DNA fragment will attach itself to the oligos on both sides, creating a bridge which is amplified, called bridge amplification. After the amplification steps, a cluster is created, which derives from one DNA fragment. After cluster generation, the SBS starts (Illumina, Inc. 2017).

The SBS is based on nucleotides with fluorophores that are being incorporated into their complementary base at the DNA strand (Illumina, Inc. 2017). One cluster contains identical DNA fragments, so the same nucleotide with the same fluorophore will be incorporated and emit one color, which is then detected (Illumina, Inc. 2017). The next nucleotide will then be incorporated, emit its color corresponding to the nucleotide until the fragment is sequenced. After the first fragment is sequenced, a new bridge amplification follows, and the reverse strand is sequenced by the same method. A standard is used with the samples, PhiX. PhiX helps to distinguish the different clusters in the flowcell (Illumina, Inc. 2017). Due to 16S' conserved regions, the cluster will emit the same color when sequencing these regions. The PhiX have different bases than the conserved region of 16S and helps the detector to distinguish clusters emitting the same color (Illumina, Inc. 2017).

The 3rd generation of sequencing technologies is being developed and is expected to soon be available. The two most anticipated technologies are the PacBio and Oxford Nanopore MinION sequencer. Both technologies promise to obtain longer read lengths than we are capable of today, and a decrease per-base cost (Pacific Biosciences of California 2018; Mikheyev A. S. & Tin M. M. Y. 2014). The MinION is a handheld device, making it easier to be able to sequence outside of the laboratory.

1.5 My project:

There has been an emerging interest in studying the gut microbiota because of its effects on human health. The short chain fatty acids produced by the gut microbiota might play a key role in our immune development. Changes in the normal gut microbiota composition have been shown to be correlated with different diseases, and these changes may alter the SCFA composition. Earlier research on SCFAs has focused on adults and their health. However, the role of SCFAs in the transition from an infant- to an adult like gut microbiota is still not known.

To be able to determine the effects of SCFAs in children between an infant and adult-like gut microbiota, a short chain fatty acids profile for the children needs to be established. Therefore, the main objectives of this master thesis project are:

1. Establish the gut microbiota composition of children between infancy and 12 months of age

2. Establish a SCFAs profile for the same children

3. Correlate the SCFAs profiles to the bacterial profiles

The fecal samples were obtained from a mother-children cohort study, PreventADALL (Prevent Atopic Diseases and Allergy) and was analyzed for microbial and SCFAs composition by sequencing and gas chromatography. PreventADALL's aim is to find simple and low-cost strategies to prevent allergic diseases during the early stages of life. PreventADALL is an international collaboration between research groups, with the main coordinators located at the University of Oslo (UiO). PreventADALL started their sampling in 2015 on woman 18 weeks into pregnancy and thereafter included their children. Urine, blood-tests, breastmilk, saliva and fecal samples are some of the sampling done for the mother-child pairs, in addition to documenting the children's diet throughout the time period. Sampling of the children started at birth and continued in intervals until the age of 36 months (0, 3, 6, 12, 24 and 36 months). From these, 500 fecal samples were received, representing 100 mothers and their children. The samples represented the children's development from infancy to 12 months.

2. Materials & Methods:

The fecal samples were diluted 1:10 in DNA shield buffer to prevent degradation of DNA and stored at -80°C until use. To profile SCFAs for the age groups, a gas chromatograph was used, a Trace 1310 equipped with an autosampler (ThermoFisher Scientific). To analyze the gut microbiota, 16S rRNA genes from bacteria in the fecal samples were extracted and sequenced using Next Generation Sequencing (NGS), on an Illumina MiSeq. A flowchart illustrating the workflow of the master thesis is shown in Figure 2.1
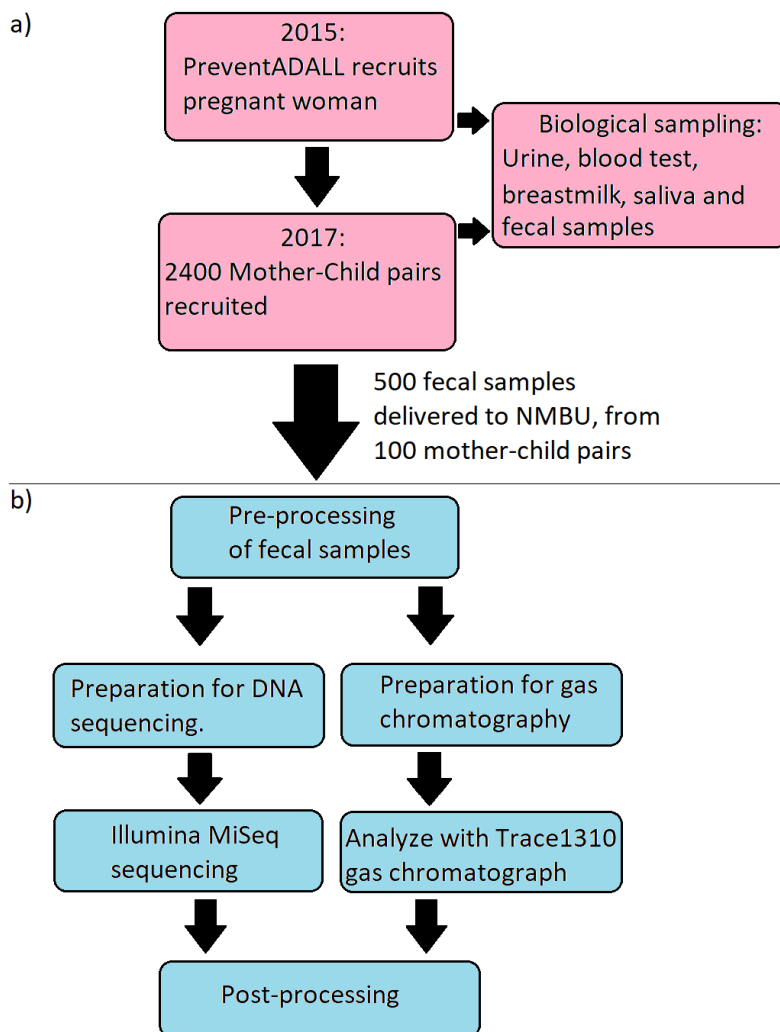
Figure 2.1. Flowchart. The flowchart shows the workflow of the project. a) represents sampling by the PreventADALL project. b) represents the workflow in this thesis. PreventADALL started collecting biological samples in 2015 and reached 2400 mother-child pairs in 2017. From the biological samples, fecal material was used for this thesis. PreventADALL retrieved biological material from children at 0 (newborn), 3, 6, 12, 24 and 36 months. One hundred samples from each age group up to 12 months were analyzed, resulting in 500 samples in total. The samples were prepared for DNA sequencing, sequenced on Illumina MiSeq, and analyzed for SCFAs on a Trace1310 gas chromatograph. Post-processing includes the use of the QIIME pipeline to process sequences, while statistics were done using R, and Spearmann correlations using MatLab.

2.1 Sample Preparation:

Fecal samples were thawed on ice, and vortexed to homogenize the samples. The samples were then pulse centrifuged at 1200 rpm for 8 seconds, making the extraction easier by aggregating bigger fecal particles. From the 1:10 diluted samples, 300µl were used for gut microbiota sequencing and 100µl for SCFA profiling.

2.2 Short Chain Fatty Acids Profiling

The 1:10 diluted fecal samples were diluted 1:1 with MilliQ-water, and then a 1:1 ratio of internal standard, containing 2% formic acid with 500µM 2-methylvaleric acid. Formic acid was used to prevent peak tailing, ghosting of the peaks, and to activate the SCFAs by lowering the samples' pH (Tangerman A & Nagengast F. M. 1996). In addition, formic acid contains few C-H groups, leaving it often undetected by gas chromatography (Waksmonski, M. 2015). 2-methylvaleric acid was included in the internal standard to obtain absolute quantitative concentrations. The diluted samples were centrifuged at 13 000 rpm for 10 minutes, making a pellet of the bigger particles, leaving the supernatant easily accessible. The supernatant was transferred to filter columns with 0.2µM filters (VWR, USA) to filtrate smaller particles, and centrifuged at 10 000 rpm for 5 minutes.

The eluate was transferred to 300µl GC vials (VWR, USA) and applied on the gas chromatograph. The fecal samples were analyzed on a Trace 1310 equipped with an autosampler (TermoFisher Scientific, USA) with ramping temperatures: from 90°C to 150°C for 6 minutes, and 150°C to 245°C for 1,9 minutes. Detailed GC specifications can be seen in Appendix A. In between every 10 samples a standard was run twice to inspect shifts or variabilities of the run. The standard consisted of 300µM acetic acid, 12µM propionic acid, 8µM isobutyric acid, 12µM butyric acid, 8µM isovaleric acid, 8µM valeric acid, 25µM of the internal standard and 1% formic acid. All acids used in the standard were purchased from Sigma-Aldrich, Germany.

2.3 DNA Purification

2.3.1 Bacterial Lysis

Bacterial DNA extraction was performed by a combination of chemical and mechanical cell lysis. Samples were mixed with 3 different sizes of glass beads. To 300µl sample, 0,2g of acid-washed glass beads (<106µm, Sigma-Aldrich, Germany), 0,2g acid-washed glass beads (425-600µm, Sigma-Aldrich, Germany) and 2x 2.5-3.5mm acid-washed glass beads (Sigma-Aldrich, Germany) were added. A combination of bead sizes performs better than only one size, but DNA shearing is aggravated (Bakken, L. R. & Frostegård, Å. 2006). This combination is to ensure a higher diversity and less bias when mechanically disrupting the bacterial cells. The samples were processed twice on a FastPrep 96 (MP Biomedicals, USA) at 1800 rpm for 40 seconds, and centrifuged at 13 000 rpm for 5 minutes. The centrifugation aggregates cell remains and glass beads, removing the residue before further DNA treatment.

The chemical lysis was performed using lysis buffer and Proteinase K. The mechanically lysed samples were treated with lysis buffer and proteinase K on a KingFisher Flex Robot. The lysis buffer lyses the cells due to the high salt concentration, while Proteinase K digests proteins in the samples, such as nucleases which destroys DNA (Thermo Fisher Scientific, Proteinase K).

2.3.2 DNA Extraction

Paramagnetic particles were used to extract DNA from the lysed samples. DNA binds to paramagnetic particles in a high salt concentration, such as NaI or $NaClO_4$, creating a salt bridge between the DNA and the particles (Boom, R. et al. 1990). DNA was extracted on a KingFisher Flex Robot, using LGC mag midi kit (LGC Genomics, UK). Ethanol was added to the samples and precipitated the DNA. The samples were then washed using washing buffers containing salts, binding the DNA to the paramagnetic particles. Water was used to eluate the DNA by disrupting the salt bridge between the DNA and paramagnetic particles. DNA was then extracted from the eluate. Quantitative tests were later performed to examine the DNA extraction.

2.3.3 PCR Product Clean-up

Sera-Mag beads were used to remove nucleotides, primers, and polymerases after PCR. Clean-up after amplicon PCR was performed on a Biomek 3000 (Beckman Coulter, USA) by adding 1x volume of beads to the volume of DNA sample, following AMPure's protocol. Clean-up of the pooled library was performed manually by using 0.1 % Sera-Mag beads, following the AMPure protocol.

2.4 Polymerase Chain Reactions

2.4.1 Quantitative PCR

Quantification of bacteria was performed by quantitative PCR using primers targeting the V3 and V4 regions of 16S rRNA genes (PRK primers). Samples were amplified and quantified on a LightCycler480 II with PRK primers (Yu, Y. et al. 2005). The reactions contained: 1x HotFirePol EvaGreen qPCR supermix (Solis BioDyne, Germany), 0.2µM PRK Forward & Reverse primer and 2µl template DNA. The samples were amplified with the following program: 95°C for 15 minutes, followed by 45 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 45 seconds.

2.4.2 Qualitative PCR

Amplification of bacteria was performed by PCR using PRK primers. DNA template (2µl) with 1x HotFirePol Blend Master Mix Ready to Load (Solis BioDyne, Germany), and 0,2 µM PRK forward and reverse primer (Yu, Y. et al. 2005) were mixed. The samples were amplified using the following program: 95°C for 15 minutes followed by 25 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 45 seconds, before a final step at 72°C for 7 minutes. For the meconium samples, the number of cycles was increased to 30.

2.4.3 Index PCR

Index PCR was performed in order to sequence several samples simultaneously, by attaching Illumina adapters to 16S rRNA fragments. A combination of 16 different forward primers and 30 reverse primers was used (Appendix B), making it possible to barcode 480 samples. Each reaction consisted of 1x FirePol Master Mix Ready to Load (Solis BioDyne, Germany), 0.2µM forward & reverse primers, nuclease free-water (VWR, USA) and 1µl DNA. The samples were amplified with: 95°C for 5 minutes followed by 10 cycles of 95°C for 30 seconds, 55°C for 60 seconds, and 72°C for 45 seconds, before a final step of 72°C for 7 minutes.

2.4.4 Droplet Digital PCR (ddPCR)

Droplet Digital PCR (ddPCR) was used to quantify the amount of DNA in the samples. A dilution series from 10^{-5} to 10^{-11} was used for quantification. The diluted samples were mixed with 1x Super mix for EvaGreen (BioRad, USA), 0.2µM Illumina colony forward & reverse primer, 2.4µl DNA template and PCR water. Droplet generation was performed using BioRad QX200^{TM} – Droplet Generator, following BioRad's instructions, and the plate was sealed using BioRad PX1^{TM} PCR Plate Sealer. The samples were amplified on PCR using the following program: 95°C for 5 minutes followed by 40 cycles of 95°C for 30 seconds, 60°C for 30 seconds, and 72°C for 45 seconds. The last two steps took place at 4°C for 5 minutes and 90°C for 5 minutes. Following PCR amplification, the amplicons were quantified on BioRad QX200^{TM} Droplet Reader. Absolute concentration was calculated by using the ddPCR.

2.5 Sequencing

For Illumina MiSeq sequencing, a pooled 4nM sample was used. The sample was diluted to 6pM following Illumina's instructions, except nuclease-free water was used instead of Tris in the preparation. A PhiX control was combined with the 6pM sample, resulting in the final concentration containing 15% PhiX. The samples were applied to an Illumina MiSeq.

2.6 DNA Quantity and Quality Control

2.6.1 Qubit

Verification of DNA extraction was performed by Qubit measurement, following the manufacturer's recommendations.

2.6.2 Gel Electrophoresis

Samples were run on a 1,5% agarose gel at 85 volts for 20 minutes. The gel was visualized using Molecular Imager Gel DOC^{TM} XR Imaging Systems, to review PCR products qualitatively.

2.6.3 Quantification and Normalization

To quantify and normalize the DNA in all samples, a Cambrex-FLX800 CSE was used. All samples were added Qubit reagent, according to Qubits instructions. A standard curve was made using the fluorescence data from the lowest to the highest fluorescence value. DNA concentrations were calculated based on the standard curve and normalized to a pooled library. The samples were normalized and pooled using a Biomek 3000. The maximum pipetting volume on the Biomek 3000 was 10µl, resulting in 3 samples not being added in their adequate volume. Samples without sufficient DNA were not normalized, to prevent diluting the pooled library.

2.7 Data Processing

2.7.1 QIIME

The sequencing file was received as a FASTQ file. To process the sequencing data, the QIIME pipeline was used. QIIME started with preprocessing the FASTQ file by decompressing it, extracting the barcodes, assembling forward and reverse reads and splitting the library into their respective samples (Huang, H. 2014). From the new file created, OTU processing followed. OTU processing checked for chimeras and created OTUs based on the sequence similarities between the fragments sequenced. Each OTU represented 97% or higher similarity between the sequences. To apply the taxonomy to OTUs, the SILVA database was used, which is a quality-controlled database of rRNA sequences from Bacteria, Archaea and Eukarya domains (Pruesse, E. et al. 2007). A consensus sequence from each OTU was used in the database and the OTUs were assigned their taxonomy. After the OTU processing, alpha and beta-diversity indexes were calculated for the samples.

2.7.2 Paired T-Test

Paired t-tests were performed to analyze if a statistical significance was present between SCFAs or bacterial orders between age groups. The paired t-tests were performed using Rcommander with a 95% confidence interval. The R version used with packages was R version 3.4.3, mixlm version 1.2.3, R commander version 2.4-0 and RcmndrPlugin.NMBU version 1.8.8.

2.7.3 Spearmann Correlations

Spearmann correlations were used to associate bacterial profiles with the different SCFAs. The correlation was performed by using the Benjamini-Hochberg method, with a p-value less than 0.05. The analysis was performed by Knut Rudi in the MatLab programming environment (MathWorks Natick, USA).

3. Results:
3.1 16S rRNA Gene Sequencing

To avoid DNA contaminants, samples having a Cq-value below 4 to the negative control on quantitative PCR were considered having sufficient DNA. From the total of 500 samples, 70 samples were discarded because of high Cq-values, all belonging to the meconium group. By combining gel electrophoresis and a standard curve based on fluorescence values (obtained from Cambrex-FLX800 CSE), 29 samples scattered amongst the age groups were found to have inadequate amounts of DNA and were discarded, resulting in a total of 401 samples before sequencing.

Two sequencing runs were performed, where the first resulted in 180K/mm^2 clustering, with 8,726,156 ssDNA fragments. The second run gained a clustering of 460K/mm^2, resulting in 22,152,156 sequenced ssDNA fragments. A cut-off was set at 5,000 dsDNA fragments per sample, resulting in 352 samples with sufficient quality. This was distributed as; meconium n=10, 3 months n=79, 6 months n=76, 12 months n=94 and mothers n=93.

3.1.1 Binning of Sequences in OTUs

Sequences acquired were binned in OTUs by the QIIME pipeline. The total count of OTUs was 1061 with a sampling depth of 5,000. The taxonomy assigned to the OTUs derives from the QIIME pipeline and is presented in Figure 3.1, showing the abundance of bacterial orders in the different age groups.

The difference in means between the bacterial orders was determined by paired t-tests for the most dominant bacterial orders within each age group. *Enterobacteriales* decreased significantly (p=0,0049) from meconium to 3 months of age, while *Bifidobacteriales* significantly increased (p=0,0088) for the same age group. The *Clostridiales* order more than doubled in proportion between the age of 6 and 12 months (p=<2*e^{-16}). At the age of 12 months, the gut microbiota was composed of 66,6% *Clostridiales*, where *Faecalibacterium* (14%), *Gnavus* group (8,8%) and *Lachnospiricieae's* rectale group (6,9%) were the most abundant. *Bifidobacteriales* significantly decreased (p=<2*e^{-14}) from the age of 6 to 12 months. Genus and Family level of the bacterial taxonomy is presented in Appendix C, while raw data for the OTU table & p-values are presented in Supplementary Tables and Figures.

Figure 3.1. Amount of bacterial orders within each age group. The bar chart shows the amount of bacterial orders acquired from sequencing with processing by the QIIME pipeline. The bacterial orders are divided between the age groups: meconium, 3 months, 6 months, 12 months and mothers. The most dominant orders of bacteria are displayed top right, with their respective colors. The asterisks represent their p-value of a paired t-test measuring the mean between the groups. One asterisk represents a p-value between 0.05 – 0.005, two represents a p-value between 0.0049 to 0.0005, and three asterisks represents p-value <0,00049. Genus and Family level of the bacterial taxonomy is presented in Appendix C, while raw data for the OTU table and p-values are shown in Supplementary Tables and Figures.

3.1.2 Alpha-Diversity:

Diversity within each age group was determined by using alpha-diversity indexes. Species richness and evenness within each age group were analyzed using the alpha-diversity indexes derived from the QIIME pipeline. Observed species, Shannon-Wiener index and inverted Simpson's index were calculated and are presented in Figure 3.2 as a), b) and c) respectively.

The lowest amount of species found was 26, belonging to the meconium age group, as shown in Figure 3.4 a). The number of unique species observed continually increased as the child aged, reaching a total of 72 species for 12 months. The highest amount of observed species was found in the mother group, holding 183 species. There was a significant increase in observed species for each age group between 3 months and mothers ($p=0,015$).

The Shannon-Wiener index represents both unique species and their evenness (Shannon, C.E & Weaver, W. 1949). The three youngest age groups (Meconium, 3 months and 6 months) all had a low Shannon-Wiener index value, representing low microbial diversity with an uneven distribution of bacteria, as shown in Figure 3.2 b). As for the 12 months-group, there was a significant increase in diversity and evenness from the 6 months-group ($p=0,015$). Mothers had the highest Shannon-Wiener index value, representing the highest diversity and evenness amongst the groups.

The inverted Simpsons index in figure 3.2 c) has the same interpretation of the index values as the Shannon-Wiener index from 3.2 b). There was a significant increase in evenness and microbial diversity between each age group from meconium to mothers, except for 3 and 6 months ($p=0,87$), which resemble each other as opposed to in the Shannon-Wiener index. Corresponding to the Shannon-Wiener index, mothers show the highest diversity and evenness amongst the groups.

To estimate the number of species in the samples based on the observed species found, the Chao1 index was used (data not shown). The mean species estimate between all pairs showed a statistical significance between the means ($p=0,015$), except between 3 and 6 months of age ($p=1$).
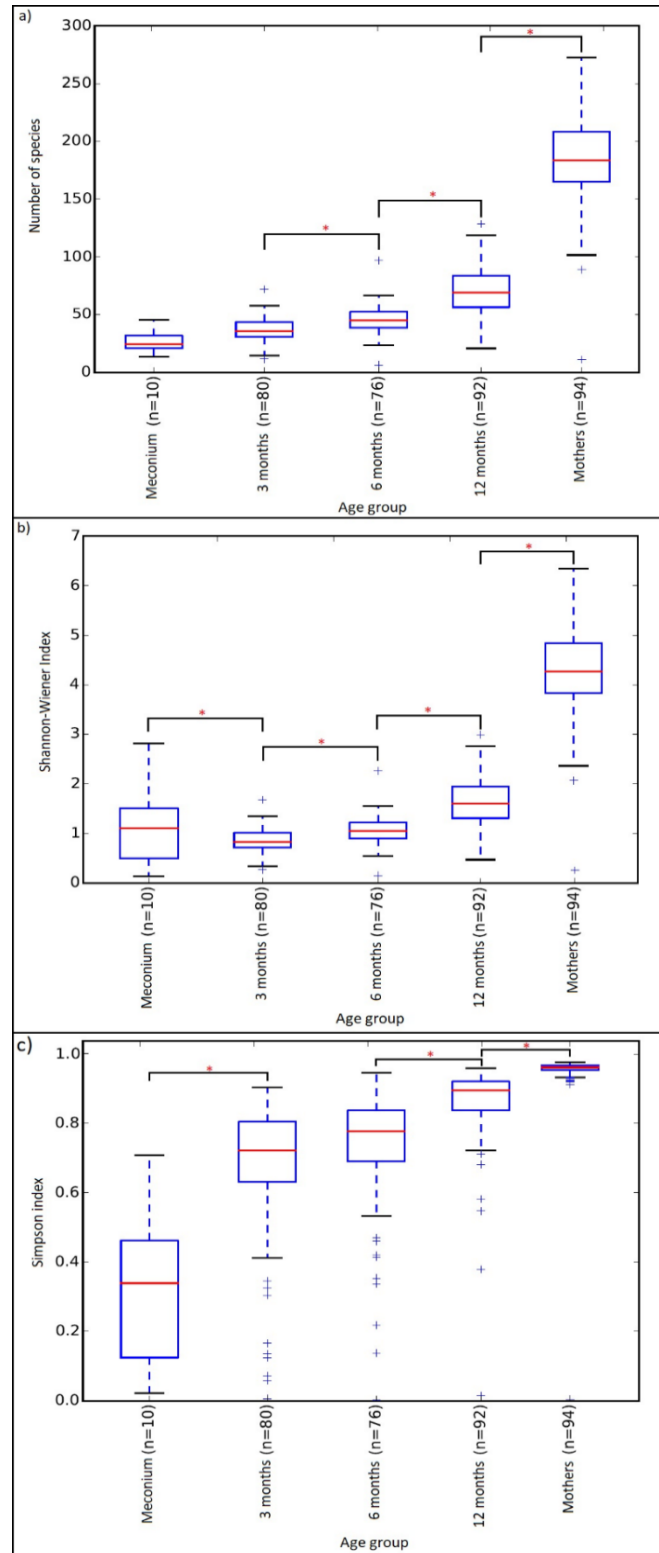
Figure 3.2. Alpha-diversity indexes. The alpha-diversity indexes illustrated are species richness (observed species) in a), Shannon-Wiener index in b) and Simpson index in c). The y-axis for a) shows the number of unique species observed within the age groups, while for b) it shows Shannon-Wiener index and for c) the inverse Simpson's index. A low Shannon-Wiener- & Simpsons-index represents low diversity, with uneven distribution, while a high number represents higher diversity with even distribution. The x-axis shows the age groups, with n=number of samples used to determine the alpha-diversity index. The asterisk represents a p-value between 0.05-0.005.

3.1.3 Beta-Diversity

Diversity between the age groups was determined using Beta-diversity indexes for the communities. The beta-diversity indexes presented in Figure 3.3 derives from the QIIME pipeline. They are presented as Principal Component Analysis (PCoA) plots in figure 3.5 as a) Binary-Jaccard and b) Unweighted Unifrac. Bray-Curtis, Euclidian distance, and Weighted Unifrac indexes are presented in Supplementary Table and Figures, S.1.

The Binary-Jaccard index evaluates similarities between age groups by quantifying species to the relative sum of unique species within the age group (Jaccard, P. 1908). As seen from figure 3.3 a), meconium was scattered, but adjacent to 3 months of age. The 3 months-group was clustered between meconium and 6 months-group, while 6 months was in between the 3 and the 12 months-group. Most children resemble each other more within each age group than between age groups. The age groups gradually increased in similarities to their mothers, from meconium to 12 months. Figure 3.5 b), Unweighted-Unifrac, show the same gradual increase of similarity between the age groups and mothers.

As shown in figure S.1 a), Bray-Curtis (Supplementary Tables and Figures), meconium resembles 3 months of age and is located furthest away from the other age groups. The 3 months and 6 months groups were scattered but intertwined, showing that these age groups resembled each other in number of species shared, relative to the total species in the communities. Mothers and 12 months of age were clustered on the opposite side from the other age groups.

Euclidian distance and Weighted Unifrac did not show any apparent clustering and are presented in figure S.1 b) and c) respectively in Supplementary Tables and Figures.
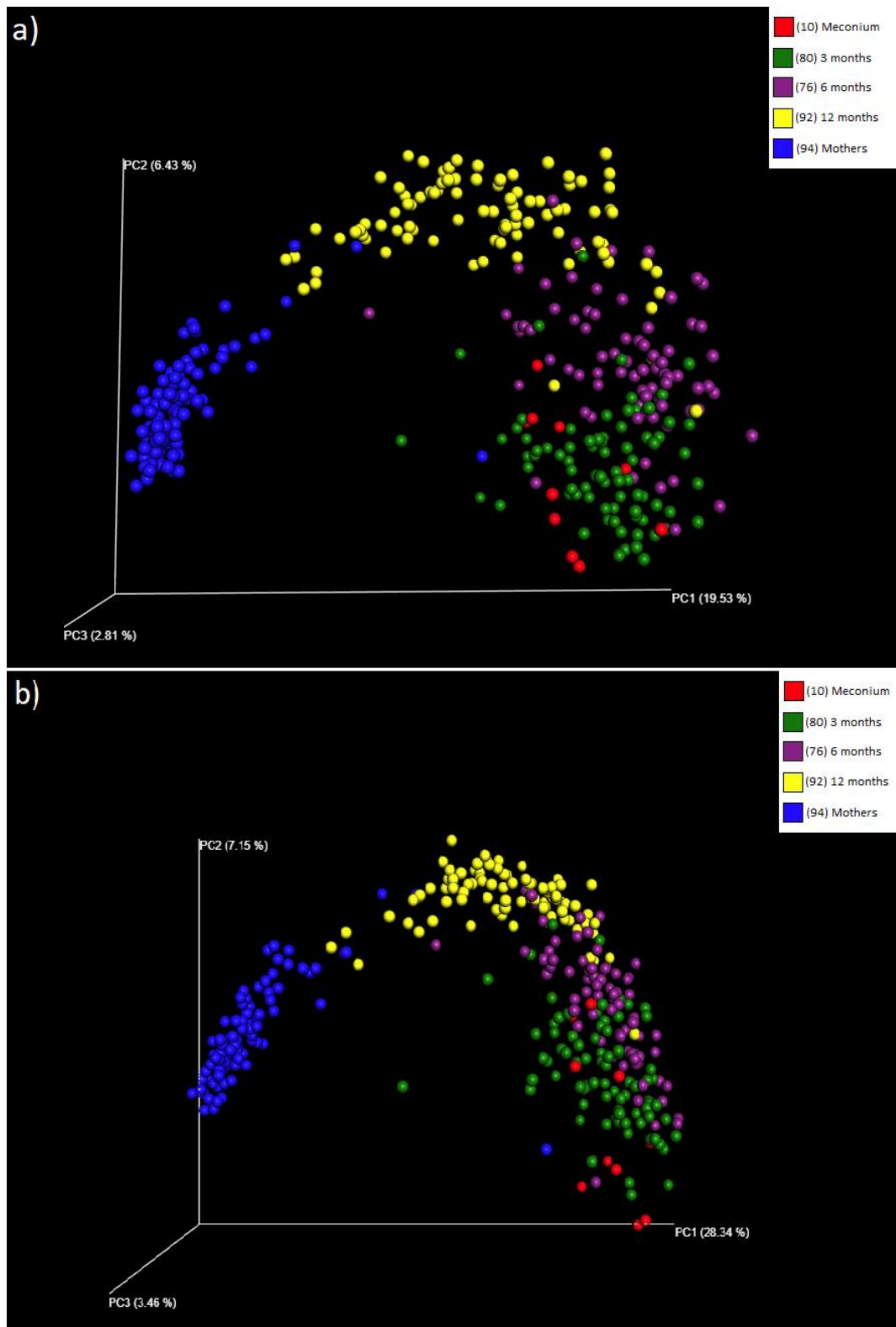
Figure 3.3. Beta-diversity indexes. The figure illustrates the beta-diversity indexes, a) Binary-Jaccard and b) Unweighted Unifrac. Each age group has their respective color shown top right. The sample sizes used for the indexes are shown in parenthesis before their respective colors.

3.2 Short Chain Fatty Acids

3.2.1 SCFAs Profile

The short chain fatty acids profile for each age group was determined using a gas chromatograph (Trace 1310). The SCFA distribution within each age group is presented in Figure 3.4., where a) shows acetate, propionate, butyrate, and others, while b) shows isobutyrate, isovalerate, and valerate.

Acetate was the most dominant SCFA for all the age groups, ranging from 67,42% (12 months) to 90,11% (3 months) of the total SCFAs found within the groups. There was a statistically significant increase in acetate between meconium and 3 months (p=0,004), while there was a significant decrease between 6 months and 12 months of age (p=<2*e^{-16}).

Butyrate was virtually non-existent for the meconium group and first appeared in the 3 months-group. There was a significant increase in butyrate both between the age of 3 months and 6 months(p=2*e^{-8}), and 6 months and 12 months (p=<2*e^{-16}). Butyrate rose more than fourfold between the age of 6 and 12 months, increasing to 18,87% from 4,13% of the total amount of SCFA detected.

Propionate was present throughout the age groups. Propionate significantly increased from 3,1% to 6,7% between 3 and 6 months (p=7*e^{-5}), and significantly increased from 6,7% to 11,1% between 6 and 12 months (p=1,8*e^{-5}).

Meconium's SCFAs profile showed that it was the age group which was the most dissimilar to their mothers. When the child aged, the similarities between mother and child increased, and 12 months of age was the most closely related in SCFAs in comparison to their mothers.

Meconium had the highest abundance of isobutyrate, isovalerate and valerate compared to the total amount of SCFAs within the age group, representing 7,1% of the total SCFA amount found. There was a gradual decrease of isobutyrate and isovalerate from meconium to 6 months of age.

Even though the 12-month age group resembled the mothers regarding the dominant SCFAs, the less abundant SCFAs represents significant differences of the SCFAs profiles, as seen in Figure 3.4 b). The children at 12 months of age had a larger percentage of their total SCFAs as isovalerate and valerate than the mother group did, while the mothers had a higher abundance of isobutyrate relative to their total amount of SCFAs. The isovalerate was the most dominant SCFA for 12 months of age, which corresponded to 5,5% of the total SCFAs found.
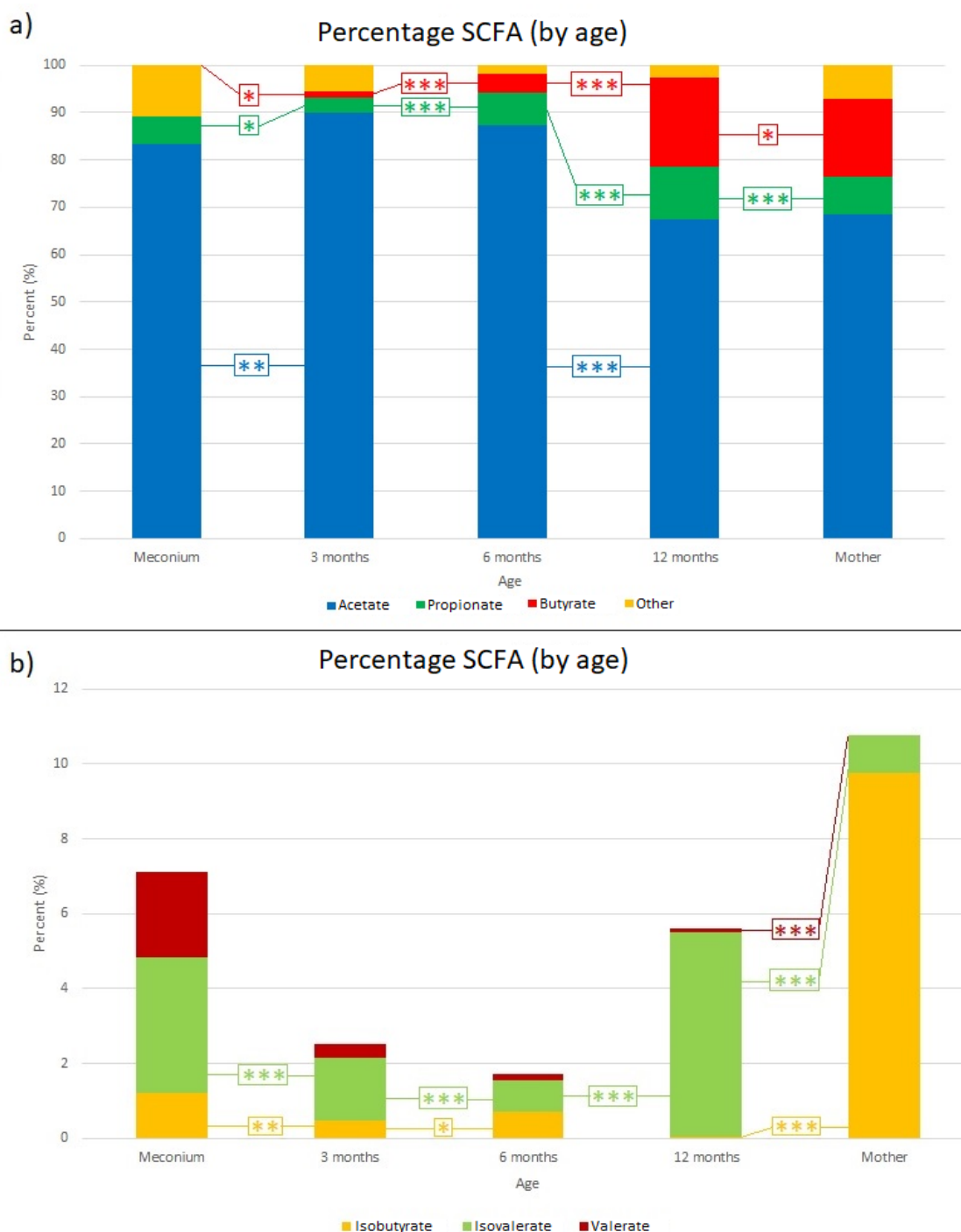
Figure 3.4. Percentage SCFAs grouped by the children's age. The bar chart shows the amount of SCFA in each age group by percent, based on the average total SCFAs. The y-axis shows the percentage of the SCFA, while the x-axis represents the different age groups. 6 a) shows acetate, propionate, butyrate and other SCFAs. The SCFAs presented in 6 b) show isobutyrate, isovalerate, and valerate. The SCFAs profiles are based on n=100 for all age groups except mothers with n=99. P-values between the groups are based on paired t-tests and are shown as asterisks. One asterisk equals a p-value between 0.05 – 0.005, two asterisks represents a p-value between 0.0049 to 0.0005, while three asterisks represents p-value < 0,00049. A data table showing percentage is shown in Supplementary Tables and Figures.

3.2.2 SCFAs, Relative to the Bacterial Load

To correlate the SCFAs to the bacteria from the gut microbiota, the amount of SCFAs compared to the bacterial load from each sample from the age groups was determined. This was done by calculating the average amount (Log10 µM) of SCFAs present per bacteria. Figure 3.5 shows the average amount of SCFAs (Log10 µM) present per 10^3 bacteria.

The highest ratios of SCFAs present per bacteria is found in the meconium age group. Acetate was the SCFA with the highest abundance per bacteria from the gut microbiota for all the age groups analyzed. Propionate is the 2nd most abundant SCFA in 6 months, while butyrate dominates in 12 months. As seen from Figure 3.5, mothers had the lowest ratio between SCFAs present and bacterial load.
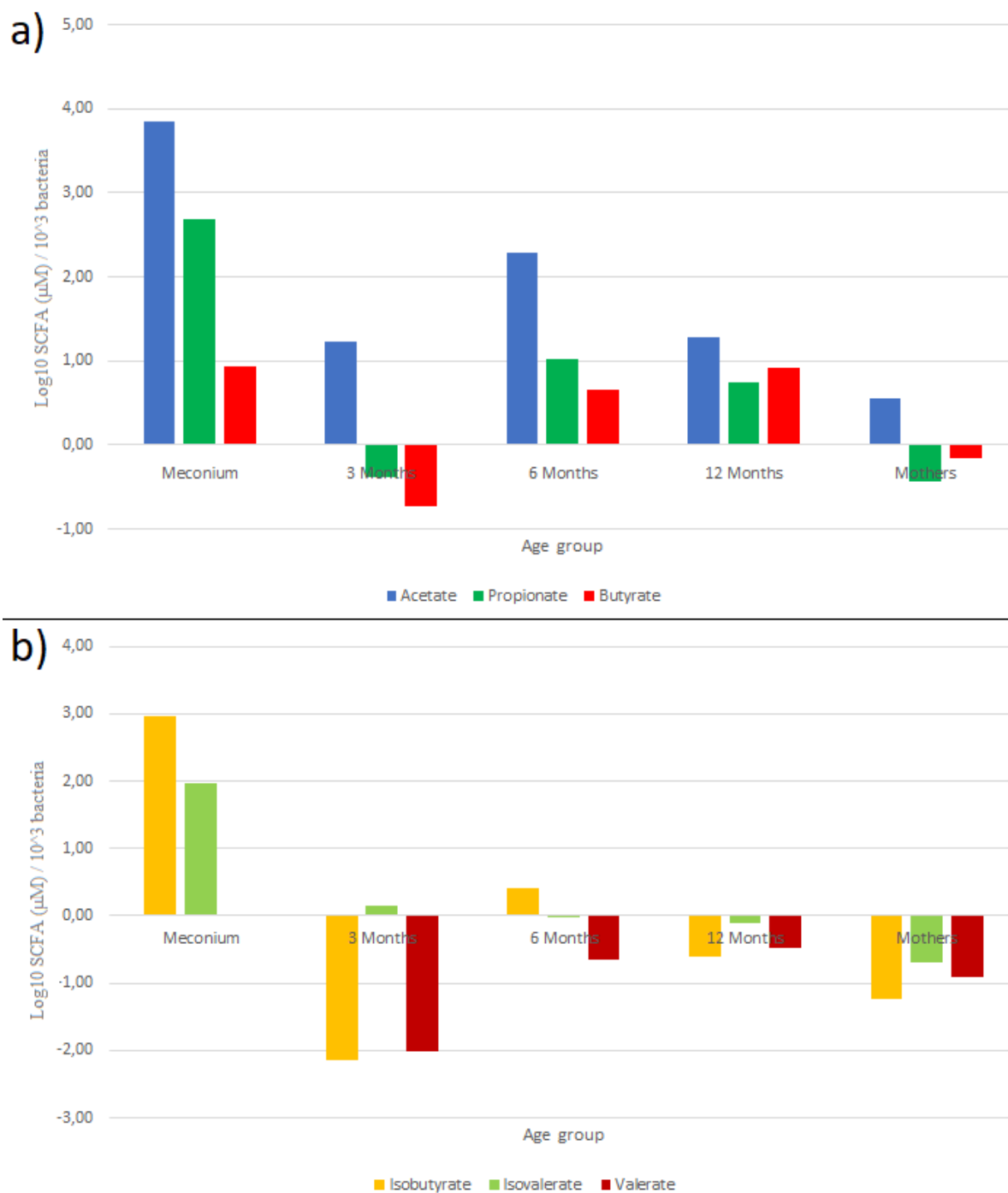
Figure 3.5. SCFAs relative to bacterial load. The illustration shows the average amount of SCFAs found per 1000 bacteria for 6 SCFAs. The y-axis shows Log10 SCFA (μM), while x-axis shows the SCFA of interest with their respective colors. 3.7 a) shows the dominant SCFAs; acetate, propionate, and butyrate, while b) shows isobutyrate, isovalerate, and valerate. The bacterial load for each sample was determined by calculating copy numbers of 16S rRNA genes from Cq-values retrieved from quantitative PCR. One bacterium in the calculations is based on having 4,2 16S rRNA gene copies (Větrovský, T & Baldrians, P. 2013).

3.2.3 Correlation Between Bacterial Orders and SCFAs

Co-occurrence between SCFAs and bacterial orders were analyzed using FDR corrected Spearmann correlations. The correlation pattern between SCFAs and bacterial orders is illustrated in Figure 3.6. The positive correlation is based on co-occurrence where the bacterial order and SCFA are high in quantity, while negative correlation is based on high amounts of the bacterial order with low amounts of the SCFA within the age group.

Some SCFAs had the same positive or negative correlation to a bacterial order in the different age groups. A positive correlation between propionate and *Bacteroidales* was found in both 3 and 12 months of age. Butyrate abundance was positively correlated to the *Clostridium* order for both 6 and 12 months, but *Clostridium* had a negative correlation to propionate in the 12 months-group. Butyrate had a negative correlation to both *Bifidobacteriales* and *Enterobacteriales* at the age of 6 months.

Acetate had a negative correlation to the *Enterobacteriales* order at the age of 3 months, but a positive correlation to the bacterial order at the age of 12 months. Isovalerate had a negative correlation to the *Enterobacteriales* at 3 months.

The only correlation found in the mother group was a negative correlation between valerate and *Burkholderiales*.
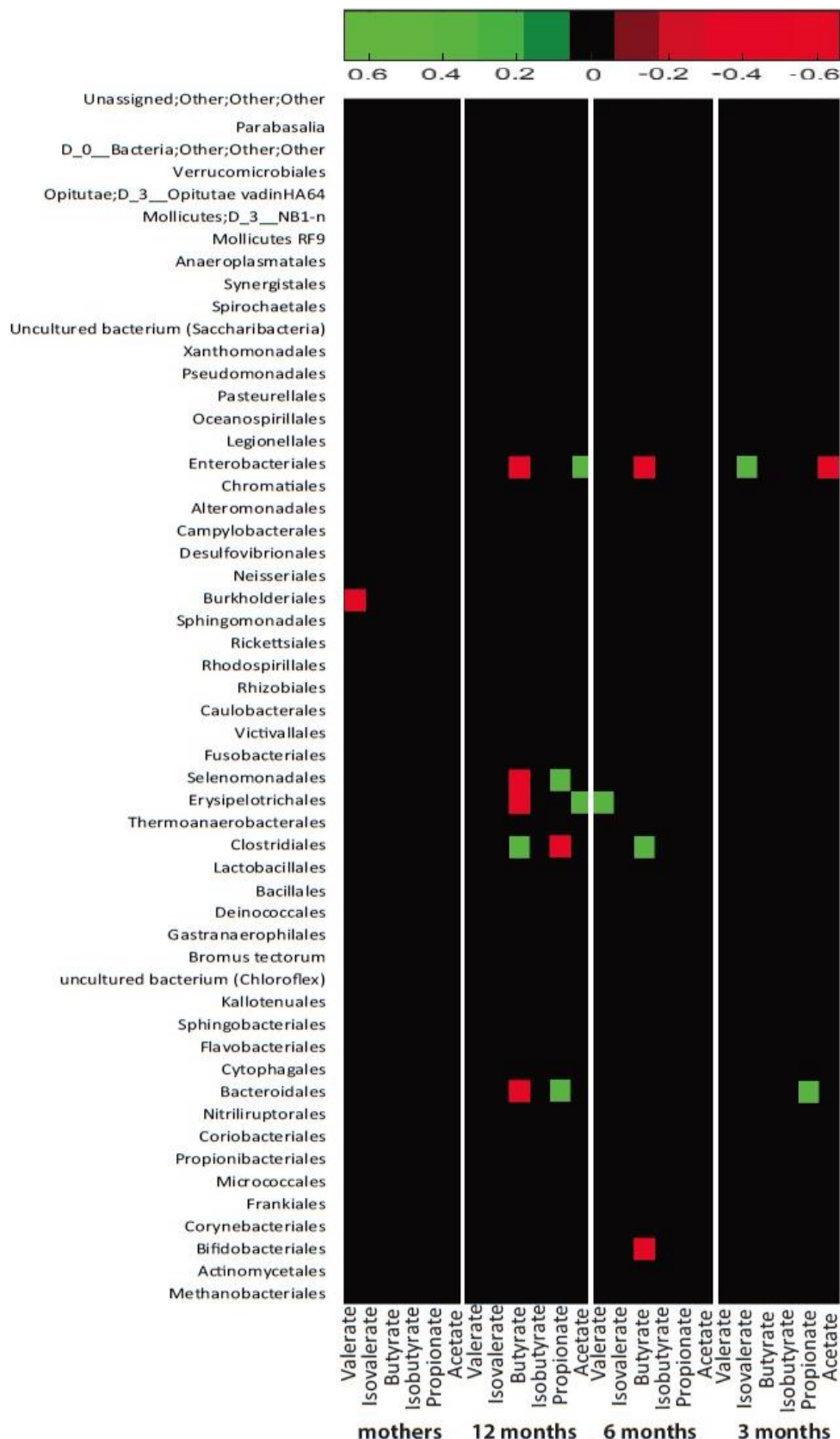
Figure 3.6. SCFA correlation to bacterial orders. The plot shows positive and negative correlations between SCFAs and bacterial orders, divided by age groups. Green indicates a positive correlation (from 0,2 to 0,6), and red indicates a negative correlation (-0,2 to -0,6), while black indicates no correlation. The y-axis consists of bacterial orders retrieved from the QIIME pipeline, while the x-axis represents the SCFAs tested. The plot is based on n=341 samples, distributed among the groups as 3 months n=79, 6 months n=76, 12 months n=94 and mothers n=92. There were too few samples from the meconium group to determine correlations. The figure was made by Knut Rudi and edited by me.

3.2.4 Reproducibility:

To determine variations of SCFA concentrations within each sample, a pilot study on 10 fecal samples from mothers was performed. The fecal samples were run in triplicates with 2-hour intervals between each sample run. The standard deviation for each SCFA is shown in Figure 3.7.

The variation of n=10 for acetate ranges from 5,5 to 43 µM. For propionate, the variation ranges between 2,1 to 22,4 µM. For isobutyrate, the variation is between 0,8 to 6,4 µM, while butyrate ranges from 0,8 to 31,2 µM. Isovalerate ranges between 0,7 to 5,4 µM, while valerate is between 0,3 to 3,1 µM. The largest variations within a sample with regards to percentage variation of the mean concentration is found within propionate and isobutyrate, with one sample each having a variation of 173,2% for both.
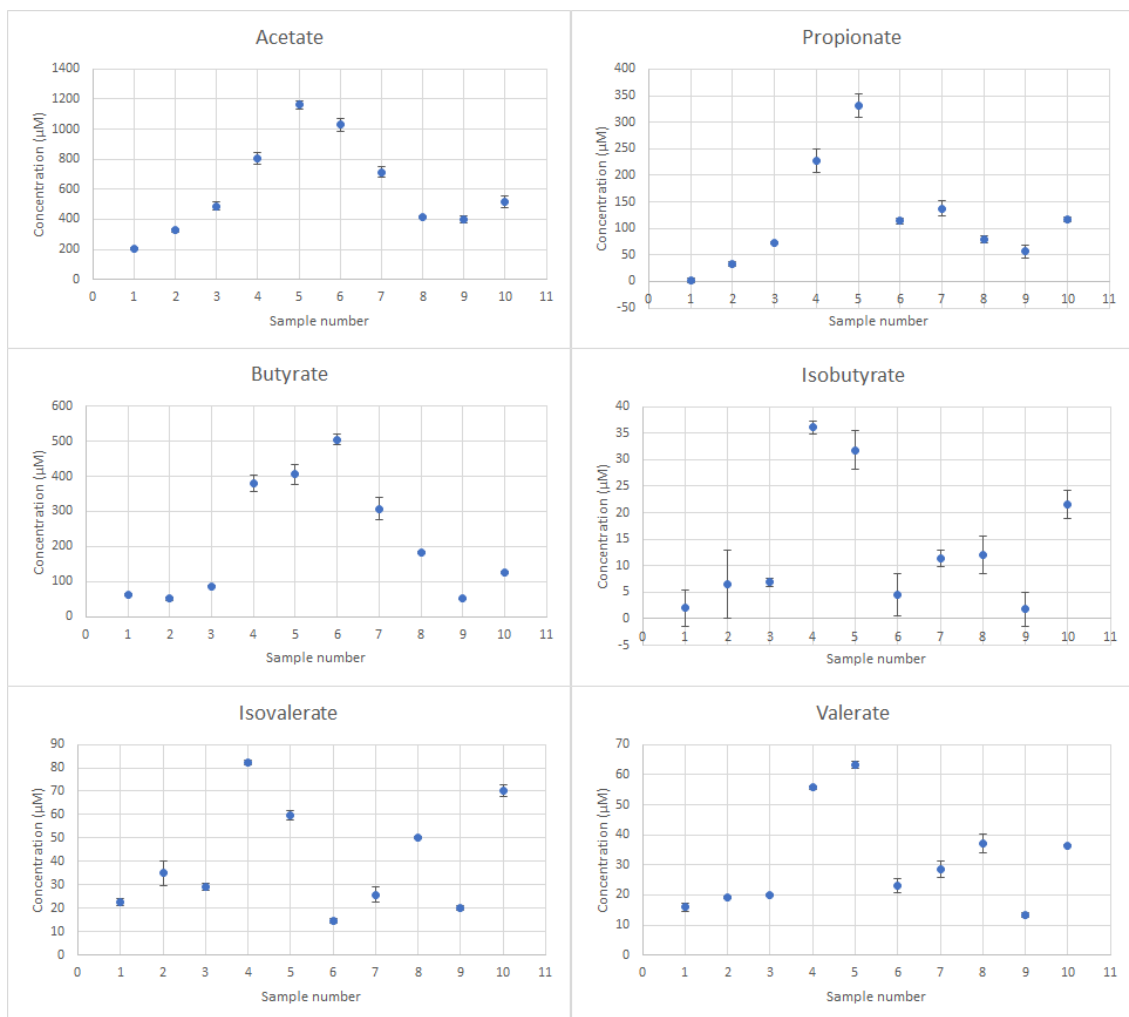
Figure 3.7. Reproducibility. Dot plots showing concentration (µM) of the SCFA on the y-axis, and sample numbers (1-10) on the x-axis. Variations were calculated by standard deviation from the replicate samples and are shown as black lines through the blue dots, representing the range of concentration variation.

4. Discussion:

4.1 High Abundance of *Enterobacteriaceae* in Meconium

In the present study, sequencing revealed that the meconium gut microbiota consisted mostly of *Enterobacteriaceae* and *Firmicutes*. The *Escherichia-Shigella* represented 63% of the total bacteria detected, as seen in Figure C.1 (Appendix C). Earlier studies of meconium found *Firmicutes* and *Bacilli* to dominate the gut microbiota, which differs from the findings in the present study (Moles, L. et al. 2013). The likely explanation for the high abundance of *Escherichia-Shigella* is the sampling time of meconium combined with a fecal transfer route of *E. coli*.

The child's first environmental encounter after birth is the perineum and vagina, which has been shown to harbor an increased number of *E. coli* before parturition and most likely acts as the first link in the transfer route (Bettelheim, K. et al. 1976). The transfer route has been further supported by the detection of high amounts of *E. coli* in the children's oral cavity after birth (Bettelheim, K. et al. 1974). These findings, combined with the low production of gastric acid produced in infants at birth suggests *E. coli*'s possibility of both entering and colonizing the GI tract (Ebers, D. W. et al. 1956; Bettelheim, K. et al. 1976). The meconium samples presented in this study varied within the range of first feces and up to 5 days after birth, while the majority was sampled within the first 24 hours after birth. The amount of *E. coli* was found to be lowest in the feces obtained right after birth (data not shown). Due to the sampling delay after birth, the *E. coli* would have had the chance to colonize the GI tract.

4.2 Meconium's High Ratio Between SCFAs Relative to Bacterial Load

The SCFAs in meconium were dominated by acetate, which represented 83% of the total SCFAs detected. These findings combined with the high abundance of *E. coli* colonizing the gut microbiota might imply that the selection process for facultative and strict anaerobes starts readily after birth. The high abundance of *Enterobacteriaceae* may help deplete the gut for its oxygen, and in addition, produce acetate. Correlations between SCFAs and bacterial orders were not detected due to insufficient DNA extraction from meconium. Acetate is known to be produced by a vast array of enteric bacteria, and the acetate detected in meconium might have originated from *Enterobacteriaceae, Bacteroidales*, *Bifidobacterium* or *Lactobacillus* (Louis, P., Hold, G. L. & Flint, H. J. 2014; den Besten, G. et al. 2013; Fukuda, S. et al. 2012; Pessione E. 2012).

Meconium had the highest ratio between SCFA concentration compared to the bacterial load. The high ratio detected might occur because of infant's short large intestine.
As the SCFAs are readily absorbed throughout the colon, the surface volume and large intestine length might be the factors affecting SCFA absorption (Roy, C. C. et al. 2006). The length of the large intestine has a linear increase compared to children's gestational age (Fitzsimmons, J. & Chinn, A. S. TH. 1988). Since the length and surface volume of the infants are shorter at birth compared to older children, the high ratio of SCFAs compared to bacterial load may be an indication of a lower absorption rate of SCFAs, rather than a high production rate from the bacteria present.

4.3 Significant Increase of Facultative and Strict Anaerobic Bacteria in 3 months

The gut microbiota at 3 months was revealed to mostly consist of *Bifidobacteriales* and *Bacteroidales*, with lower proportions of *Clostridiales* and *Enterobacteriales*. The increase of facultative and strict anaerobic microorganisms early in gut microbiota development has been established by earlier studies (Matamoros, S. et al. 2013). The significant increase of *Bifidobacteriales* and *Bacteroidales* can most likely be explained by *Enterobacteriaceae's* properties of depleting the gut for its oxygen, making the colon a suitable habitat for facultative and strict anaerobic bacteria (Matamoros, S. et al. 2013). The increase of *Bifidobacterium* is probably the result of the diet which affects the gut microbiota colonization. *Bifidobacterium* is well-known to dominate the gut microbiota early in life as a result of the breastmilk's properties, which selects for the bacteria (Fernández, L. et al. 2013; Avershina, E. L. et al. 2016).

The gas chromatography revealed that the SCFAs profile was dominated by acetate, followed by a low proportion of propionate and butyrate. The increase in propionate was positively correlated to *Bacteroidales*, which might confirm earlier observations of *Bacteroidales* being important propionate producers of the gut microbiota (Liou, A. P. et al. 2013). Butyrate, which first appeared in 3-months, did not correlate to a bacterial order within the age group. *Clostridium* is a well-known butyrate producer in the gut microbiota, and the increase of butyrate detected might coincide with the significant increase of *Clostridium* between meconium and 3 months of age, although no correlations were found (den Besten, G. et al. 2013).

4.4 The Significant Increase of *Clostridium* and Propionate in 6 Months

The gut microbiota at the age of 6 months resembled 3 months of age in terms of composition, seen for both taxonomy and beta-diversity. The dominant bacterial order still consisted of *Bifidobacteriales*. The main differences were the significant decrease of *Bacteroidales* and the significant increase of *Clostridium*. Calculations by paired t-tests showed a statistically significant increase in both propionate and butyrate between the age groups. The increase of propionate did not show any apparent positive or negative correlations to bacterial orders in the present study. The *Bacteroidales* positively correlated to propionate in earlier age groups, but not in the 6-month age group, implying that there might be other mechanisms maintaining the increased propionate levels. The significant increase might be explained by the available substrates in the gut microbiota which alters the metabolism of some bacteria.

The presence of *Lactobacilliales* combined with high uptake of breastmilk makes lactate readily available for the bacteria in the gut by cellular respiration or lactic acid fermentation. *Clostridium catus* alter its metabolism based on the presence of lactate. A study discovered that *C. catus* produces both propionate and butyrate, depending on the substrates available (Reichardt, N. et al. 2014). *C. catus* produced propionate in the presence of lactate through the acrylate pathway (Reichardt, N. et al. 2014). Similar mechanisms have been detected for *Roseburia inulinivorans*, which produced propionate in the presence of fucose using the propanediol pathway, but is generally a butyrate producer (Scott, K. P. et al. 2006). *Roseburia* first appeared in the gut microbiota in the 6-months age group in the present study. These and related mechanisms might explain the significant increase of propionate detected between the 3- and 6-month age group.

The positive correlation found between butyrate and *Clostridium* in the 6-month age group might explain the increase of butyrate. In addition, an increase of known butyrate producers such as the *Eubacterium* (rectale group), *Blautia* and *Ruminococcus* might explain the increased butyrate proportion.

4.5 Correlation Between Propionate and Butyrate to Bacterial Order in 12 Months

The bacterial profile for the 12-month age group revealed that the *Clostridiales* order was the most dominant order, followed by *Bacteroidales*. The *Enterobacteriales* and *Bifidobacteriales* decreased significantly between the age of 6 and 12 months. The significant increase of *Clostridium* might be explained by breastmilk weaning and the introduction of solid foods. This combination has been shown in earlier studies to increase the amounts of strict anaerobes, such as the *Clostridium* order (Bäckhed, F. et al. 2015).

Whether it's the introduction of solid food, the cessation of breastfeeding, or a combination of the two which promotes the microbial shift is not yet known. The introduction of solid food gives availability of new fiber sources and other substrates which selects for *Clostridium*, *Ruminococcus,* and *Faecalibacterium*, while ceasing of breastfeeding decreases the abundance of *Bifidobacterium* and *Enterobacteriaceae* (Laforest-Lapointe I & Arrieta M-C. 2017). Evidence suggests that the cessation of breastmilk has a larger impact on the selection than the introduction of solid food (Bäckhed, F. et al. 2015). The microbial shift was associated with the introduction of solid foods, but the shift did not occur until the infants stopped breastfeeding, suggesting that breastmilk weaning might be the causation for the microbial shift, rather than the introduction of solid foods (Bäckhed, F. et al. 2015).

The *Clostridium* order significantly increased and was found to have a positive correlation to butyrate, while the *Bifidobacterium* decreased. The increase of butyrate detected might be a result of the combination of breastmilk weaning, introduction to solid foods and the decrease of *Bifidobacterium*. In earlier studies, the *Bifidobacterium* was negatively correlated to *Clostridium* (Wang, M. et al. 2015). The decrease of *Bifidobacterium*, which most likely is the result of the cessation of breastmilk, might allow *Clostridium* to flourish in the gut microbiota (Wang, M. et al. 2015). This series of events might explain the increase of butyrate, which corresponded to the increase of *Clostridium*.

In the present study, the *Bacteroidales* was not affected by the change in diet, and still represented one of the most dominant bacterial orders even after weaning had commenced, which corresponds to earlier findings (Fallani, M. et al. 2011). The *Bacteroidales* was found to have a positive correlation to propionate detected in the same age group. The high proportion of propionate might be explained by how the *Bacteroidales* order is unaffected by the change in diet and can continue their production of propionate.

4.6 Mothers as a Comparative Group

The SCFAs profile and microbial composition of mothers were used to identify if there were any similarities between the children as they aged, to their mothers. The findings in the present study correspond to earlier experiments where children increase their similarities to their mothers as they age, reaching their adult-like gut microbiota at an age of 2-3 years (Avershina, E. L. et al. 2016; Rodríguez, J. M. et al. 2015). In the present study, the children increase in similarity to their mothers for both microbial diversity and evenness which seem to correspond to the SCFAs detected.

4.7 Diversity Between the Age Groups

Alpha- and Beta-diversity was calculated within and between the age groups. The meconium was found to be composed of low diversity and evenness. This increased gradually as the child aged, becoming more similar to their mothers, which corresponds to earlier findings (Avershina, E. et al. 2016).

The clustering pattern seen from Euclidian distance and Weighted Unifrac did not correspond to previous findings, and it was therefore decided not to analyze these plots and indexes further. No obvious explanation was found for the clustering pattern presented.

4.8 Feces as a Proxy for Determination of SCFAs Production

In the present study, fecal material was used to analyze the SCFAs ratio in different age groups. Feces as material to study SCFAs production is not supported as a representative material and has long been discussed. One likely explanation for this is that the SCFAs found in the fecal material represents those that are not absorbed by the colon rather than the amount produced by the gut microbiota. Because the SCFAs are readily absorbed throughout the colon, less than 5% of the SCFAs produced by the gut microbiota is said to be excreted through feces (Nyangale, E. P. et al. 2012). However, because the SCFAs are readily absorbed at the same rate throughout the colon, the fecal material can be used to represent the ratios of SCFAs (Schmitt, M. G. et al. 1976). New methods for analyzing the production of SCFAs from the gut microbiota need to be developed to further investigate how the quantity of SCFAs is affected by microbial shifts and diets, and how this may affect the children's health early life.

4.9 Technical Considerations

4.9.1 Strengths of the Present Study

One of the strengths of the present study is the large sample-size of mother-children cohort analyzed. The cohorts all have detailed information regarding their delivery, diets and other characteristics throughout their first 3 years of life. Because of the extensive documented information, new projects analyzing several factors of the same children is possible and can then be linked altogether too see how various factors affect the children early life.
The longitudinal sampling made it possible to analyze microbial shifts throughout the first year of life, and to find correlations to SCFAs detected from the samples.

4.9.2 Reproducibility of Gas Chromatography Results

Three replicates of 10 fecal samples derived from mothers were tested in 2-hour intervals to analyze the standard deviation. The standard deviation of the samples had a low variation with few outliers. The 2-methylvaleric acid was used as an internal standard, applied in known concentrations and used as a factor to be able to determine the absolute concentration of the SCFAs in the fecal samples.

4.9.3 The Need for an Optimized Protocol for DNA Extraction in Meconium

The low amounts of DNA extracted from the meconium samples might indicate that the DNA extraction method was not optimal to use on these samples, or that there is, in general, a low abundance of DNA in the meconium. The DNA extraction method used was the same for the rest of the age groups, with the only difference being an increase of PCR cycles. This was to ensure that samples were treated as equal as possible. By using a more optimized protocol for DNA extraction from meconium, the results might have differed. Because of the low detection of bacterial DNA extracted from meconium, an optimized protocol needs to be established for further analyses. A study compared different DNA extraction kits and their efficacy on DNA extraction, and the PM kit was found to be the best kit for this purpose, compared to 3 other kits (Stinson, L. F. et al. 2018). Kits like the PM or optimized protocols on DNA extraction of meconium samples should be considered for DNA extraction of meconium samples.

4.9.4 Unknown Amount of Feces in the Diluted Samples

The fecal samples received from PreventADALL was not weighed, leaving the weight of the fecal samples unknown. As a result, the concentration of bacteria and SCFAs could not be determined.

5. Conclusion and Further Research:

The majority of the microbial compositions within each age group corresponded with earlier studies based on gut microbiota composition. The SCFAs profile significantly differed as the child aged and corresponded with the microbial shifts observed. Like the gut microbiota, the SCFAs profiles in the children increased in similarity to their mothers as they aged. In the present study, the positive correlations between SCFAs and their respective bacterial orders strengthen recent knowledge of these bacteria being SCFAs producers in the gut microbiota.

The significant increase in propionate and butyrate between the age groups of 6 and 12 months was positively correlated to *Bacteroidales* and *Clostridium*, respectively. The positive correlation indicates that a selection of these bacteria can be essential for immune maturation early life. Between 6 and 12 months, the shift from an infant- to an adult-like gut microbiota and SCFAs composition might be initiated and influenced by weaning and introduction to solid foods. In conclusion, this work lays the foundation for further research investigating immunological effects connected to the gut microbiota and their SCFAs.

The human gut microbiota serves important functions, and their substrate production deserves an increased attention. The importance of SCFAs has been established in adults, and the focus should shift towards the gut microbiotas' SCFAs production early life. Further research on this matter may be key to prevent immunological disorders that develop in infants. The increased proportion of butyrate and propionate between 6 and 12 months may indicate an important period for immune maturation. Determining a set of core-species for SCFAs production within this time period, with focus on propionate and butyrate producer, might provide insight in maintaining proper immune development early life.

6. References:

Al-Lahham SaH, Peppelenbosch MP, Roelofsen H, Vonk RJ, Venema K. 2010. Biological effects of propionic acid in humans; metabolism, potential applications and underlying mechanisms. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids 1801:1175-1183.

Anand PAA, Vennison SJ, Sankar SG, Gilwax Prabhu DI, Vasan PT, Raghuraman T, Jerome Geoffrey C, Vendan SE. 2010. Isolation and characterization of bacteria from the gut of Bombyx mori that degrade cellulose, xylan, pectin and starch and their impact on digestion. Journal of Insect Science 10:107-107.

Avershina E, Lundgård K, Sekelja M, Dotterud C, Storrø O, Øien T, Johnsen R, Rudi K. 2016. Transition from infant- to adult-like gut microbiota. Environmental Microbiology 18:2226-2236.

Bakken, L. R., Frostegård, Å. 2006. Nucleic Acid Extraction from Soil. In: Nannipieri, P., Smalla K. (eds) Nucleic Acids and Proteins in Soil. Soil Biology, vol 8. Springer, Berlin, Heidelberg.

Bäckhed F, Roswall J, Peng Y, Feng Q, Jia H, Kovatcheva-Datchary P, Li Y, Xia Y, Xie H, Zhong H, Khan Muhammad T, Zhang J, Li J, Xiao L, Al-Aama J, Zhang D, Lee Ying S, Kotowska D, Colding C, Tremaroli V, Yin Y, Bergman S, Xu X, Madsen L, Kristiansen K, Dahlgren J, Wang J. 2015. Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. Cell Host & Microbe 17:690-703.

Bettelheim K, Breadon A, Faiers MC, O'Farrell SM, Shooter RA. 1974. The origin of O serotypes of Escherichia coli in babies after normal delivery. The Journal of Hygiene 72:67-70.

Bettelheim K, Lennox-King S. 1976. The acquisition of escherichia coli by new-born babies. Infection 4:174-179.

Binder HJ. 2010. Role of Colonic Short-Chain Fatty Acid Transport in Diarrhea. Annual Review of Physiology 72:297-313.

Boom R, Sol CJ, Salimans MM, Jansen CL, Wertheim-van Dillen PM, van der Noordaa J. 1990. Rapid and simple method for purification of nucleic acids. Journal of Clinical Microbiology 28:495-503.

Busing, S.A. 2005. Real-time PCR. Encyclopedia of diagnostic genomics and proteomics. 11: 17-25

Bøyre, Kjell. (2014, 21. November). Enterobacteriaceae. I store norske leksikon. https://snl.no/Enterobacteriaceae.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7:335.

Corrêa-Oliveira R, Fachi JL, Vieira A, Sato FT, Vinolo MAR. 2016. Regulation of immune cell function by short-chain fatty acids. Clinical & Translational Immunology 5:e73.

Cambustion (2018) Fast FID Principles. Cambustion Limited. Cambridge, UK. Found on the World Wide Web: 20.01.18 on: https://www.cambustion.com/products/hfr500/fast-fid-principles

Cullen TW, Schofield WB, Barry NA, Putnam EE, Rundell EA, Trent MS, Degnan PH, Booth CJ, Yu H, Goodman AL. 2015. Antimicrobial peptide resistance mediates resilience of prominent gut commensals during inflammation. Science (New York, NY) 347:170-175.

D'Argenio V, Salvatore F. 2015. The role of the gut microbiome in the healthy adult status. Clinica Chimica Acta 451:97-102.

den Besten G, van Eunen K, Groen AK, Venema K, Reijngoud D-J, Bakker BM. 2013. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. Journal of Lipid Research 54:2325-2340.

DiGiulio DB, Romero R, Amogan HP, Kusanovic JP, Bik EM, Gotsch F, Kim CJ, Erez O, Edwin S, Relman DA. 2008. Microbial Prevalence, Diversity and Abundance in Amniotic Fluid During Preterm Labor: A Molecular and Culture-Based Investigation. PLOS ONE 3:e3056.

Dinarello CA. 1999. IL-18: A TH1 -inducing, proinflammatory cytokine and new member of the IL-1 family. Journal of Allergy and Clinical Immunology 103:11-24.

Dominguez-Bello MG, Costello EK, 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. Proceedings of the National Academy of Sciences 107:11971-11975.

Ebers DW, Smith DI, Gibbs GE. 1956. GASTRIC ACIDITY ON THE FIRST DAY OF LIFE. Pediatrics 18:800-802.

Escobar-Zepeda A, Vera-Ponce de León A, Sanchez-Flores A. 2015. The Road to Metagenomics: From Microbiology to DNA Sequencing Technologies and Bioinformatics. Frontiers in Genetics 6:348.

Fallani M, Amarri S, Uusijarvi A, Adam R, Khanna S, Aguilera M, Gil A, Vieites JM, Norin E, Young D, Scott JA, Doré J, Edwards CA, team tI. 2011. Determinants of the human infant intestinal microbiota after the introduction of first complementary foods in infant samples from five European centres. Microbiology 157:1385-1392.

Fernández L, Langa S, Martín V, Maldonado A, Jiménez E, Martín R, Rodríguez JM. 2013. The human milk microbiota: Origin and potential roles in health and disease. Pharmacological Research 69:1-10.

Fitzsimmons J, Chinn A, Shepard TH. 1988. Normal Length of the Human Fetal Gastrointestinal Tract. Pediatric Pathology 8:633-641.

Flint HJ, Duncan SH, Scott KP, Louis P. 2014. Links between diet, gut microbiota composition and gut metabolism. Proceedings of the Nutrition Society 74:13-22.

Fukuda S, Toh H, Taylor TD, Ohno H, Hattori M. 2012. Acetate-producing bifidobacteria protect the host from enteropathogenic infection via carbohydrate transporters. Gut Microbes 3:449-454.

Funkhouser LJ, Bordenstein SR. 2013. Mom Knows Best: The Universality of Maternal Microbial Transmission. PLoS Biology 11:e1001631.

Grob, R.L & Barry, E. F. (2004) Modern Practice of Gas Chromatography 4th. edition. John Wiley & Sons Inc. Publication

Grönlund MM, Gueimonde M, Laitinen K, Kociubinski G, Grönroos T, Salminen S, Isolauri E. 2007. Maternal breast-milk and intestinal bifidobacteria guide the compositional development of the Bifidobacterium microbiota in infants at risk of allergic disease. Clinical & Experimental Allergy 37:1764-1772.

Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, Bright IJ, Lucero MY, Hiddessen AL, Legler TC, Kitano TK, Hodel MR, Petersen JF, Wyatt PW, Steenblock ER, Shah PH, Bousse LJ, Troup CB, Mellen JC, Wittmann DK, Erndt NG, Cauley TH, Koehler RT, So AP, Dube S, Rose KA, Montesclaros L, Wang S, Stumbo DP, Hodges SP, Romine S, Milanovich FP, White HE, Regan JF, Karlin-Neumann GA, Hindson CM, Saxonov S, Colston BW. 2011. High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. Analytical Chemistry 83:8604-8610.

Huang, H. (2014). QIIME workflow. Found on the world wide web 27.03.2018. https://sites.google.com/site/knightslabwiki/qiime-workflow

Illumina, Inc. 2017. An introduction to Next-Generation Sequencing Technology. Found on the World Wide Web 23.03.2018: https://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf

Jaccard, P. (1908). Nouvelles recherches sur la distribution florale. *Bulletin de la Soci été Vaudoise des Sciences Naturelles* **44**, 223-270

Jiménez E, Fernández L, Marín ML, Martín R, Odriozola JM, Nueno-Palop C, Narbad A, Olivares M, Xaus J, Rodríguez JM. 2005. Isolation of Commensal Bacteria from Umbilical Cord Blood of Healthy Neonates Born by Cesarean Section. Current Microbiology 51:270-274.

Jiménez E, Marin ML, Martin R, Odriozola JM, Olivares M, Xaus J, Fernandez L, Rodriguez JM. 2008. Is meconium from healthy newborns actually sterile? Res Microbiol 159:187-93.

Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, Angenent LT, Ley RE. 2011. Succession of microbial consortia in the developing infant gut microbiome. Proceedings of the National Academy of Sciences of the United States of America 108:4578-4585.

Kondo T, Kishi M, Fushimi T, Kaga T. 2009. Acetic Acid Upregulates the Expression of Genes for Fatty Acid Oxidation Enzymes in Liver To Suppress Body Fat Accumulation. Journal of Agricultural and Food Chemistry 57:5982-5986.

Laforest-Lapointe I, Arrieta M-C. 2017. Patterns of Early-Life Gut Microbial Colonization during Human Immune Development: An Ecological Perspective. Frontiers in Immunology 8:788.

LeBlanc JG, Chain F, Martín R, Bermúdez-Humarán LG, Courau S, Langella P. 2017. Beneficial effects on host energy metabolism of short-chain fatty acids and vitamins produced by commensal and probiotic bacteria. Microbial Cell Factories 16:79.

Linde AG. (2018) Gas Chromatography (GC). Linde Gases Division. Pullach, Germany. Found on the World Wide Web 20.01.18: http://hiq.linde-gas.com/internet.lg.hiq.global/en/images/Application%20sheetHiQ_Gas_Chromatography%28appl%29899_92473.pdf?v=2.0

Linde AG. (2018) Flame ionisation detector (FID). Linde Gases Division. Pullach, Germany. Found on the World Wide Web 20.01.18: http://hiq.linde-gas.com/en/analytical_methods/gas_chromatography/flame_ionisation_detector.html18. 19.

Liou AP, Paziuk M, Luevano J-M, Machineni S, Turnbaugh PJ, Kaplan LM. 2013. Conserved Shifts in the Gut Microbiota Due to Gastric Bypass Reduce Host Weight and Adiposity. Science translational medicine 5:178ra41-178ra41.

Liu L, Li Y, Li S, Hu N, He Y, Pong R, Lin D, Lu L, Law M. 2012. Comparison of Next-Generation Sequencing Systems. Journal of Biomedicine and Biotechnology 2012:11.

Louis P, Hold GL, Flint HJ. 2014. The gut microbiota, bacterial metabolites and colorectal cancer. Nat Rev Microbiol 12:661-72.

Louis P, Young P, Holtrop G, Flint HJ. 2010. Diversity of human colonic butyrate-producing bacteria revealed by analysis of the butyryl-CoA:acetate CoA-transferase gene. Environmental Microbiology 12:304-314.

Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, Knight R. 2012. Diversity, stability and resilience of the human gut microbiota. Nature 489:220-230.

Macfarlane S, Macfarlane GT. 2007. Regulation of short-chain fatty acid production. Proceedings of the Nutrition Society 62:67-72.

Matamoros S, Gras-Leguen C, Le Vacon F, Potel G, de La Cochetiere M-F. 2013. Development of intestinal microbiota in infants and its impact on health. Trends in Microbiology 21:167-173.

Mathews, C. K., Holde, K. E. V., Appling, D. R., Anthony-Cahill, S. J. (2013). Biochemistry, 4th edition. Pearson, Toronto.

Maxam AM, Gilbert W. 1977. A new method for sequencing DNA. Proceedings of the National Academy of Sciences 74:560-564.

Mikheyev AS, Tin MMY. 2014. A first look at the Oxford Nanopore MinION sequencer. Molecular Ecology Resources 14:1097-1102.

Miller TL, Wolin MJ. 1996. Pathways of acetate, propionate, and butyrate formation by the human fecal microbial flora. Applied and Environmental Microbiology 62:1589-1592.

Moles L, Gómez M, Heilig H, Bustos G, Fuentes S, de Vos W, Fernández L, Rodríguez JM, Jiménez E. 2013. Bacterial Diversity in Meconium of Preterm Neonates and Evolution of Their Fecal Microbiota during the First Month of Life. PLoS ONE 8:e66986.

Morrison DJ, Preston T. 2016. Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. Gut Microbes 7:189-200.

Morgan, P. 2012. Gas Chromatography Liner Selection Guide. Termo Fisher Scientific. Chesire, UK. Found on the world wide web 15.02.2018:
http://www.separatedbyexperience.com/documents/Liner_Selection_Guide.pdf

Nyangale EP, Mottram DS, Gibson GR. 2012. Gut Microbial Activity, Implications for Health and Disease: The Potential Role of Metabolite Analysis. Journal of Proteome Research 11:5573-5585.

Pacific Biosciences of California. (2018). Found on the world wide web at: www.pacb.com

Pessione E. 2012. Lactic acid bacteria contribution to gut microbiota complexity: lights and shadows. Frontiers in Cellular and Infection Microbiology 2:86.

Primec M, Mičetić-Turk D, Langerholc T. 2017. Analysis of short-chain fatty acids in human feces: A scoping review. Analytical Biochemistry 526:9-21

Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glöckner FO. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Research 35:7188-7196.

Reichardt N, Duncan SH, Young P, Belenguer A, McWilliam Leitch C, Scott KP, Flint HJ, Louis P. 2014. Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. The ISME Journal 8:1323-1335.

Ríos-Covián D, Ruas-Madiedo P, Margolles A, Gueimonde M, de los Reyes-Gavilán CG, Salazar N. 2016. Intestinal Short Chain Fatty Acids and their Link with Diet and Human Health. Frontiers in Microbiology 7:185.

Rodríguez JM, Murphy K, Stanton C, Ross RP, Kober OI, Juge N, Avershina E, Rudi K, Narbad A, Jenmalm MC, Marchesi JR, Collado MC. 2015. The composition of the gut microbiota throughout life, with an emphasis on early life. Microbial Ecology in Health and Disease 26:10.3402/mehd.v26.26050.

Roy CC, Kien CL, Bouthillier L, Levy E. 2006. Short-Chain Fatty Acids: Ready for Prime Time? Nutrition in Clinical Practice 21:351-366.

Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America 74:5463-5467.

Satokari R, Gronroos T, Laitinen K, Salminen S, Isolauri E. 2009. Bifidobacterium and Lactobacillus DNA in the human placenta. Lett Appl Microbiol 48:8-12.

Schmitt MG, Jr., Soergel KH, Wood CM. 1976. Absorption of short chain fatty acids from the human jejunum. Gastroenterology 70:211-5.

Schochetman G, Ou C-Y, Jones WK. 1988. Polymerase Chain Reaction. The Journal of Infectious Diseases 158:1154-1157.

Scott KP, Martin JC, Campbell G, Mayer C-D, Flint HJ. 2006. Whole-Genome Transcription Profiling Reveals Genes Up-Regulated by Growth on Fucose in the Human Gut Bacterium "Roseburia inulinivorans". Journal of Bacteriology 188:4340-4349.

Shannon, C.E & Weaver, W. (1949). *The mathematical theory of communication*. The University of Illinois Press, Urbana, 117p.

Stinson LF, Keelan JA, Payne MS. 2018. Comparison of Meconium DNA Extraction Methods for Use in Microbiome Studies. Frontiers in Microbiology 9:270.

Tangerman A, Nagengast FM. 1996. A Gas Chromatographic Analysis of Fecal Short-Chain Fatty Acids, Using the Direct Injection Method. Analytical Biochemistry 236:1-8.

Termo Fisher Scientific. Proteinase K Solution (20 mg/mL). Found on the world wide web 16.02.2018: https://www.thermofisher.com/order/catalog/product/AM2548

Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett C, Knight R, Gordon JI. 2007. The human microbiome project: exploring the microbial part of ourselves in a changing world. Nature 449:804-810.

Velázquez M, Davies C, Marett R, Slavin JL, Feirtag JM. 2000. Effect of Oligosaccharides and Fibre Substitutes on Short-chain Fatty Acid Production by Human Faecal Microflora. Anaerobe 6:87-92.

Větrovský T, Baldrian P. 2013. The Variability of the 16S rRNA Gene in Bacterial Genomes and Its Consequences for Bacterial Community Analyses. PLOS ONE 8:e57923

Waksmonski, M. 2015. Short Chain Fatty Acid Headspace Analysis Optimization for GC Using Alcohol Esterication. Shimadzu Scientific Instruments, Columbia, USA. Found on the World Wide Web 16.02.2018: https://solutions.shimadzu.co.jp/an/n/en/gcms/sio215019.pdf .

Wong, Julia M. W. RD; de Souza, Russell RD; Kendall, Cyril W. C. PhD; Emam, Azadeh MSc; Jenkins, David J. A. MD. 2016. Colonic Health: Fermentation and Short Chain Fatty Acids. Journal of Clinical Gastroenterology vol (40:3) pp. 235-243.

Woting A, Blaut M. 2016. The Intestinal Microbiota in Metabolic Disease. Nutrients 8:202.

Wang M, Li M, Wu S, Lebrilla CB, Chapkin RS, Ivanov I, Donovan SM. 2015. Fecal Microbiota Composition of Breast-fed Infants is Correlated with Human Milk Oligosaccharides Consumed. Journal of pediatric gastroenterology and nutrition 60:825-833.

Yu, Y., Lee, C., Kim, J. and Hwang, S. 2005. Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnology and bioengineering*, **89**, 670-679.)

Xu J, Gordon JI. 2003. Honor thy symbionts. Proceedings of the National Academy of Sciences 100:10452-10459.

Supplementary Tables & Figures:
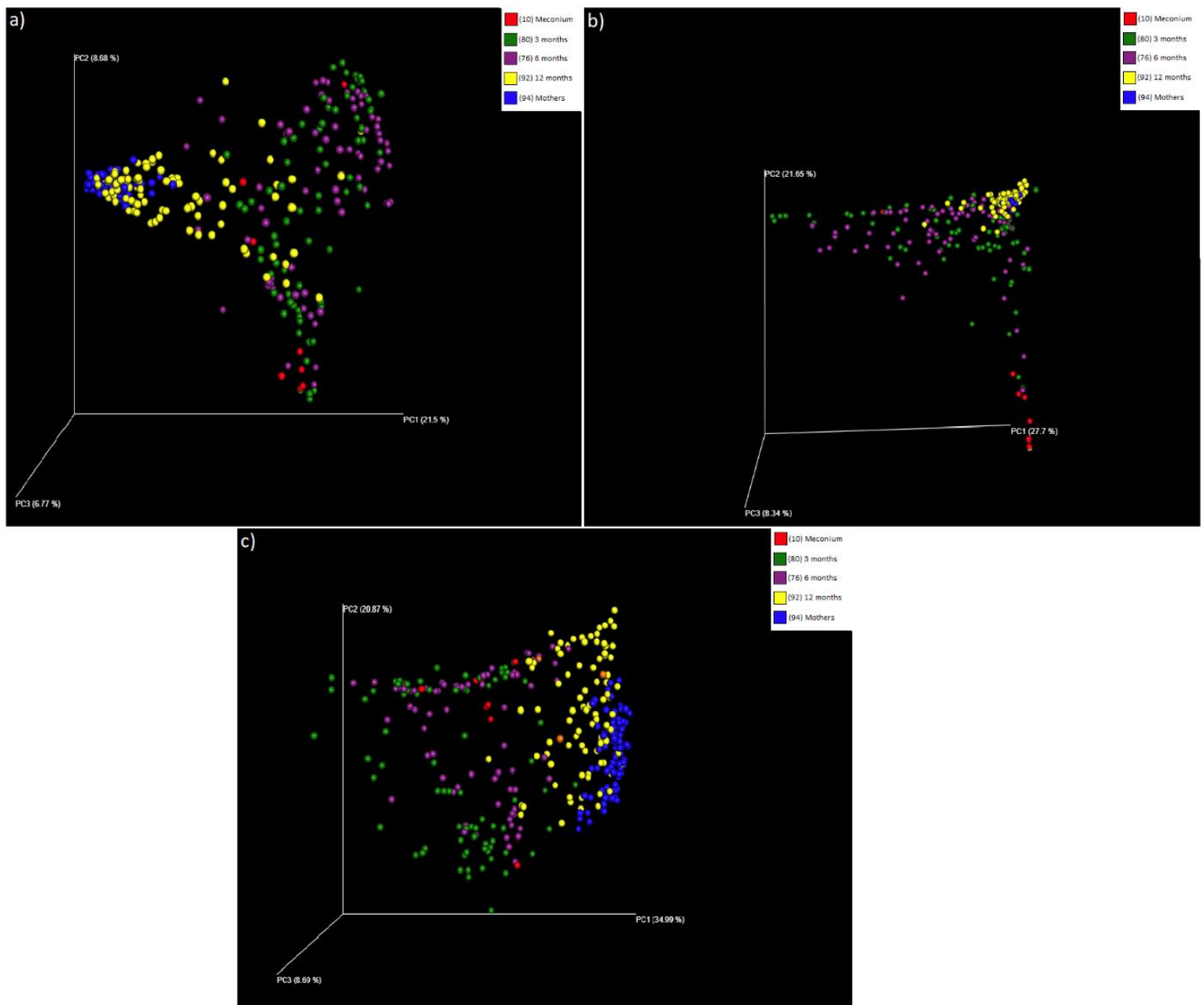
Figure S.1. Beta-diversity indexes. Illustration of a) Bray-Curtis, b) Euclidian distance and c) Weighted-Unifrac. Each age group is represented by a color, showing top right with sample sizes in parenthesis.

Table S.1: Amount of bacterial orders shown in percent. The proportion of bacteria is shown in percent within each age group, acquired from the QIIME pipeline. The table shows the number of the dominant orders of bacteria retrieved from sequencing.

Age group

Bacterial order	Meconium:	3 months:	6 months:	12 months:	Mothers:
Enterobacteriales	63,5 %	14,1 %	12,7 %	2,2 %	1,2 %
Clostridiales	5,7 %	16,2 %	25,7 %	66,6 %	60,1 %
Bacteroidales	10,2 %	23,5 %	14,3 %	15,8 %	26,7 %
Bifidobacteriales	5,5 %	33,1 %	34,1 %	7,8 %	2,8 %
Lactobacilliales	7,2 %	3,2 %	3,6 %	1,7 %	0,8 %

Table S.2. SCFA composition in percent. The table shows the proportion of SCFA in percent for the three most dominant SCFA, while the other group includes isovalerate, isobutyrate and valerate.

Age group

SCFA:	Meconium	3 months	6 months	12 months	Mother
Acetate	83,46 %	90,11 %	87,38 %	67,42 %	68,60 %
Propionate	5,69 %	3,09 %	6,76 %	11,19 %	7,93 %
Butyrate	0,08 %	1,18 %	4,13 %	18,87 %	16,33 %
Other	10,77 %	5,62 %	1,72 %	2,52 %	7,13 %

Table S.3. p-values of SCFAs between age groups. The table gives an overview of the different p-values obtained of a paired t-test between SCFA proportion between the age groups.

Age		Short chain fatty acid					
Group 1	Group 2	Acetate	Propionate	Butyrate	Isobutyrate	Isovalerate	Valerate
Meconium	3 months	0.004189*	0.0256*	0.01463*	0.002413*	1.0e-5*	0.05529
3 months	6 months	0.1626	7.3e-5*	2.8e-8*	0.04403*	1.0e-5	0.09082
6 months	12 months	< 2.2e-16*	1.8e-5*	< 2.2e-16*	0.1563	1.8e-4*	0.7108
12 months	Mothers	0.1319	6.0e-4*	0.03499*	2.2e-6*	5.7e-11*	< 2.2e-16*

*P-value = 95% significance

Table S.4. p-values for bacterial orders between age groups. The table gives an overview of the different p-values obtained between the age groups* for the dominant bacterial orders.

Age		Bacterial order				
Group 1	Group 2	Enterobacteriales	Bacteroidales	Bifidobacteriales	Clostridiales	Lactobacilliales
Meconium	3 months	0.004814*	0.0727	0.008855*	0.4036	0.4615
3 months	6 months	0.4566	0.01319*	0.3906	5.3e-4*	0.8806
6 months	12 months	2.5e-91*	0.4691	2.074e-14*	< 2.2e-16*	0.1591
12 months	Mothers	3.7e-4*	1.29e-11*	1.0e-8*	0.003081*	0.05371

*p-value = 95% significance

Appendix A: Gas Chromatograph Specifications

Instrument: Trace 1310 with autosampler (ThemoFisher Scientific)

Injector:

Mode: split

Temperature: 250°C

Carrier gas: Helium

Column flow: 2.5 ml/min

Split flow: 200 ml/min

Purge flow: 3 ml/min

Injection volume: 0.2µl

Liner: 4mm x 6.3mm x 78.5mm (Catalog# 233115., Restek)

Syringe: 10µl syr FN 50 mm C, Ga 23, cone tip (catalog# 365D3741, ThermoFisher Scientific)

Column:

Stabilwax DA 30m, 0.25mm ID, 0.25µM (Restek)

Temperature program: 90°C to 150°C (6 minutes), 150°C to 245°C (1.9 minutes)

Time per sample: 14.9 minutes

Detector:

Type: FID

Temperature: 275°C

Hydrogen: 30 ml/min

Air: 300 ml/min

Makeup gas: 30 ml/min

Software Analyzer: Chromeleon 7

Appendix B: Primer sequences

PRK Illumina primer sequences for Index PCR:

PRKi forward (5'-3'):

1. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctagtcaaCCTACGGGRBGCASCAG

2. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctagttccCCTACGGGRBGCASCAG

3. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctatgtcaCCTACGGGRBGCASCAG

4. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctccgtccCCTACGGGRBGCASCAG

5. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctgtagagCCTACGGGRBGCASCAG

6. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctgtccgcCCTACGGGRBGCASCAG

7. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctgtgaaaCCTACGGGRBGCASCAG

8. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctgtggccCCTACGGGRBGCASCAG

9. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctgtttcgCCTACGGGRBGCASCAG

10. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctcgtacgCCTACGGGRBGCASCAG

11. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctgagtggCCTACGGGRBGCASCAG

12. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctggtagcCCTACGGGRBGCASCAG

13. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctactgatCCTACGGGRBGCASCAG

14. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctatgagcCCTACGGGRBGCASCAG

15. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctattcctCCTACGGGRBGCASCAG

16. aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctcaaaagCCTACGGGRBGCASCAG

PRKi Reverse (5' – 3'):

1. caagcagaagacggcatacgagatCGTGATgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

2. caagcagaagacggcatacgagatACATCGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

3. caagcagaagacggcatacgagatGCCTAAgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

4. caagcagaagacggcatacgagatTGGTCAgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

5. caagcagaagacggcatacgagatCACTCTgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

6. caagcagaagacggcatacgagatATTGGCgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

7. caagcagaagacggcatacgagatGATCTGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

8. caagcagaagacggcatacgagatTCAAGTgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

9. caagcagaagacggcatacgagatCTGATCgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

10. caagcagaagacggcatacgagatAAGCTAgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

11. caagcagaagacggcatacgagatGTAGCCgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

12. caagcagaagacggcatacgagatTACAAGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

13. caagcagaagacggcatacgagatTTGACTgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

14. caagcagaagacggcatacgagatGGAACTgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

15. caagcagaagacggcatacgagatTGACATgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

16. caagcagaagacggcatacgagatGGACGGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

17. caagcagaagacggcatacgagatCTCTACgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

18. caagcagaagacggcatacgagatGCGGACgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

19. caagcagaagacggcatacgagatTTTCACgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

20. caagcagaagacggcatacgagatGGCCACgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

21. caagcagaagacggcatacgagatCGAAACgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

22. caagcagaagacggcatacgagatCGTACGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

23. caagcagaagacggcatacgagatCCACTCgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

24. caagcagaagacggcatacgagatGCTACCgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

25. caagcagaagacggcatacgagatATCAGTgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

26. caagcagaagacggcatacgagatGCTCATgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

27. caagcagaagacggcatacgagatAGGAATgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

28. caagcagaagacggcatacgagatCTTTTGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

29. caagcagaagacggcatacgagatTAGTTGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

30. caagcagaagacggcatacgagatCCGGTGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

PRK Primers for amplification of 16S rRNA (5' – 3'):

Forward: CCTACGGGRBGCASCAG

Reverse: GGACTACYVGGGTATCTAAT

Appendix C: Species Sequenced
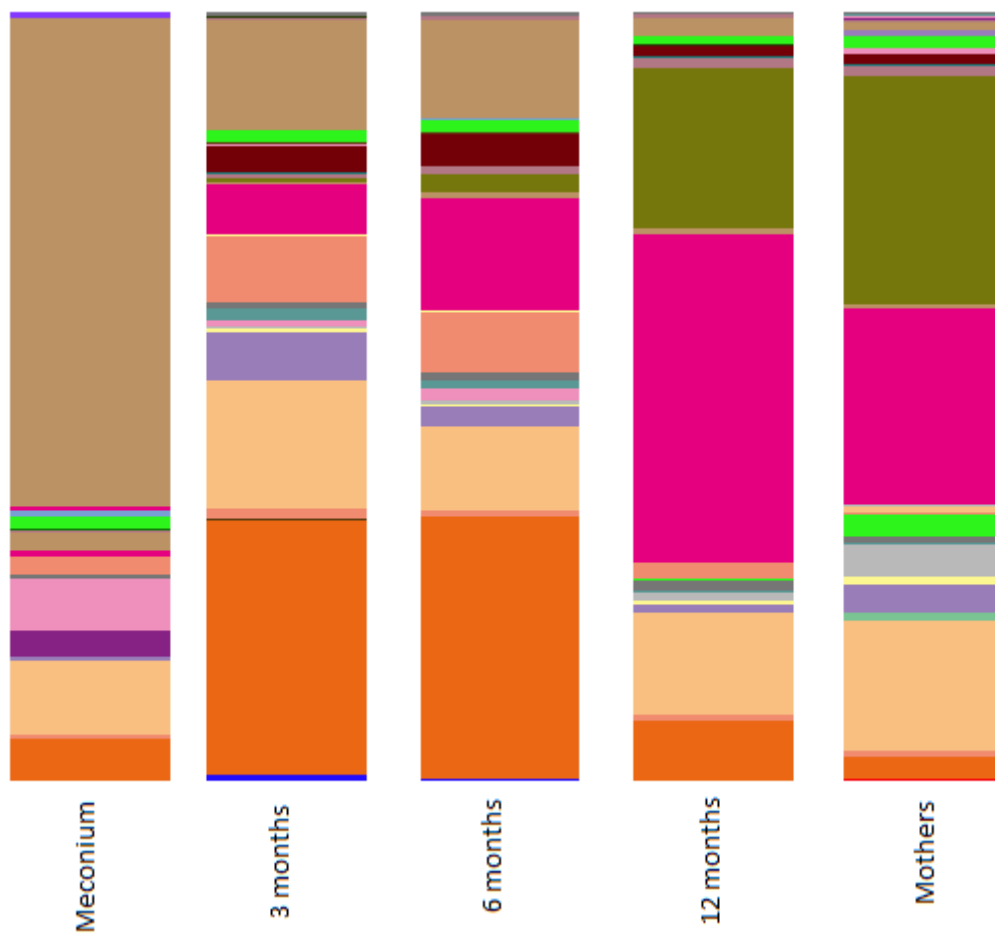
Figure C.1. Bacterial composition at family level. The bar chart shows the number of bacteria for the different age groups at family level. The families are divided by color and are explained in Table C.1.

Table C.1. Bacterial composition at family level in percent. The table shows the abundance (%) of the bacterial families within each age group retrieved from sequencing and processed by the QIIME pipeline.

| Legend	Taxonomy	Total %	Meconium %	3 months %	6 months %	12 months %	Mothers %
	D_0__Archaea;D_1__Euryarchaeota;D_2__Methanobacteria;D_3__Methanobacteriales;D_4__Methanobacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.3%
	D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Actinomycetales;D_4__Actinomycetaceae	0.2%	0.0%	0.9%	0.2%	0.0%	0.0%
	D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Bifidobacteriales;D_4__Bifidobacteriaceae	15.3%	5.5%	33.1%	34.1%	7.8%	2.8%
	D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Corynebacteriales;D_4__Corynebacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Corynebacteriales;D_4__Nocardiaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Frankiales;D_4__Nakamurellaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Micrococcales;D_4__Dermabacteraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Micrococcales;D_4__Dermacoccaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Micrococcales;D_4__Microbacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Micrococcales;D_4__Micrococcaceae	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Micrococcales;D_4__Promicromonosporaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Actinobacteria;D_2__Actinobacteria;D_3__Propionibacteriales;D_4__Propionibacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Actinobacteria;D_2__Coriobacteriia;D_3__Coriobacteriales;D_4__Coriobacteriaceae	0.8%	0.3%	1.5%	0.9%	0.7%	0.9%
	D_0__Bacteria;D_1__Actinobacteria;D_2__Nitriliruptoria;D_3__Nitriliruptorales;D_4__Nitriliruptoraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Bacteroidaceae	12.0%	9.6%	16.6%	10.9%	13.2%	16.8%
	D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Bacteroidales S24-7 group	0.2%	0.0%	0.0%	0.0%	0.0%	1.0%
	D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Porphyromonadaceae	2.6%	0.6%	6.3%	2.6%	1.2%	3.8%
	D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Prevotellaceae	0.4%	0.0%	0.5%	0.2%	0.4%	1.1%
	D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__Rikenellaceae	1.1%	0.0%	0.2%	0.5%	1.0%	4.1%
	D_0__Bacteria;D_1__Bacteroidetes;D_2__Bacteroidia;D_3__Bacteroidales;D_4__uncultured	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Bacteroidetes;D_2__Cytophagia;D_3__Cytophagales;D_4__Cytophagaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Bacteroidetes;D_2__Flavobacteriia;D_3__Flavobacteriales;D_4__Flavobacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Bacteroidetes;D_2__Flavobacteriia;D_3__Flavobacteriales;D_4__NS9 marine group	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Bacteroidetes;D_2__Sphingobacteriia;D_3__Sphingobacteriales;D_4__Chitinophagaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Chloroflexi;D_2__Chloroflexia;D_3__Kallotenuales;D_4__Kallotenuaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Chloroflexi;D_2__KD4-96;D_3__uncultured bacterium;D_4__uncultured bacterium	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Cyanobacteria;D_2__Chloroplast;D_3__Bromus tectorum;D_4__Bromus tectorum	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Cyanobacteria;D_2__Melainabacteria;D_3__Gastranaerophilales;D_4__uncultured bacterium	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Cyanobacteria;D_2__Melainabacteria;D_3__Gastranaerophilales;Other	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Deinococcus-Thermus;D_2__Deinococci;D_3__Deinococcales;D_4__Deinococcaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Bacillales;D_4__Bacillaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Bacillales;D_4__Family XI	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Bacillales;D_4__Paenibacillaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Bacillales;D_4__Staphylococcaceae	0.6%	3.4%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Lactobacillales;D_4__Aerococcaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Lactobacillales;D_4__Carnobacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Lactobacillales;D_4__Enterococcaceae	1.5%	6.8%	0.9%	1.4%	0.1%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Lactobacillales;D_4__Lactobacillaceae	0.5%	0.0%	1.5%	1.2%	0.1%	0.1%
	D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Lactobacillales;D_4__Leuconostocaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Bacilli;D_3__Lactobacillales;D_4__Streptococcaceae	0.8%	0.3%	0.8%	0.9%	1.5%	0.8%
	D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Christensenellaceae	0.5%	0.0%	0.0%	0.0%	0.0%	2.9%
	D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Clostridiaceae 1	5.9%	2.5%	8.6%	7.9%	2.3%	0.2%
	D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Clostridiales Incertae Sedis	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Clostridiales vadinBB60 group	0.1%	0.0%	0.0%	0.0%	0.0%	0.8%
	D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Defluviitaleaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Eubacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Family XI	0.0%	0.0%	0.1%	0.1%	0.0%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Family XIII	0.1%	0.0%	0.0%	0.1%	0.0%	0.2%
	D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Lachnospiraceae	21.5%	0.6%	6.5%	14.7%	42.6%	25.5%
	D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Peptococcaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
	D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Peptostreptococcaceae	0.9%	2.4%	0.4%	0.8%	0.8%	0.3%
	D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Clostridiales;D_4__Ruminococcaceae	11.0%	0.1%	0.7%	2.2%	20.8%	29.9%
	D_0__Bacteria;D_1__Firmicutes;D_2__Clostridia;D_3__Thermoanaerobacterales;D_4__Thermoanaerobacteraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Firmicutes;D_2__Erysipelotrichia;D_3__Erysipelotrichales;D_4__Erysipelotrichaceae	0.8%	0.2%	0.5%	1.1%	1.4%	1.1%
	D_0__Bacteria;D_1__Firmicutes;D_2__Negativicutes;D_3__Selenomonadales;D_4__Acidaminococcaceae	0.1%	0.0%	0.1%	0.1%	0.1%	0.3%
	D_0__Bacteria;D_1__Firmicutes;D_2__Negativicutes;D_3__Selenomonadales;D_4__Veillonellaceae	2.1%	0.0%	3.4%	4.1%	1.5%	1.4%
	D_0__Bacteria;D_1__Fusobacteria;D_2__Fusobacteriia;D_3__Fusobacteriales;D_4__Fusobacteriaceae	0.0%	0.0%	0.0%	0.0%	0.1%	0.0%
	D_0__Bacteria;D_1__Fusobacteria;D_2__Fusobacteriia;D_3__Fusobacteriales;D_4__Leptotrichiaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Lentisphaerae;D_2__Lentisphaeria;D_3__Victivallales;D_4__Victivallaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Lentisphaerae;D_2__Lentisphaeria;D_3__Victivallales;D_4__vadinBE97	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Alphaproteobacteria;D_3__Caulobacterales;D_4__Caulobacteraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Alphaproteobacteria;D_3__Rhizobiales;D_4__Bradyrhizobiaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Alphaproteobacteria;D_3__Rhizobiales;D_4__Methylobacteriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Alphaproteobacteria;D_3__Rhizobiales;D_4__Rhizobiaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Alphaproteobacteria;D_3__Rhizobiales;D_4__Xanthobacteraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Alphaproteobacteria;D_3__Rhodospirillales;D_4__Rhodospirillaceae	0.2%	0.0%	0.2%	0.0%	0.1%	0.8%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Alphaproteobacteria;D_3__Rickettsiales;D_4__Mitochondria	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Alphaproteobacteria;D_3__Sphingomonadales;D_4__Sphingomonadaceae	0.1%	0.0%	0.2%	0.2%	0.1%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Alphaproteobacteria;D_3__Sphingomonadales;Other	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Betaproteobacteria;D_3__Burkholderiales;D_4__Alcaligenaceae	1.4%	1.6%	1.5%	1.7%	1.0%	1.6%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Betaproteobacteria;D_3__Burkholderiales;D_4__Burkholderiaceae	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Betaproteobacteria;D_3__Burkholderiales;D_4__Comamonadaceae	0.2%	0.8%	0.1%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Betaproteobacteria;D_3__Burkholderiales;D_4__Oxalobacteraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Betaproteobacteria;D_3__Neisseriales;D_4__Neisseriaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Deltaproteobacteria;D_3__Desulfovibrionales;D_4__Desulfovibrionaceae	0.1%	0.0%	0.1%	0.1%	0.0%	0.5%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Epsilonproteobacteria;D_3__Campylobacterales;D_4__Campylobacteraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Alteromonadales;D_4__Alteromonadaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Alteromonadales;D_4__Pseudoalteromonadaceae	0.1%	0.5%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Chromatiales;D_4__Chromatiaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Enterobacteriales;D_4__Enterobacteriaceae	17.8%	63.5%	14.1%	12.7%	2.2%	1.2%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Legionellales;D_4__Coxiellaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Oceanospirillales;D_4__Halomonadaceae	0.1%	0.7%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Pasteurellales;D_4__Pasteurellaceae	0.4%	0.0%	0.5%	0.6%	0.5%	0.1%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Pseudomonadales;D_4__Moraxellaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Pseudomonadales;D_4__Pseudomonadaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Xanthomonadales;D_4__Xanthomonadaceae	0.0%	0.1%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Proteobacteria;D_2__Gammaproteobacteria;D_3__Xanthomonadales;D_4__Xanthomonadales Incertae Sedis	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Saccharibacteria;D_2__uncultured bacterium;D_3__uncultured bacterium;D_4__uncultured bacterium	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Spirochaetae;D_2__Spirochaetes;D_3__Spirochaetales;D_4__Brachyspiraceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Synergistetes;D_2__Synergistia;D_3__Synergistales;D_4__Synergistaceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Tenericutes;D_2__Mollicutes;D_3__Anaeroplasmatales;D_4__Anaeroplasmataceae	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
	D_0__Bacteria;D_1__Tenericutes;D_2__Mollicutes;D_3__Mollicutes RF9;D_4__uncultured bacterium	0.0%	0.0%	0.0%	0.0%	0.0%	0.3%
	D_0__Bacteria;D_1__Tenericutes;D_2__Mollicutes;D_3__Mollicutes RF9;D_4__uncultured rumen bacterium	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
	D_0__Bacteria;D_1__Tenericutes;D_2__Mollicutes;D_3__Mollicutes RF9;Other	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Tenericutes;D_2__Mollicutes;D_3__NB1-n;D_4__uncultured bacterium	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%
	D_0__Bacteria;D_1__Tenericutes;D_2__Mollicutes;D_3__NB1-n;Other	0.0%	0.0%	0.0%	0.0%	0.0%	0.3%
	D_0__Bacteria;D_1__Verrucomicrobia;D_2__Opitutae;D_3__Opitutae vadinHA64;D_4__uncultured bacterium	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Bacteria;D_1__Verrucomicrobia;D_2__Verrucomicrobiae;D_3__Verrucomicrobiales;D_4__Verrucomicrobiaceae	0.3%	0.0%	0.6%	0.5%	0.3%	0.3%
	D_0__Bacteria;Other;Other;Other;Other	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	D_0__Eukaryota;D_1__Excavata;D_2__Metamonada;D_3__Parabasalia;D_4__Tritrichomonadea	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
	Unassigned;Other;Other;Other;Other	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%