Norwegian University of Life Sciences
Faculty of Veterinary Medicine and Biosciences
Department of Chemistry, Biotechnology and
Food Science

Master Thesis 2014
60 credits

Diet Effects on the Short-Term Temporal Dynamics of the Equine Hindgut Microbiota

Camilla Thollefsen Kristoffersen

Acknowledgements

This thesis was performed at the Norwegian University of Life Science (Institute of Chemistry, Biotechnology and Food Science), with Professor Knut Rudi and associated Professor Dag Austbø as supervisors and PhD. student Ekaterina Avershina as secondary supervisor.

First, I want to thank Dag Austbø for giving me the opportunity to work with horses, my great passion in life. And, thanks to Rasmus B. Jensen for giving me the opportunity to join the sampling in conjunction with your doctoral project and additionally thanks for your contribution with the pH and SCFA measurements (Rasmus B. Jensen owns all rights to the results). Furthermore, I want to thank Knut Rudi who helped me with the planning and execution of my laboratory work and additionally performed OLS regression analysis. You are always positive and enthusiastic about me and my thesis. Ekaterina Avershina, I admire your knowledge. Thank you for always answering all of my "stupid" questions and for doing a good job with the analysis of my 16S rRNA gene sequencing results in QIIME and MATLAB (PCA and weighted UniFrac analysis). And last but not least, thanks to all the positive people in the Microbial diversity group for making me feel welcome and look forward to each day in the lab.

Thanks to my best girl Shirley who has given me relaxation and fresh air every day. To the most excellent lab partner, Melissa Schanche, MC-girls forever! We have achieved a lot together, and I had not survived my education without you in my life the past five years. Henriette Solberg, the greatest roommate; you are my handy man and always making something good on heavy days. Thanks to the rest of my favorite fellow students, you know who you are. I really appreciate you all.

Finally, a special thanks to my parents who have always been supportive no matter the context.

I am lucky to have you!

Ås, May 2014

Camilla T. Kristoffersen

II

Abstract

Horses' living conditions have changed through thousands of years; altering their diets, that no longer fit the horse's physiology. Horses were free-ranging grazing animals and their digestive system is therefore adapted to cope with large quantities of fibrous feeds. Mixed roughage and concentrate diets will consequently affect how horses utilize nutrients. In order to optimize feeding rations, it is important to understand how the hindgut microbiota reacts when different diets are presented. The aim of this thesis was therefore to investigate the short-term temporal dynamics of the equine hindgut microbiota by using 16S rRNA gene and shotgun metagenomic sequencing. The microbiota and associated metabolic products were compared to investigate the diet effects on the equine hindgut microbiota, through cecal content collected in a time period of 24 hours from four cecally cannulated horses, given two different diets. Additional fecal samples were also collected in the same time range, to investigate whether fecal samples could represent the microbial population of the cecum. The detected dominant bacterial phyla, in equine cecum, comprised of the predominating phyla *Firmicutes* and *Bacteroidetes,* followed by the phyla *Verrucomicrobia, Proteobacteria, Tenericutes, Spirochaetes, Cyanobacteria* and *Fibrobacteres*. Based on the observed temporal patterns of the microbiota, we propose a model where the fibrolytic bacteria *Fibrobacter succinogenes* produce SCFA that lead to a pH decrease in the cecum and the resurgence of lactic acid-producing *Streptococcus spp*., where ultimately growth of lactate utilizing *Anaerovibrio spp*. are believed to prevent the accumulation of lactate in the cecum for a prolonged period of time. Further, the cecal content and feces showed significant difference, suggesting that feces cannot represent the cecal microbiota in a proper way. This study provides a foundation for further understanding of the equine hindgut microbiota and its function, allowing production of feeds that are more adapted to this intestinal ecosystem and may prevent diseases in the future.

Sammendrag

Hestens levekår har endret seg gjennom tusenvis av år; endret sin diett, som ikke lenger passer hestens fysiologi. Hester var frittgående beitedyr og deres fordøyelsessystem er derfor tilrettelagt for å håndtere store mengder fiberholdig fôr. Rasjoner med grovfôr og kraftfôr vil dermed påvirke hvordan hesten utnytter næringsstoffene. For å optimalisere rasjonen, er det viktig å forstå hvordan blindtarmmikrobiotaen reagerer når ulike fôrtyper presenteres. Målet med denne masteroppgaven var derfor å undersøke den kortsiktige temporale dynamikken av hestens blindtarmmikrobiota ved hjelp av 16S rRNA gen- og shotgun metagenom sekvensering. Mikrobiotaen og tilhørende metabolske produkter ble sammenlignet for å undersøke effektene av diett på hestens blindtarmmikrobiota, gjennom oppsamlet blindtarmsinnhold i en tidsperiode på 24 timer fra fire blindtarmkanylerte hester, gitt to forskjellige dietter. Ekstra gjødselprøver ble også samlet i samme tidsperiode, for å undersøke om gjødselprøver kan representere det mikrobielle samfunnet i blindtarmen. De observerte dominerende bakterielle fyla i hestens blindtarm bestod av de mest dominerende fyla *Firmicutes* og *Bacteroidetes*, etterfulgt av fyla *Verrucomicrobia*, *Proteobacteria*, *Tenericutes*, *Spirochaetes*, *Cyanobacteria* og *Fibrobacteres*. Basert på de observerte temporale mønstrene av mikrobiota, foreslår vi en modell der den fibrolytiske bakterien *Fibrobacter succinogenes* produserer SCFA som har ført til en pH nedgang i blindtarmen og oppblomstring av melkesyreproduserende *Streptococcus spp*., hvor til slutt vekst av laktatutnyttende *Anaerovibrio spp*. antas å hindre akkumulering av laktat i blindtarmen over et lengre tidsrom. Videre viste blindtarmsinnhold og gjødsel signifikant forskjell, som tyder på at avføringen ikke kan representere blindtarmmikrobiotaen på en pålitelig måte. Denne studien gir et grunnlag for videre forståelse av hestens blindtarmmikrobiota og dens funksjon, slik at produksjonen av fôr som er mer tilpasset dette tarmøkosystemet muliggjøres og kan forebygge sykdommer i fremtiden.

Abbreviations

ADF	Acid detergent fiber
ANOVA	Analysis of variance
bp	Base pairs
CO_2	Carbon dioxide
DNA	Deoxyribonucleic acid
IHA	Institute for Animal and Aquacultural Science
MG-RAST	Metagenomic RAST server
mRNA	Messenger ribonucleic acid
NDF	Neutral detergent fiber
OLS	Ordinary least squares
OTU	Operational taxonomic unit
PC	Principal component
PCA	Principal component analysis
PCR	Polymerase chain reaction
QIIME	Quantitative Insight Into Microbial Ecology
qPCR	Quantitative polymerase chain reaction
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
SCFA	Short-chain fatty acids
SEM	Standard error of the mean
STD	Standard deviation

Table of content

1. Introduction

Horses' living conditions have changed through thousands of years; altering their diets, that no longer fit the horse's physiology (Daly et al. 2001). Horses were free-ranging grazing animals and their digestive system is therefore adapted to cope with large quantities of fibrous feeds. Mixed roughage and concentrate diets will consequently affect how horses utilize nutrients. By understanding the digestive system and nutrient utilization by the gut microbiota, feeds that are more adapted to this intestinal ecosystem can be created and disease may be prevented.

1.1 The equine digestive system

1.1.1 The gastrointestinal tract

Through the different compartments of the equine gastrointestinal tract (Fig 1-1), the feed are exposed to enzymatic degradation and microbial fermentation, where the mean total retention time has been found to range between 20 and 30 hours (Austbø & Volden 2006; Rosenfeld et al. 2006). In the mouth, the horse's teeth crush the ingested feed, while it is mixed with produced saliva. The horse's saliva contains no degradation enzymes as in humans (Julliand et al. 2006). However, it is functioning as a pH buffer (contains bicarbonate) and lubrication for the horse's esophagus. Through the esophagus the feed enters the stomach followed by acid degradation. The digesta only remains in the stomach for a short period of time (2-6 hours) and stomach contractions, initiated by newly arrived substances, moves digesta further into the small intestine (Van Weyenberg et al. 2006) where the pre-cecal digestion of protein, starch and other carbohydrate compounds takes place via enzymatic feed degradation (Santos et al. 2011). The digesta moves rapidly (30 cm/min) through the about 20 meters long small intestine, which is comprised by duodenum, jejunum and ileum. Pancreatic juices are added in duodenum to neutralize the acid from the stomach (Van Weyenberg et al. 2006). The mean pHs in duodenum, jejunum and ileum have shown to be 6.3, 7.1 and 7.5 respectively (Mackie & Wilkins 1988). Substrates that have not been absorbed in the small intestine are transported from ileum to the cecum (Santos et al. 2011). The main function in the equine hindgut, cecum and colon, is

microbial degradation and fermentation of fiber into substances utilized for the horse's energy requirements (Julliand et al. 1999). pH decreases in the cecum to about 6.7, due to the fermentation process (Mackie & Wilkins 1988). Most digesta reach the cecum and the ventral colon within three hours, and thus the main digestion takes place in the hindgut. Finally, the digesta reaches the rectum where the remaining water is absorbed (Van Weyenberg et al. 2006).

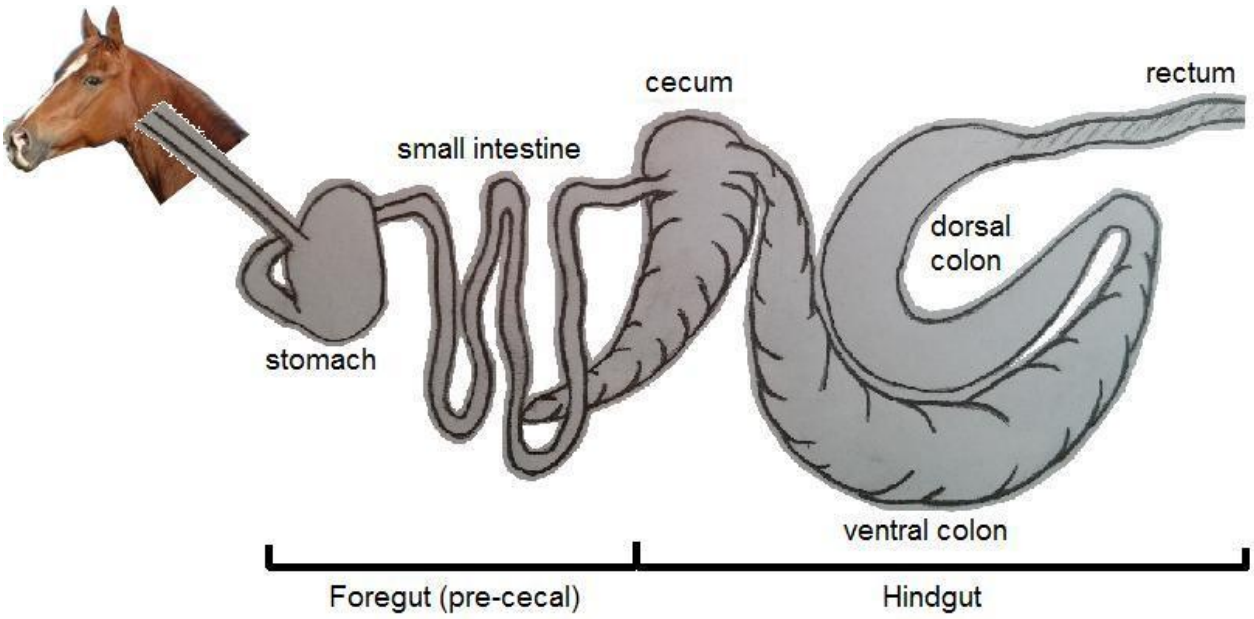


Figure 1-1: The equine gastrointestinal tract. The feed enters the foregut where enzymatic degradation takes place and are further transported to the hindgut for microbial fermentation (Kristoffersen, this thesis).

1.1.2 Hindgut microbiota

The hindgut microbiota live in symbiosis with the host by helping breaking down fiber compounds, while the host contributes with a regular carbohydrate source for the gut microbiota (Santos et al. 2011). *Firmicutes* and *Bacteroidetes* are the predominant phyla in the equine hindgut (Costa & Weese 2012; Flint et al. 2008; O' Donnell et al. 2013). O' Donnell et al. (2013) investigated the core fecal bacterial microbiome of Irish Thoroughbred racehorses and found the dominant phyla to be represented by *Proteobacteria*, *Verrucomicrobia*, *Actinobacteria*, *Euryarchaeota*, *Fibrobacteres* and *Spirochaetes*, in addition to *Firmicutes* and *Bacteroidetes*. Up to as much as 80% of the microbiota in the cecum and colon are estimated to be strict anaerobes and on average 78% of the microbiota are cellulolytic (Santos et al. 2011).

Bacterial fermentation processes in the hindgut produce short-chain fatty acids (SCFA), microbial mass, methane and fermentation heat. These patterns indicate microbial activity and digestibility of substrates; providing insight into which metabolic pathways the hindgut microbiota utilizes (Santos et al. 2011). The SCFA absorbed across the gut mucosa (Costa & Weese 2012) constitutes as much as 60-70% of the horse's energy resource (Biddle et al. 2013; Costa & Weese 2012) and may be used as substrate in body tissue metabolism (Jansson & Lindberg 2012). Acetate and butyrate can be converted into acetyl-CoA which is further used as a substrate in the citric acid cycle (aerobic metabolism). Propionate however, is mainly used in gluconeogenesis (Jansson & Lindberg 2012). The types and amounts of SCFA produced by the hindgut microbiota, depends on substrate availability, microbiota composition and intestinal passage rate (Macfarlane & Macfarlane 2003).

1.1.3 Digestive associated disease

A stable microbiota is crucial for the horse's health and imbalance in the gut microbiota may lead to severe disease (Costa & Weese 2012). Laminitis is, together with colic, a widespread disease caused by intestinal complications in the horse. A frequency study by Wylie et al. (2011) reported findings of equine laminitis ranging from 1.5% to 34%. Laminitis is a painful disease characterized by lameness, which often becomes a chronic problem (Katz & Bailey 2012). And, due to animal welfare reasons often results in euthanasia (Sloet van Oldruitenborgh-Oosterbaan 1999).

Starch induced laminitis is caused by starch overload in the cecum (Katz & Bailey 2012). Domestic horses, and especially competition horses, spend much time indoor and on the training pitch which practically leads to unfortunate eating habits. Due to the horse's high energy demands, starch-rich concentrates are often fed in greater quantities (Julliand et al. 2006). By ingestion of large concentrate rations, with high starch content, the small intestine has trouble coping with the major enzyme digestion requirement. Therefore, a large proportion of un-degraded starch will enter the cecum (Brøkner et al. 2012) leading to a change in cecal microbiota, promoting lactic acid producing bacteria (Katz & Bailey 2012). These bacteria favor readily hydrolysable carbohydrates as a substrate for fermentation, and therefore they quickly multiply in starch-rich environment and produce lactic acid and CO_2 (Daly et al. 2012), which

further leads to a pH drop. This acidic environment may cause loss of barrier function in the cecum through degenerative changes in the epithelial cells. The loss of barrier function may in turn result in influx of unknown endotoxin into the circulatory system (Katz & Bailey 2012), leading to blood supply disturbance in the laminar region. This process makes the pedal bone separate from the hoof wall, due to degradation of the laminae (Sloet van Oldruitenborgh-Oosterbaan 1999).

1.2 Nucleic acid based technologies

The availability of nucleic acid based methods, which may replace culture-dependent analysis, is increasing. These molecular methods are generally less time- and labor-intensive, which leads to increased efficiency in the laboratory and thus increased amount of data.

In general, deoxyribonucleic acid (DNA) from a sample is isolated, and the desired genes are amplified by polymerase chain reaction (PCR), and further sequenced. The DNA isolation process separates the cell's DNA from the cell's other components. Often, the samples do not contain enough amount of DNA to appropriate sequencing; therefore, the DNA in the samples must be amplified by PCR. Additional reasons for using PCR amplification may be cases were sequencing primers are needed to be incorporated for allowing the sequencing reaction to initiate. However, single-molecule sequencing are also possible (Harris et al. 2008). The sequencing processes result in information about the DNA nucleotide sequences, which further can be used to identify bacteria taxonomic relations or their functions depending on the selected DNA sample.

1.2.1 Quantitative polymerase chain reaction

Quantitative polymerase chain reaction (qPCR) is a method which can detect and quantify microorganisms independent of cultivation (Yu et al. 2005). Due to the speed, sensitivity and reproducibility, qPCR is widely accepted (Mackay et al. 2002). qPCR works in the same way as qualitative PCR (denaturation of double stranded DNA, primer annealing and elongation by inserting complementary bases), except that the use of fluorescence labeling makes it possible to monitor the concentration of the product through the amplification cycles, where the fluorescence

intensity reflects the amplicon copy number in real time (Valones et al. 2009). There are different types of fluorescent reagents that can be used in qPCR, like dyes which bind to double stranded DNA (e.g. EvaGreen) and DNA sequence-specific probes (e.g. TaqMan) (Valones et al. 2009). The initial concentration of DNA can be estimated by examining the changes in the PCR product concentration through the amplification cycles (Zhang & Fang 2006).

Response curves showing the amplification phase for each individual reaction, describe the difference between each sample's initial template DNA amounts. The amount of template DNA is reflected by the number of cycles required to reach a specific fluorescence signal level (Kubista et al. 2006). Accordingly, the Ct-values correspond to the cycle number were the fluorescence level reach the threshold (Fig. 1-2).

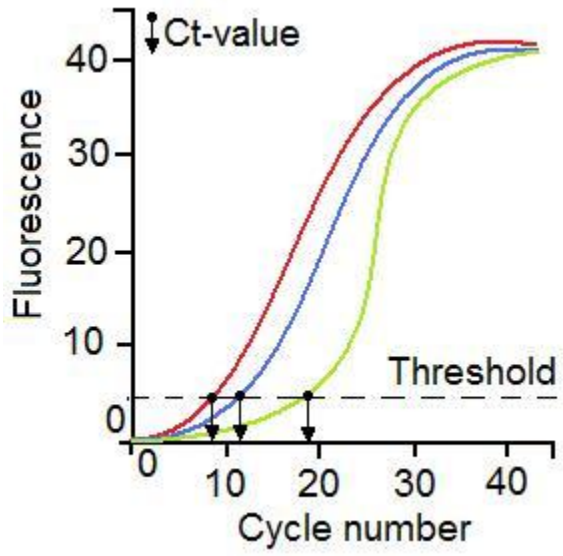


Figure 1-2: Quantitative polymerase chain reaction response curves. Ct-values are registered when the sample fluorescence signal reach the threshold (Kristoffersen, this thesis).

1.2.2 Next-generation sequencing technologies

Over the past 10 years there has been a tremendous increase of sequenced genomes, which is due to the development and improvement of next-generation sequencing technologies (Forde & O'Toole 2013). The first next-generation sequencing system on the marked was the 454 GenomeSequencer FLX instrument, developed by 454 Life Sciences (Ansorge 2009). However,

Illumina, another next-generation sequencing system, has received great popularity recently (Nelson et al. 2014). The various next-generation sequencing platforms often have a common workflow. Modified DNA fragments, with platform-specific PCR and sequencing primers, form the sequencing library. Further, the sequencing library is amplified to form clusters of copies, originating from each DNA fragment. Finally, all fragments are sequenced in parallel, where each cluster generates information about the DNA fragment nucleotide sequence (Meaburn & Schulz 2012). The sequence identity is obtained by comparing the query sequence with a database e.g. BLAST (Petrosino et al. 2009) and the DNA is quantified based on sequence reads (Ahn et al. 2011).

Illumina/Solexa

The Illumina technology enables outputs of 2x300 base pair (bp) read length, with up to 25 million sequencing reads (http://www.illumina.com/systems/miseq.ilmn, 09.05.14 17:22). Originally this sequencing technology was developed by the company Solexa, which later was acquired by Illumina (http://www.illumina.com/technology/solexa_technology.ilmn, 09.05.14 17:30). Illumina perform solid-phase amplification that achieves amplified templates, which are required to read fluorescence signals during the sequencing reaction. The DNA templates attach to a solid surface, which leads to a spatial separation of the templates and also enables thousands of sequencing reactions simultaneously (Fig. 1-3) (Metzker 2010). A single stranded DNA molecule anneals to a complementary slide-primer and the extension reaction synthesize a copy of the template molecule. The double stranded DNA molecule denaturizes and covalently binds to another slide-primer forming a bridge, which further leads to synthesis of a new copy. This process continues for several cycles until a cluster of copied DNA strands are accomplished (Bentley et al. 2008). Universal sequencing primers can then bind to the free ends in the clusters and begin the sequencing reaction (Metzker 2010).

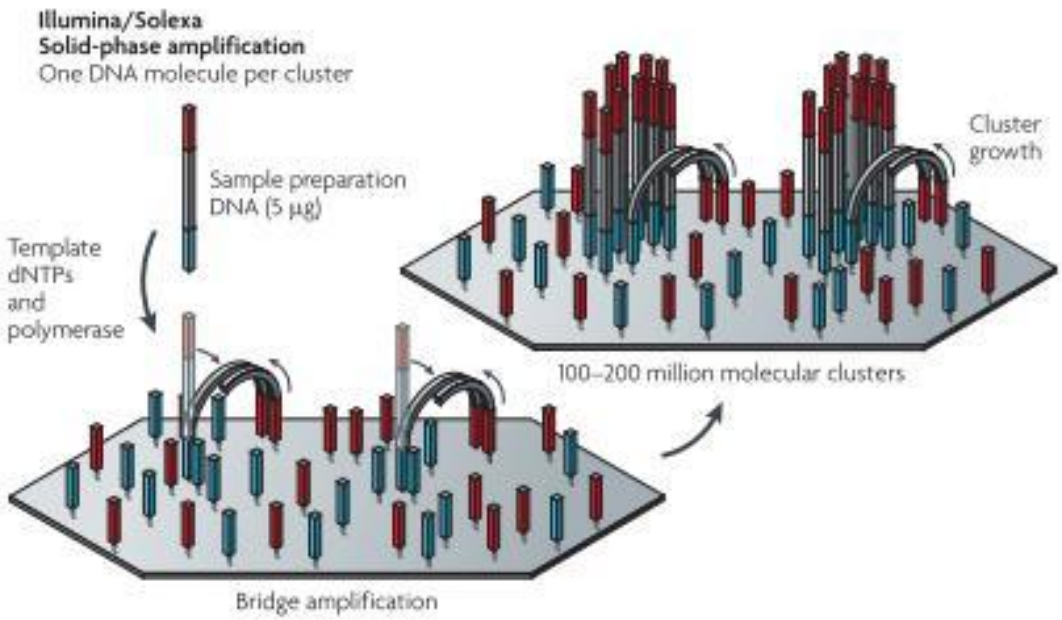


Figure 1-3: Illumina/Solexa solid-phase amplification. The template binds to covalently-attached forward and reverse primers on the slide, and by bridge amplification produce millions of separated clusters originating from a single DNA template (Metzker 2010).

The cyclic reversible termination method uses modified nucleotides with a protective group that terminate DNA synthesis allowing nucleotide reading and further continued DNA synthesis, when the protective group is removed (Metzker 2005). The sequencing cycles consist of three steps: incorporation of nucleotide, imaging and removing the terminator and the fluorescence marking. The available nucleotides are labeled with four different colors, and the DNA polymerase incorporates the matching nucleotide. The modified nucleotides do not contain a free 3'-OH group, and due to the 3'-blocking only one nucleotide can be incorporated by the DNA polymerase in each cycle. The remaining nucleotides are then removed and a color of fluorescence signal in each cluster is observed. This color identifies the incorporated nucleotide in each cluster. Further, the fluorescence marking is removed and the terminator is cleaved off, which leads to further extension by the DNA polymerase when new modified nucleotides are added. The process continues in several cycles with nucleotide incorporation, imaging and 3'-unblocking (Fig. 1-4) (Metzker 2010).

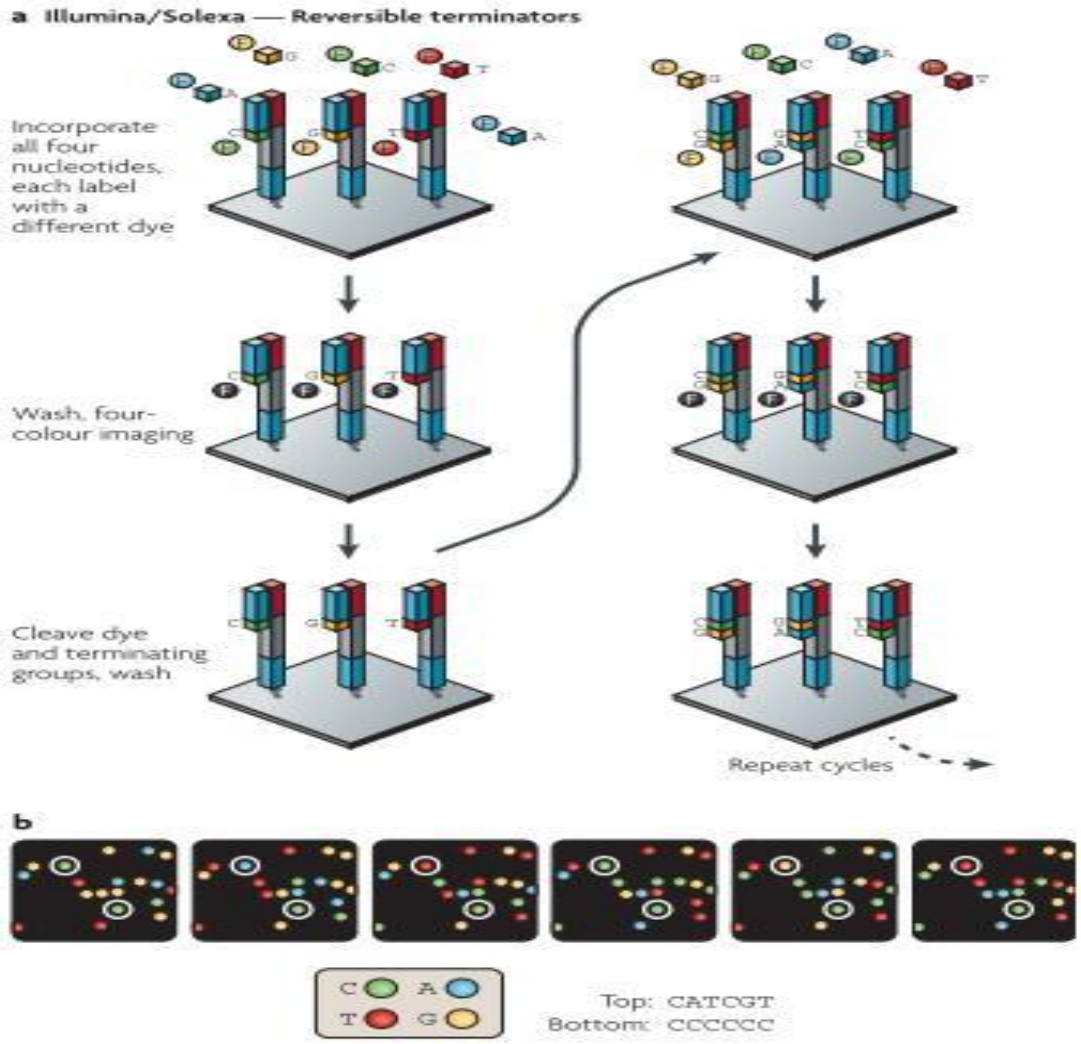


Figure 1-4: Four-color cyclic reversible termination method. A) Cycles of reversible termination identify the template nucleotide sequence. **B)** An image from each cycle provide information about, by four-color fluoresces coding, which base that has been incorporated in the sequencing reaction. The sequences represent the two highlighted clusters (Metzker 2010).

1.3 Metagenome sequencing

Culture-independent investigations of a mixed microbial community, that reside in a specific environment, are referred to as metagenomics (Petrosino et al. 2009). The insights into microbial communities have increased the recent years due to metagenomics, where 16S rRNA genes from the whole microbial community or all DNA from environmental samples are sequenced (Meyer et al. 2008).

1.3.1 Data generation

16S rRNA gene sequencing

Due to importance of 16S ribosomal ribonucleic acid (rRNA) for the cell during the translation of messenger RNA (mRNA) into protein, the 16S rRNA gene is well conserved in all organisms. Since certain areas of this gene are more susceptible to mutations (Olsen & Woese 1993), the 16S rRNA gene contains both conserved and variable regions. This information can be used to classify microorganisms on different taxonomic levels (Zhang & Fang 2006). The conserved regions are useful for sequence homology recognition (used for primer design), but give no phylogenetic information. Organisms that are distantly related can be distinguished by examining slightly variable regions, but organisms that are closely related can commonly be distinguished by examining highly variable regions (Olsen & Woese 1993). Comparing detected sequences with reference sequence databases, like National Center of Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/), enables bacterial identification.

Shotgun metagenomics

Instead of the widespread rRNA gene sequencing, whole-genome shotgun sequencing of metagenomic DNA may in future become more attractive (Davenport & Tummler 2013). The metagenomic gene pool encodes functional categories, individual pathways and fitness traits, which provide insight into the microbial community's specific features (Davenport & Tummler 2013).

A common way to prepare metagenomic libraries is by fragmenting the DNA either by mechanical force or by enzymatic digestion, followed by end-repairing and adapter ligation (van Dijk et al. 2014). Methods that combine both steps also exist, like Illumina Nextera XT DNA, where transposomes fragment and adds adapters at the same time in a limited cycle PCR reaction (Illumina 2012). Further, a size selection step is performed to remove remaining adapters and for selecting molecules of desired size. Due to often low template DNA quantities, PCR amplification is performed, and additionally may be performed to add additional adapter sequences, resulting in molecules completely ready for bridge amplification and sequencing (van Dijk et al. 2014).

1.3.2 Data analysis

New and improved sequencing methods are in constant development. However, analysis of this increasing amount of raw data creates problems. Therefore, bioinformatic tools play a crucial role in the interpretation of these data.

Taxonomic analysis

The 16S rRNA gene is usually used to generate information about the taxonomy of a single bacteria or a metagenome sample containing a whole community of bacterial species. The amplicon sequences are compared to a database containing previous sequenced species and are assign taxonomic classification. Based on high-throughput amplicon sequencing, Quantitative insight into microbial ecology (QIIME) can be used to compare and analyze microbial communities with billions of sequences from thousands of samples. The program converts raw data by clustering sequences into so-called operational taxonomic units (OTUs), assigns taxonomy and constructs phylogenetic trees (Caporaso et al. 2010).

Functional analysis

In contrast to taxonomical analysis, functional analysis usually converts the raw sequences to annotated proteins and wherefrom provide functional information about the sequenced sample. Metagenomic RAST server (MG-RAST) is publicly available software for analyzing metagenome sequencing data, based on the SEED framework for comparative genomics (Meyer et al. 2008). Users may upload fasta formatted raw sequence data and the data will be normalized and processed by comparing to known sequence databases, such as NCBI BLAST, SQLite and Grid Engine. The MG-RAST software will automatically generate a summary of each uploaded sample and give it a unique internal ID. The server provides the ability to access phylogenetic and metabolic reconstructions, and other various data types. It also provides the ability to compare the metabolism and annotations in one or more of the uploaded metagenomes (Meyer et al. 2008). By end of March 2014, the MG-RAST server contained almost 17 000 publically available metagenomes in a total of about 114 000 uploaded metagenomes.

1.4 Aim of this thesis

Horses have become part of the everyday domestic animal household, but also become a big part of sports and betting. Norwegian Rikstoto had, in 2012, a total totalizator turnover of 3.9 billion Norwegian kroner (https://www.rikstoto.no/Hjem/OmRikstoto/, 09.05.14 17:20). A lot of money is spent on equestrian sports and the horses' health is important to enable them to provide maximum in sporting events. The main cause of death in domestic horses is diseases related to the horse's gastrointestinal tract. In addition, gastrointestinal diseases are a major cause of morbidity and economical loss in the horse industry (Daly et al. 2001). The horse's health is partly determined by feeding, where improper feeding may lead to disease due to microbial shifts. The equine hindgut microbiota is complex and has a crucial role in health and despite this, the understanding of the microbiota composition and function remains limited (Costa & Weese 2012). Gaining more understanding of the equine hindgut microbiota and how it affects the host may contribute to creation of more equine adapted feed, and further prevent different gut related disorders. **The aim of this thesis was therefore to investigate the short-term temporal dynamics of the equine hindgut microbiota according to dietary changes.**

Knowledge about equine hindgut microbiota quantity, characters and metabolic activity is limited, due to difficulty in obtaining samples (Dougal et al. 2012). Due to the difficulty in collecting equine hindgut samples, fecal samples are often used (Schoster et al. 2013), but whether these fecal samples provide a complete insight into the gut microbiota dynamics is rather questionable. **Therefore, a sub goal in this thesis was to compare cecal content and feces to decide whether feces samples may describe the temporal dynamics of the equine hindgut microbiota in a proper way.**

The approaches chosen, in this thesis to address these goals, were qPCR, Illumina sequencing of the 16S rRNA gene and shotgun metagenomes. qPCR was used to decide the quantity bacterial stability and, 16S rRNA gene metagenome and shotgun metagenomic sequencing were performed to decide the bacterial and functional diet effects of the equine hindgut microbiota throughout the collection time range.

2. Materials and methods

2.1 Study design and experimental setup

Samples were collected from four coldblood geldings, belonging to the Norwegian University of Life Science (NMBU), which had a cecal cannula placed close to the ileocecal-junction. This cannula makes it possible to sample cecal content without euthanizing the horse, and also contribute to the possibility of collecting samples in time series and diet comparisons for each individual horse.

The experiment was carried out in a crossover design with 2x2 horses and two diets. Cecum and feces samples were collected from the four horses fed the two different diets, only hay and hay plus pelleted barley. Sampling was carried out for 24 hours, starting just before the morning meal. To prevent intestinal complications, due to change in diets, the horses had a diet adaptation period of about two weeks between sampling days. Cecal content was sampled from all four horses every hour, plus fecal samples every second hour. A total of 192 cecum samples and 96 feces samples were collected in this study (Fig.2-1).

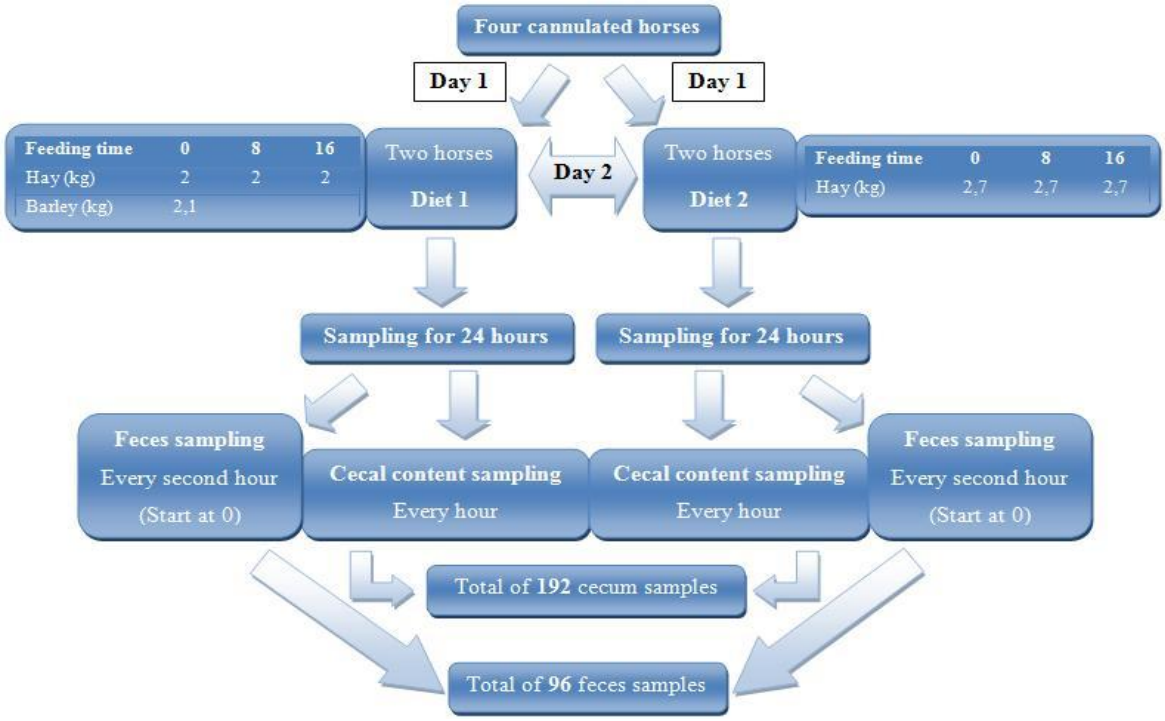


Figure 2-1: Study design. Four cannulated horses were given two different diets. A total of 288 cecal and fecal samples were collected every hour and every second hour respectively during 24 hour.

A flow chart of methods used is presented in Fig. 2-2. All reactions that are not temperature referred, has been performed at room temperature.

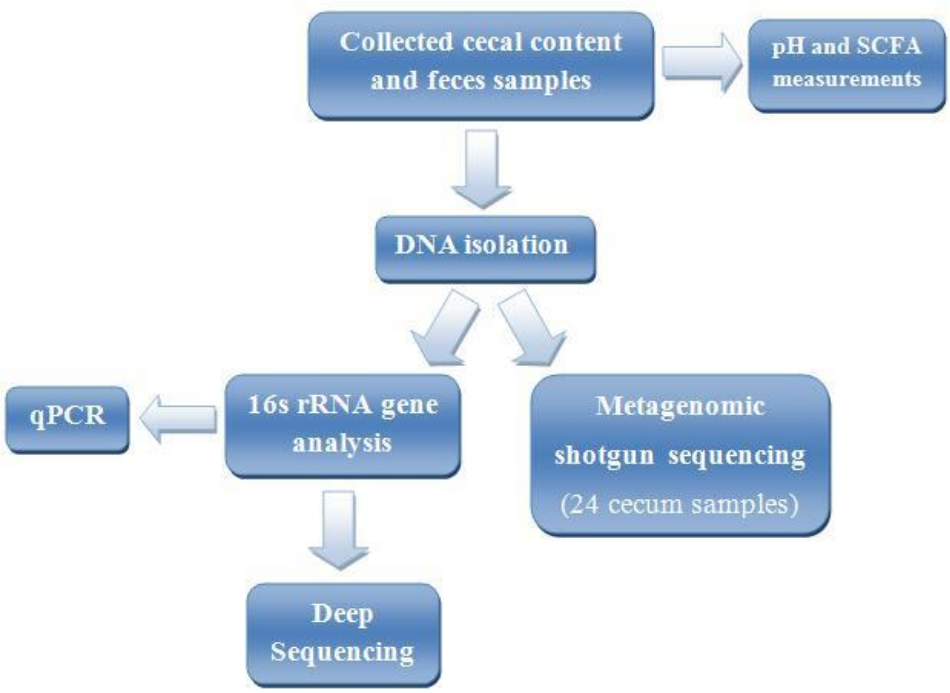


Figure 2-2: Experimental setup. pH and SCFA were measured in all samples. DNA was isolated from collected cecal content and feces samples. 16S rRNA gene analysis was conducted both by qPCR and deep sequencing. In addition, metagenomic shotgun sequencing was performed on 24 selected cecum samples.

2.2 Sampling and feed analysis

S.T.A.R. buffer (stool transport and recovery buffer; Roche, Germany) was added to all cecum and feces samples in 2:1 and 3:1 ratios respectively. The S.T.A.R. buffer prevents degradation of nucleic acids, and inactivates bacteria which protects against possible pathogenic bacteria (Espy et al. 2006). The samples were then frozen at -40 ºC until further use.

Sampling was done in collaboration with Rasmus B. Jensen (Ph.D. student at the University of Copenhagen), who measured pH and analyzed levels of SCFA in all samples. Cecal content and feces pH were measured with a pH electrode (SenTix® 41, WTW GmbH, Weilheim, Germany) immediately after sampling. The pH electrode was calibrated (at pH 4 and 9) between each

measurement. Additional subsamples were stored at -20ºC for analyses of SCFA composition. The SCFA analyses were performed at NMBU.

Nutritional content of the hay and the barley, used in this experiment, were analyzed at the Institute for Animal and Aquacultural Science (IHA) with laboratory manager Kari Norberg. Neutral detergent fiber (NDF) (IHA-nr:1041), consisting of hemicellulose, cellulose, lignin and silicate was measured by digesting the sample in a neutral detergent solution added sodium sulfite and thermostable α-amylase. The samples were then filtered, washed, dried and weighed. Acid Detergent Fiber **(**IHA-nr: MSP1037), consisting of cellulose, lignin and silicate was analyzed by same procedure as NDF with the exception of adding acid detergent solution instead of neutral detergent solution. Starch **(**IHA-nrMSP 1159), made up of maltose units, constitutes the major carbohydrate portion of grain. α-amylase were added to break down the starch three-dimensional structure to water soluble short chains. Amyloglucosidase enzyme was used for further degradation to glucose. Glucose concentration was then determined using a color reaction.

The hemicelluloses content in both hay and barley was calculated according to formula (2.1).

$$(2.1)\ Hemicellulose = NDF - ADF$$

2.3 DNA isolation

Genomic DNA was isolated using MagLGC™ Total Nucleic Isolation Kit. To ensure disruption of cell walls, samples were first subjected to mechanical lysis using glass beads. Samples were thawed and homogenized by vortexing and then 300 µl of the sample was transferred into a micro tube (Sarstedt, Germany) with acid-washed glass beads (<106 µm, 0.1 g) (Sigma-Aldrich, Germany). All the tubes were processed twice in the MagNa Lyser (Roche, Germany) at 2000 rpm for 40 seconds with 40 seconds rest between runs. Samples were kept cold during rest, to prevent DNA degradation. Finally the tubes were centrifuged at 13500 rpm for 5 minutes.

Further, to remove cellular proteins, 50 µl of lysis buffer and 5 µl of proteinase were added to 50 µl supernatant. The samples were then incubated in the KingFisher® Flex robot

(ThermoScientific, USA) at 55 °C for 10 min. From this step all samples were isolated in two parallels to detect variation between isolation runs.

The DNA extraction step was also performed on KingFisher® Flex robot (ThermoScientific, USA) and DNA extraction plates were prepared using an epMotion 5070 pipetting robot (Eppendorf, Germany). All steps were performed according to the manufacturer's recommendations.

The genomic DNA was stored at -20 ºC until further use.

2.4 Polymerase chain reaction

2.4.1 Quantitative polymerase chain reaction

Mainly as a control of the DNA isolation, qPCR with PRK primers (Appendix A: PRK341F and PRK806R) targeting prokaryotic 16S rRNA gene (Yu et al. 2005) was performed. Each reaction contained 1x HOT FIREPol® EvaGreen qPCR Mix Plus (Solis BioDyne, Estonia), 0,2uM forward and reverse primers (Life Technologies™, USA) and 1µl genomic DNA. The qPCR were performed by a LightCycler 480 (Roche, Germany) with the initial denaturation at 95 ºC for 15 min continued by 40 cycles of 95 ºC for 30 seconds and 60 ºC for 1 minute. Fluorescence was measured in the end of each cycle. Ct-values were calculated using LinRegPCR software (Ruijter et al. 2009) and further processed in Microsoft Excel (Microsoft, USA).

2.4.2 Sequencing PCR

To investigate the microbial content in the samples, Illumina sequencing of 16S rRNA gene was performed. The genomic DNA was amplified with PRK primers (Yu et al. 2005) using nested PCR approach, in order to increase the amplification success rate. In the second PCR run, primers were modified by addition of Illumina-specific adapters (PRKillumina primers; Fig. 2-3) (Hagen 2012; Jordhøy 2012). The library consisted of 16 forward primers and 36 reverse primers which made a total of 576 possible primer combinations (Appendix A: PRKi F and PRKi R).

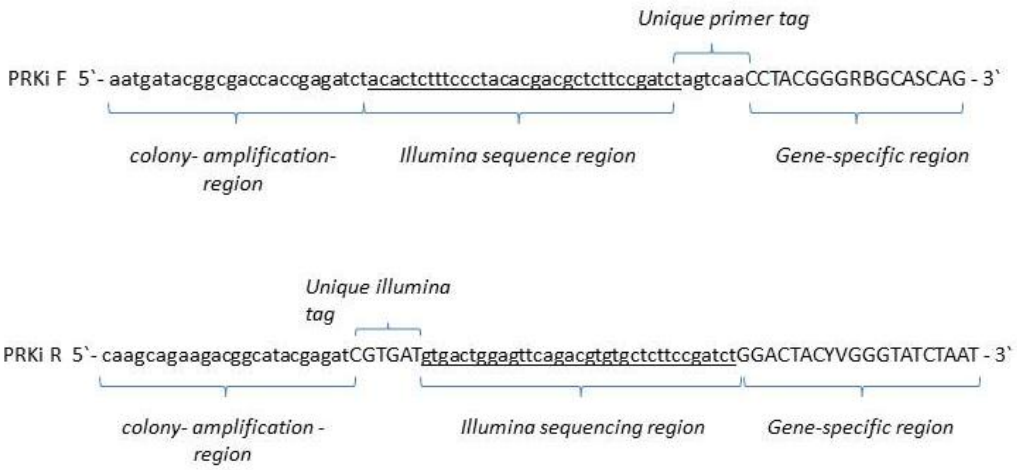


Figure 2-3: PRKillumina Forward (PRKi F) and PRKillumina Reverse (PRKi R) primers (Hagen 2012; Jordhøy 2012). These unique primers are modified with an illumina adapter; the 3'end consists of the gene-specific part while, the 5'end consists of the adapters with a colony amplification region and a sequencing region.

Each PCR reaction contained 1.25 U HotFirePol® DNA polymerase, 1x HotFirePol® buffer B2, 2.5 mM $MgCl_2$, 200 µM dNTPs (Solis BioDyne, Estonia), 0.2 µM PRK341F and PRK806R primers (Life Technologies™, USA) and 1 µl of template DNA. Amplification was done by using a 2720 Thermal Cycler (Applied Biosystems, USA) with initial denaturation at 95 ºC for 15 minutes and 25 cycles of denaturation at 95 ºC for 30 seconds, annealing at 50 ºC for 30 seconds and elongation at 72 ºC for 45 seconds. Finally, polymerization was finished at 72 ºC for 7 minutes.

The PCR products were then diluted 1:100 and these dilutions were used in the second PCR amplification step with unique PRKillumina primer combination for each sample. In this step, 10 cycles were used and annealing time was increased to 1 minute to ensure annealing of long primer oligonucleotides.

2.5 DNA quality and quantity control

Through the entire laboratory process, samples were qualitatively and quantitatively controlled.

2.5.1 PicoGreen

DNA concentrations were measured by using Quant-iT™ PicoGreen® dsDNA Assay Kit (Life technologies™, USA). PicoGreen reagents were prepared according to manufacturer protocol in a black 96 well nunc® microtiter plate (ThermoFisher, USA) and added 5 µl DNA. The use of black plate prevents background fluorescence and well to well scatter. The plate was incubated for 5 minutes allowing PicoGreen to bind the DNA molecules. In order to compare the DNA quantity with fluorescence measurements, a standard of known DNA concentrations (bacteriophage-λ DNA) was used. Fluorescence was measured by a FLX 800cse Microplate reader (BioTek, USA) with excitation at 480 nm, emission at 528 nm and a sensitivity of 50 as setup.

2.5.2 Qubit

DNA concentration was calculated by performing a Qubit® dsDNA HS Assay (Life Technologies™, USA). Preparations were done according to the manufacturer protocol (198 µl Working solution and 2 µl of sample DNA) and read in a Qubit™ fluorometer.

2.5.3 Gel electrophoresis

PCR products were controlled by 1% Agarose gel electrophoresis (90 V; 30 min) where the DNA fragments are separated by size. Due to DNA's negative charge, the fragments will migrate to positive pole in an electric field. The gel pores slows the migration process for larger fragments, leading to separation by fragment size. A 100 bp ladder (Solis BioDyne, Estonia) was added to determine DNA fragment sizes. GelRed™ dye (VWR, USA), which binds to DNA and fluoresces when exposed to UV light, was used to visualize the bands with a Molecular Imager® Gel Doc™ XR Imaging (Bio-Rad laboratories, USA).

2.6 Sequencing

2.6.1 16s rRNA gene metagenome sequencing

Gel electrophoresis results were used to normalize the PCR product library. All samples were ranged by the gel band strength (strong, medium and weak) and pooled (2 µl, 5 µl and 10 µl for strong, medium and weak bands) accordingly. E.Z.N.A® Cycle-Pure kit (Omega bio-tek, USA) was used to purify the mixed PCR products twice. Finally, the samples were sent to University of Oslo for sequencing on a MiSeq® Sequencer (Illumina, USA).

2.6.2 Shotgun metagenomic sequencing

To investigate the gene content in the samples, shotgun metagenomic analysis with Nextera® XT DNA sample preparation kit (Illumina, USA) was performed. 24 cecum samples were selected based on the pH results. Samples were taken at start (time point 1) and end point (time point 3), in addition to samples which correspond to a drop in pH response in either diet (time point 2).

The Nextera® XT transposome fragments input DNA and simultaneously adds adapter sequences to the ends of these fragments, which further enable PCR amplification. In addition to amplification of the input DNA, the PCR step adds unique indexes and sequences required for cluster formation during the sequencing run.

Nextera® XT DNA sample preparation kit (Illumina, USA) was used according to manufacturer's recommendations with some exceptions. As recommended, the DNA library was purified with AMPure® XP beads (Beckman Coulter, USA) to remove remaining nucleotides and primer dimmers as well as to select PCR fragments with desired length. However, the amount of AMPure® XP beads (Beckman Coulter, USA) was increased to a ratio of 1:1. Further, samples were normalized based on DNA concentrations, measured by Qubit, instead of bead-based normalization recommended in the manual. 10 ng of DNA from each sample was added the library pool and then sequenced on a MiSeq® sequencing platform (Illumina, USA).

2.7 Data analysis

All p-values were calculated by Student's t-Test (two-tailed distribution and two-sample unequal variance), if not referred otherwise in the result part.

2.7.1 Quantitative Insight Into Microbial Ecology

The 16S rRNA gene sequencing results were analyzed using QIIME (Appendix B). Sequences were quality filtered and clustered at 99% identity level against the Greengenes v.13.5 database (Caporaso et al. 2010). Weighted UniFrac analysis was used to assess β-diversity. Mean centered principal component analysis (PCA) was performed to detect differentially responding OTUs. The analysis was done in MATLAB (MathWorks, USA) using PLS toolbox (Eigenvector, USA). Kruskal-Wallis and analysis of variance (ANOVA) statistics were performed on all weighted UniFrac diversity calculations.

2.7.2 MG-RAST

The shotgun metagenomic sequencing results were uploaded to and analyzed in MG-RAST (Meyer et al. 2008) for organism and functional classification. Default settings were used with maximum e-value of 1e-5, a minimum identity of 60% and a minimum alignment length of 15 amino acids. The M5NR database was used for taxonomic assignment and the Subsystems database was used for functional annotation.

3. Results

3.1 Diet composition

The nutritional content of the hay and the barley used in this project is shown in Table 3-1 (for full analysis results see Appendix C: Table A-1). The hay contained higher proportion of NDF than barley. The ADF was also higher in hay than in barley, while starch was only present in barley. The morning ration of the hay diet contained 1493.1 g NDF, 828.9 g ADF, 664.2 g hemicelluloses and 0 g starch. While, the morning ration of the hay and barley diet contained 1404.2 g NDF, 689.6 g ADF, 714.6 g hemicelluloses and 1047.9 g starch. The hay ration had higher NDF and ADF, while the hay supplemented with barley ration contained higher proportion of hemicelluloses and starch.

Table 3-1: Nutritional content of hay and barley used in this project.

	NDF (g/kg)	ADF (g/kg)	Hemicelluloses[(2.1)] (g/kg)	Starch (g/kg)
Hay	553	307	246	0
Barley	142	36	106	499

NDF=Neutral detergent fiber

ADF=Acid detergent fiber

[(2.1)] calculated by formula 2.1

3.2 pH and SCFA measurements

pH was measured in both cecal content and feces during the time period of 24 hours (Fig. 3-1). A pH drop, in the time interval 2 to 10 hours after the morning meal, was observed in the cecum when the horses were fed the hay and barley diet. Calculated by t-Test, time point 2, 6, 7 and 22 in the cecum samples showed significant diet difference (p-values were 0.037, 0.040, 0.036 and 0.046 respectively). Significant total diet difference were also detected in cecum (p=0.033). Feces samples did not show any trend or significant diet difference at any time point.

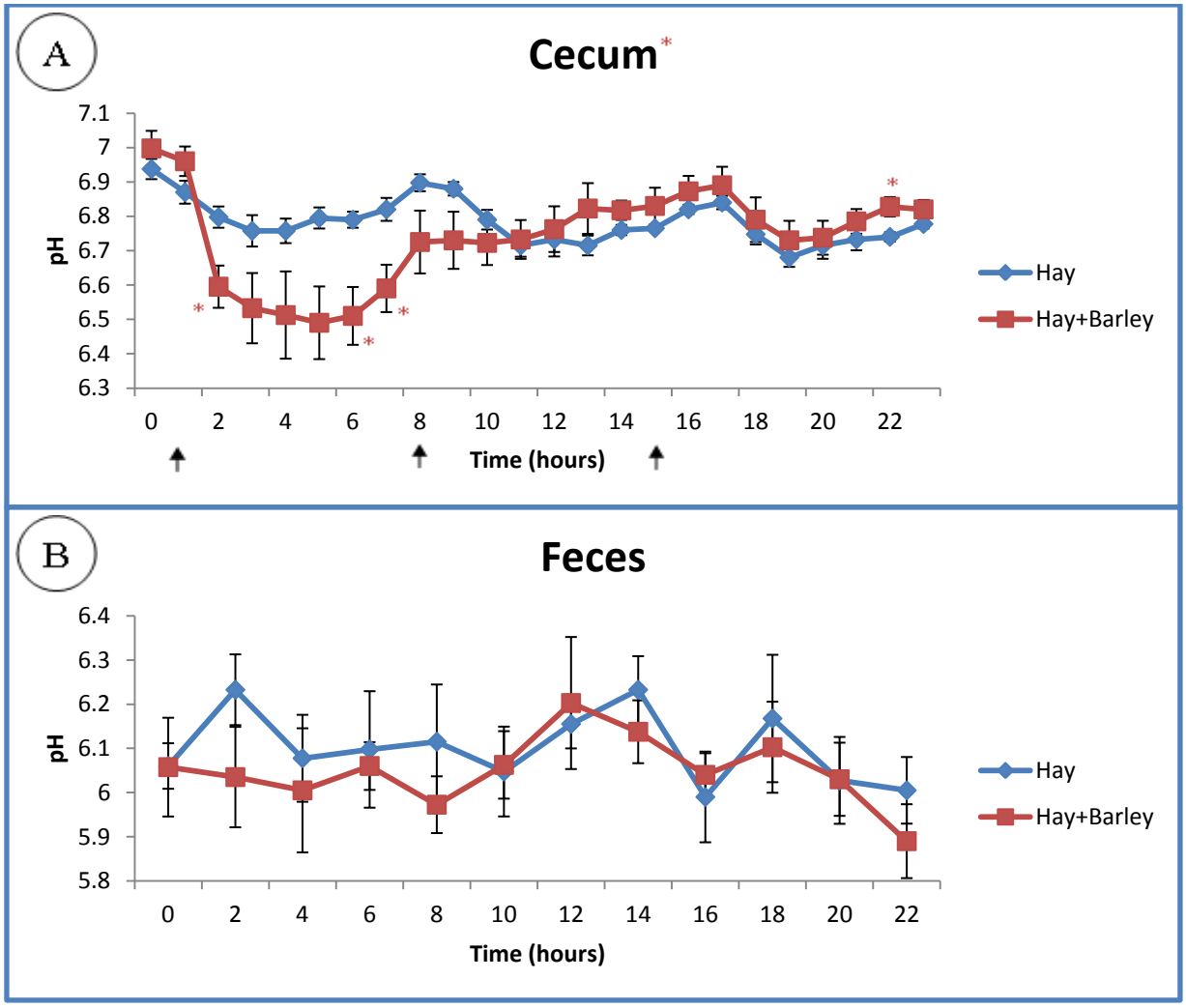


Figure 3-1: Measured pH in cecum (A) and feces (B) for both diets during 24 hours (mean ± SEM). Time is defined as hours after sampling start/morning meal. The arrows indicate the feeding time points. Asterisk associated with the diagram title indicate total significant diet difference, while asterisk associated with graphs indicate significant diet difference at specific time points (*0.010<p<0.050).

Fig. 3-2 presents SCFA amounts measured in all cecum samples. Total SCFA quantity remained relatively stable when the horses ate hay diet (Fig. 3-2A). However, when the horses were fed hay supplemented with barley, there was a major peak in total SCFA. This peak showed almost doubling of the total SCFA amount in the time range 2 to 8 hours after the morning meal, with significant diet difference at time point 2, 4 and 5 (p-values were 0.013, 0.010 and 0.007 respectively). Total SCFA amounts also showed significant total diet difference ($p=0.002$) although it seemed unaffected of the different hay quantity fed at 8 hours and 16 hours. In addition to total SCFA, specific SCFA (acetate, propionate and butyrate) amounts were measured in all the cecum samples.

The acetate quantity showed significant diet difference ($p<0.001$) with stable trend through small peaks after each feeding with only hay. In contrast, when the horses were fed the hay and barley diet, the acetate quantity dropped, between 0 and 16 hours after the morning meal, with significant diet difference at time point 2 to 14, where p-values ranged between 0.004 and 0.045 (Fig. 3-2B).

The propionate amounts transpired no clear peaks (Fig. 3-2C). However, slightly higher propionate amounts was detected in the hay supplemented with barley diet, compared to only hay diet ($p<0.001$), the first 16 hours after the morning meal with significant diet difference at time point 8 to 14 (0.016, 0.007, 0.007 and 0.017 respectively).

Butyrate amounts measured within the cecum of hay and barley fed horses peaked in the period from 2 to 10 hours while, when fed only hay diet, the butyrate amount remained quite stable (Fig. 3-2D). However, no significance was detected either in total diet comparison or at specific time points.

All the major diet differences in SCFA amount occurred throughout the first 16 hours and the butyrate/acetate and propionate/acetate ratio increased when the horses were fed hay and barley.

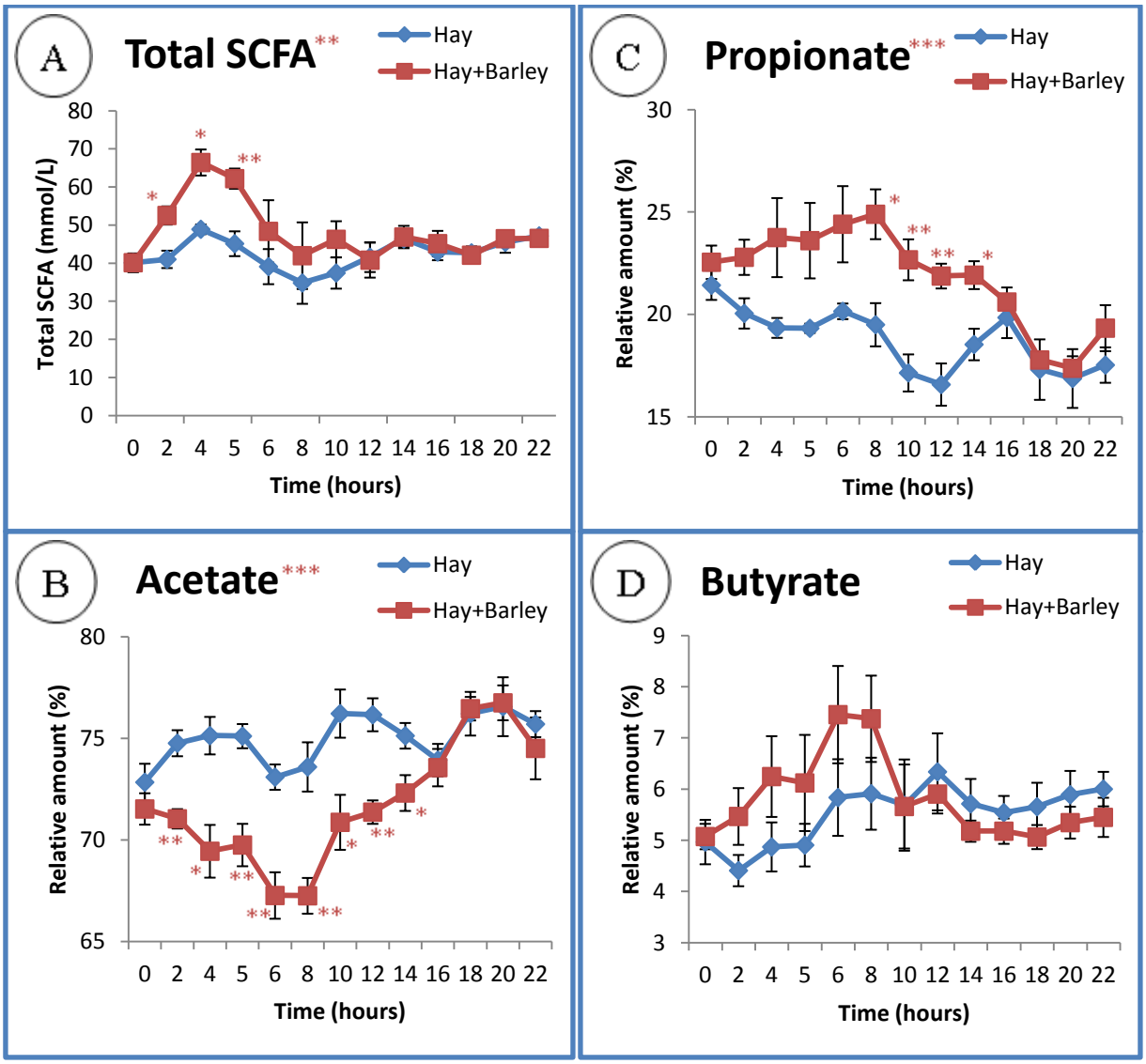


Figure 3-2: Measured SCFA in cecum during 24 hours (mean ± SEM). Time is defined as hours after sampling start/morning meal. **A)** Total SCFA (mmol/l) **B)** Acetate (mol/100mol) **C)** Propionate (mol/100mol) **D)** Butyrate (mol/100mol). The asterisk associated with the diagram title indicate total significant diet difference, while the asterisk associated with the graphs indicate significant diet difference at the specific time point (*0.010<p<0.050 **0.001<p<0.010 ***p<0.001).

3.3 Metagenome analyses

3.3.1 16S rRNA gene metagenome analyses

Sequence analysis was performed in QIIME to generate an OTU table. The sequence data were first quality filtered by removing reads with an average score less than Q25 and, additionally, removing reads shorter than 200 bp or with at least one nucleotide mismatch in the barcode region. The 16S rRNA gene sequencing generated a total of 5 309 669 sequences in 528 samples that passed the quality filtering. After quality filtering, sequences were clustered with 99% homology threshold and assigned taxonomy based on a closed reference search against Greengenes v 13.5 database to construct the OTU table. The OTU table was further edited by removing samples that contained less than 2 000 sequences in total, which removed a total of 43 samples. The numbers of detected sequences were then converted into relative amount (%) by dividing the number of detected sequences for each OTU on the total number of detected sequences for each sample. All samples combined, the final OTU table showed a total of 7 769 detected OTUs.

α-diversity comparison

Rarefaction curves, generated from QIIME, illustrated the α-diversity of observed species within different sample categories (Fig. 3-3). According to sample origin, observed species within the feces samples showed higher α-diversity than observed species within the cecum samples (Fig. 3-3A). The α-diversity within the individual horses was quite similar, with samples collected from horse 3 showing the lowest species diversity (Fig. 3-3B). Whereas, the α-diversity within samples with different diet demonstrated no clear species difference (Fig. 3-3C).

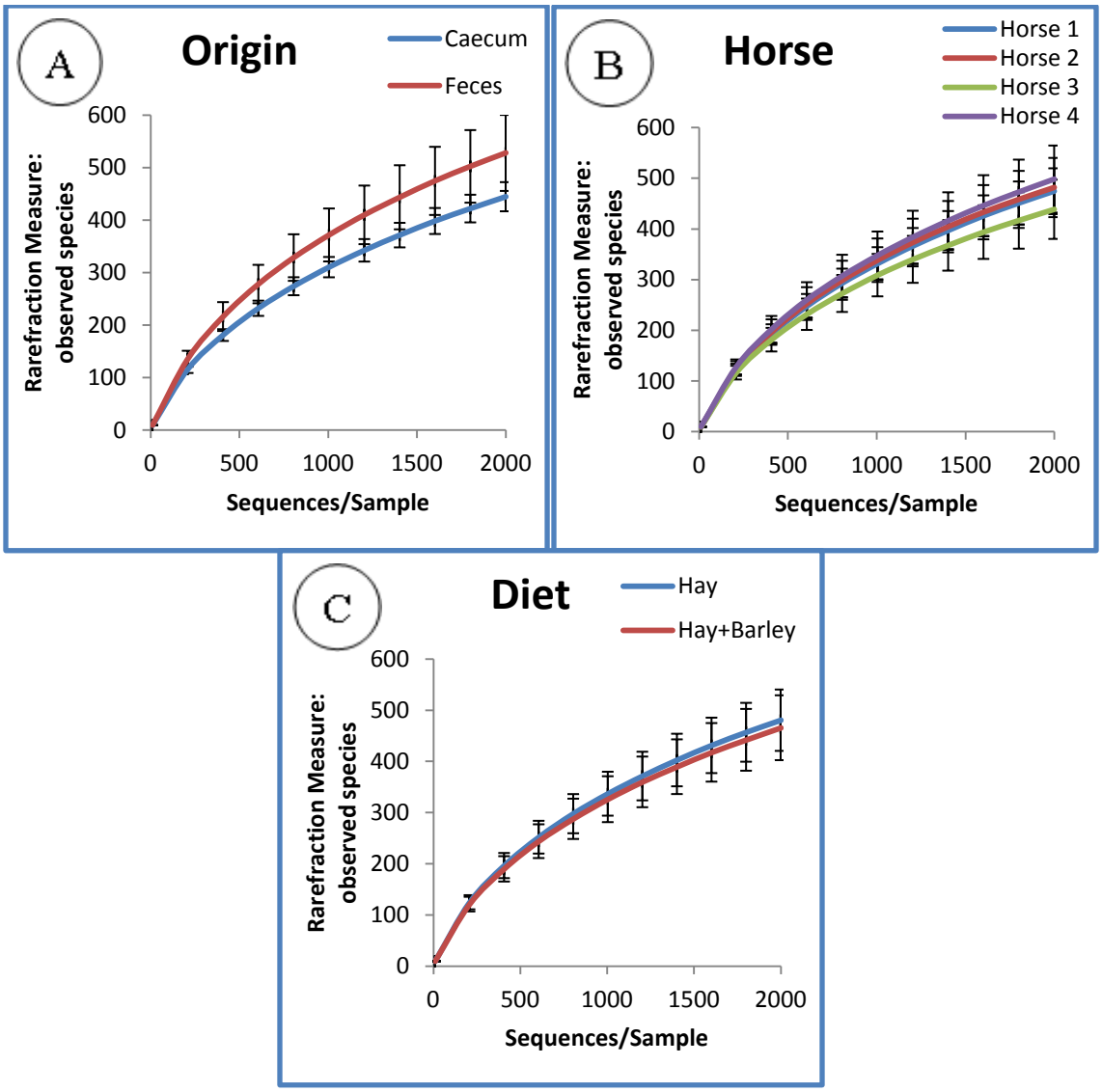


Figure 3-3: Rarefaction curves with observed species categorized as sample origin (A), individual horses (B) and diet (C).

β-diversity comparison

Weighted UniFrac diversity calculations showed higher within-group β-diversity between feces samples than between cecum samples ($p<0.001$), and even higher diversity was detected when comparing the two types of samples (Fig.3-4A) (Appendix D: Table A-2). Additionally, cecum and feces samples clustered separately in the weighted UniFrac plot (Appendix D: Fig. A-1A). All horses showed significantly different β-diversity between each other, in both feces (p-value between horse 1 and 4 were 0.041, while all other p-values<0.001) and cecum (all p-values<0.001). Furthermore, respective horses showed significant β-diversity difference between cecum compared to between feces. Horse 3 showed the lowest β-diversity between cecum samples and the highest between the feces samples and in contrast, horse 2 showed highest cecum diversity and lowest feces diversity (Fig.3-4B) (Appendix D: Table A-3). The weighted UniFrac diversity calculations, according to diet, showed significant diet difference between both cecum and feces samples (both p-values<0.001) with higher diversity in the hay diet compared to the hay plus barley diet (Fig.3-4C). Furthermore, respective diets showed significant diversity difference in cecum compared to feces (both p-values<0.001) (Appendix D: Table A-4).

In addition to UniFrac, PCA analysis of OTU abundance was used in order to assess an effect of diet, time and origin of samples. In the PCA plot by horse, horse 3 was located separately from the other three horses (Fig. 3-5A). However, such separate clustering of horse 3 was not detected in the feces samples (Fig. 3-5B). Diet effect in the cecum sample, was also observed in the PCA plot colored by diet (Appendix E: Fig. A-2). PC3 separated the diets, showing positive values for hay diet and negative values for the hay and barley diet, with some overlap in the transition between positive and negative PC3 scores. Such trend was indistinct in feces samples, with any form of clustering undetected. No pattern was observed in the PCA plot by time (Appendix E: Fig. A-3).

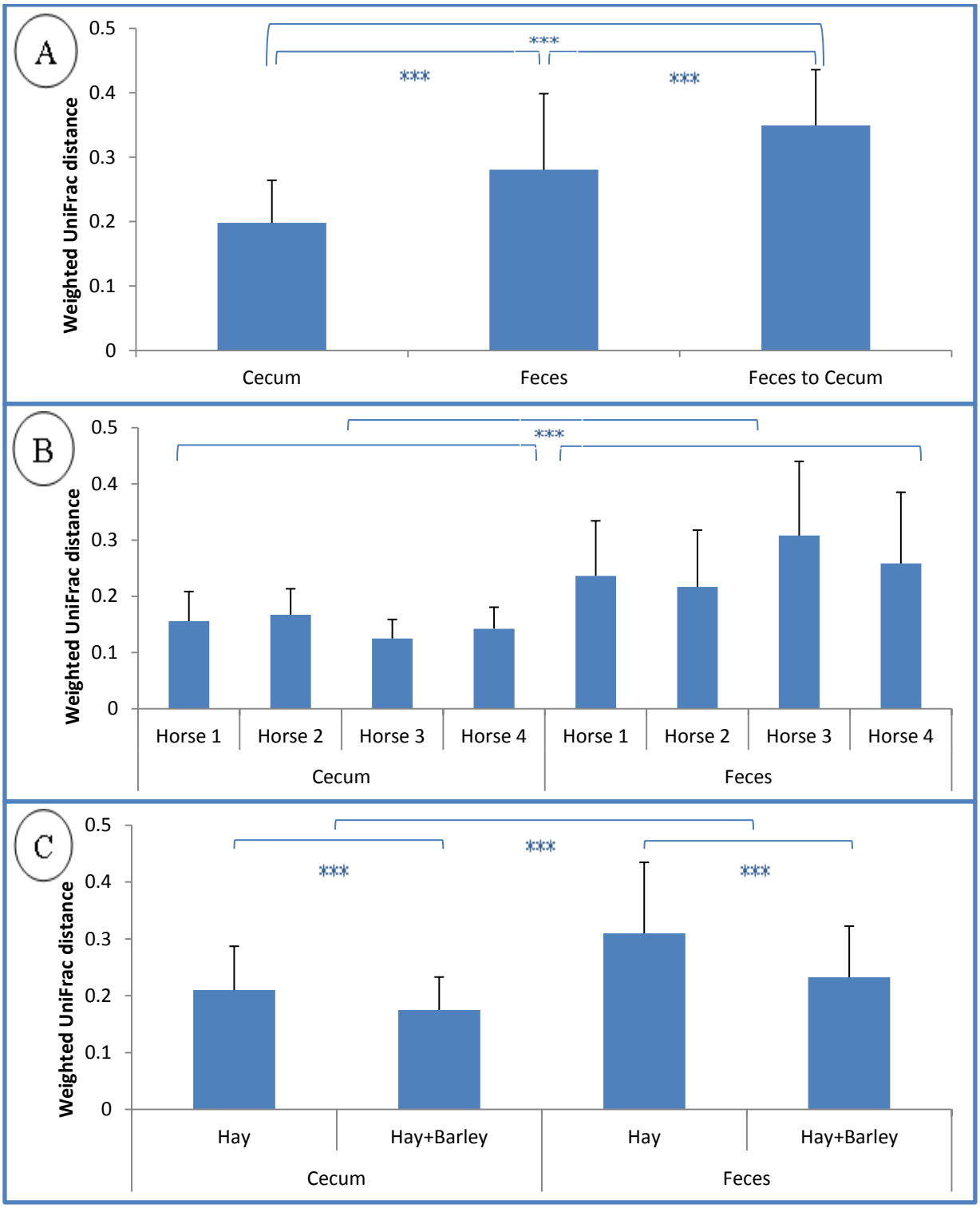


Figure 3-4: Weighted UniFrac diversity index diagram (mean + STD). A) Compare diversities within feces to those within cecum and those between these two groups. B) Compare diversities between the various horses. C) Compare diversities between different diets. ***$p < 0.001$

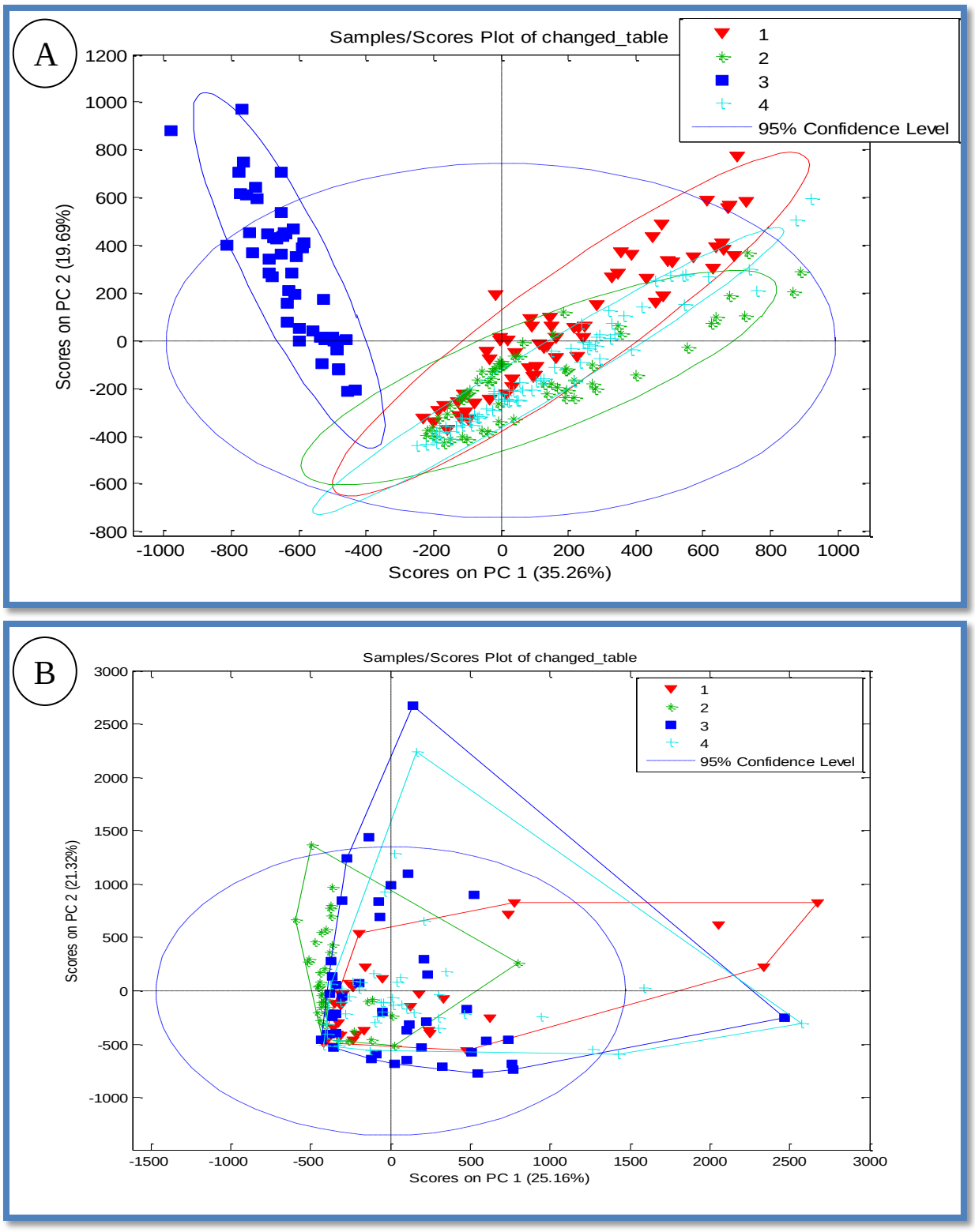


Figure 3-5: PCA plot by horse (PC1 vs. PC2). The numbers indicate horse number A) Cecum samples B) Feces samples

Dominant taxonomic groups

The phyla *Firmicutes* and *Bacteroidetes* dominated both cecum and feces independent of diet (Fig. 3-6A). The remaining phyla, which comprised over 1% of total amount in either diet or sample origin, were *Verrucomicrobia*, *Proteobacteria*, *Tenericutes*, *Spirochaetes*, *Cyanobacteria*, *Fibrobacteres*, *Actinobacteria* and *Euryarchaeota* (Fig. 3-6B).

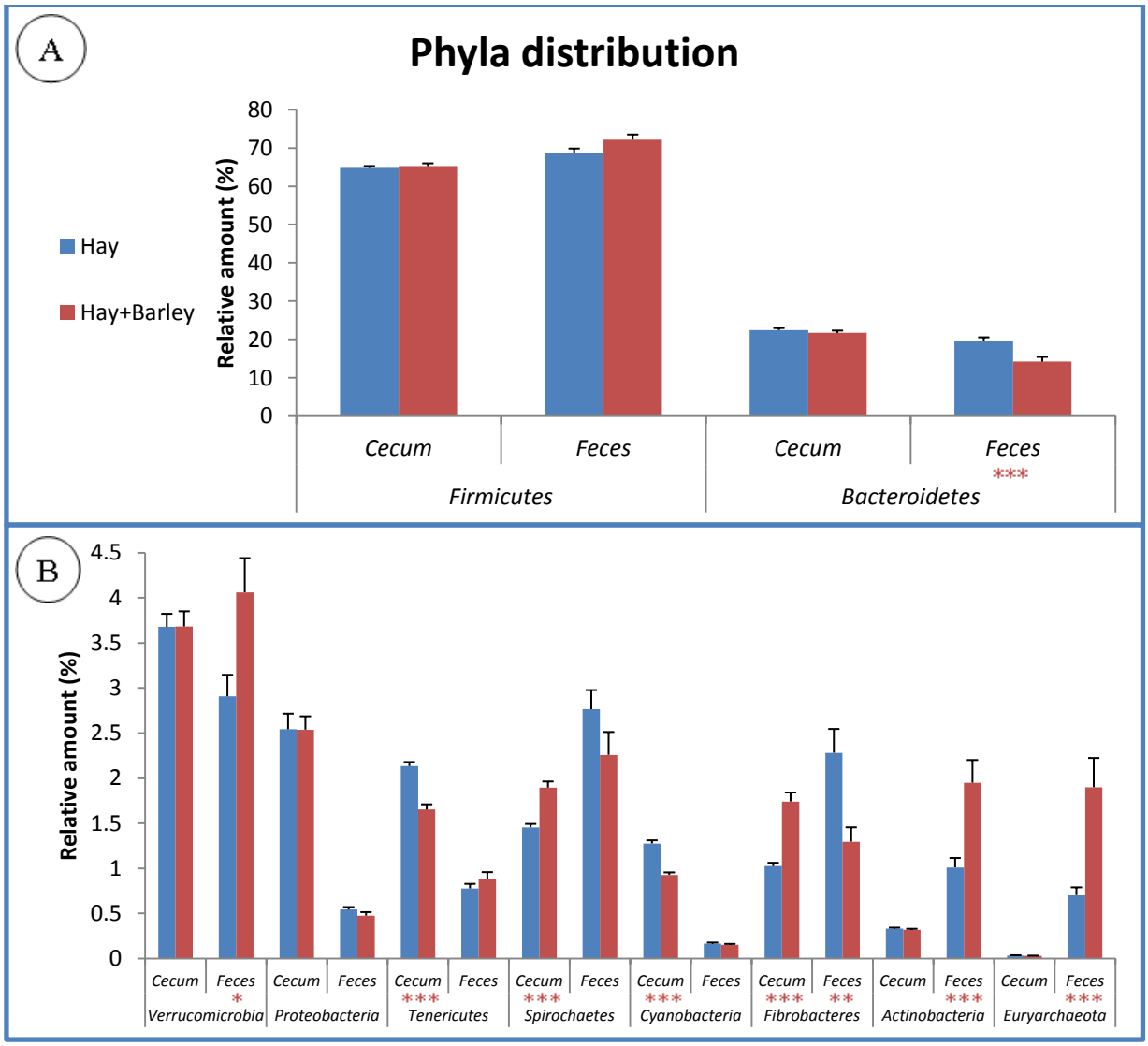


Figure 3-6: Phyla distribution in cecum and feces according to diet (mean + SEM). A) Relative amounts (%) of the most dominant phyla: *Firmicutes* and *Bacteroidetes*. B) Relative amounts (%) of remaining phyla over 1% in either diet or sample origin. *Significant diet difference (*0.010<p<0.050 **0.001<p<0.010 ***p<0.001).

Through PCA analysis in MATLAB, loadings for both the cecum and feces samples were generated (Appendix F). The OTUs with loadings above 0.1 or below -0.1 was selected and summed with other OTUs with the same taxonomy. Student's t-Test, based on diet in cecum samples, was performed on all selected bacterial groups. In addition, a limit of 1% of the total bacterial load in cecum, in either diet, was set as minimum for the bacterium to be considered dominant. A total of six bacterial groups passed the criteria (Table 3-2).

Table 3-2: Selected bacterial groups based on the selection criteria.

Taxonomy[1)]	Phylum	Loadings | Total [2)]	Relative amount [3)]	p-value[4)]
f_Lachnospiraceae	*Firmicutes*	4 | 2056	23.93	0.004
g_Anaerostipes		1 | 18	2.19	<0.001
g_Anaerovibrio		1 | 4	6.08	
g_Streptococcus		2 | 126	4.02	
g_Fibrobacter s_succinogenes	*Fibrobacteres*	1 | 12	3.5	
g_Treponema	*Spirochaetes*	3 | 40	2.78	

1) f = family, g = genus, s = species.

2) Loadings = number of OTUs with loadings above 0.1 or below -0.1 in either cecum or feces samples.
Total = total summed OTUs with same taxonomy.

3) Mean maximum relative amounts (%) of the bacterial groups in cecum detected in either diet.

4) p-values calculated by cecum diet based t-Tests.

The *F. succinogenes* group showed significant diet difference in both cecum and feces samples (cecum: $p<0.001$ feces: $p=0.002$). *F. succinogenes* contents in the cecum remained relatively stable when the horses were fed only hay and consisted of about 1 ± 0.5% of the cecal microbiota (Fig. 3-7A). Greater variation in relative amount of this bacterium was detected for the hay and barley diet, where the average ranged from 0.4% to 3.5%. Almost immediately after the feed reached the cecum, a large peak of *F. succinogenes* occurred. The bacterium reached a maximum of 3.5% at 3 hours (the time point showing most significant diet difference with a p-value of 0.002) and decreased wherefrom to 0.4% at 9 hours. After new feeding at 8 hours, a small peak of this bacterium that reached same quantities as the hay diet occurred. Finally, after feeding at

16 hours, last cycle peak with a maximum of 2.6% relative bacteria amount were reduced close to the hay diet by sampling end. When investigating each single horse separately, they all showed the same trend with two high peaks of *F. succinogenes* amount according to hay plus barley diet. Similar *F. succinogenes* amount was detected in the feces samples but, this bacterium showed no correlating patterns to cecum observations (Fig. 3-7B).

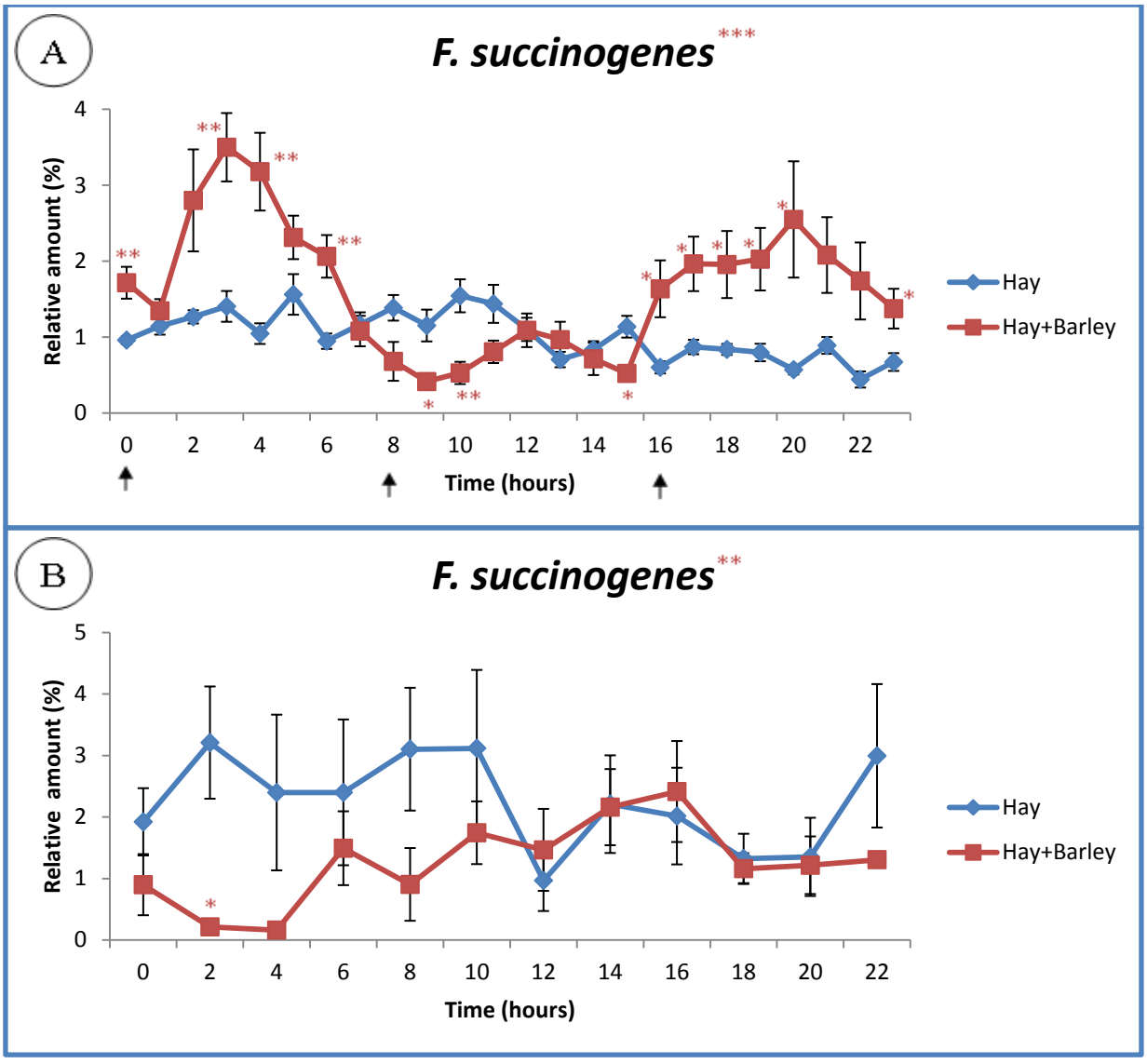


Figure 3-7: Relative amounts of *F. succinogenes* according to diet and different time measurements (mean ± SEM). Time is defined as hours after sampling start/morning meal. A) Cecum samples. Arrows indicate feeding time points in both diets. B) Feces samples. Asterisk associated with the diagram title indicate total significant diet difference, while asterisk associated with the graphs indicate significant diet difference at specific time points (*0.010<p<0.05 **0.001<p<0.010 *** p<0.001).

The S*treptococcus spp*. group also showed significant total diet difference in both cecum and feces samples (both p-values<0.001). Within this group, the group S*treptococcus luteciae* dominated the cecum samples with a maximum in the hay and barley diet with 2.5% relative abundance. A small and relatively stable amount of the *Streptococcus spp*. group was detected in the cecal samples when the horses had been fed only hay (0.2-1.4%) (Fig. 3-8A). However, when the horses were fed hay supplemented with barley, this group showed a peak between time point 5 and 10 with a maximum relative amount of 4% at 8 hours. Cecum correlating patterns of the *Streptococcus spp*. group were detected in the feces samples but, in greater quantities (up to 27%) than of cecum samples (Fig. 3-8B).

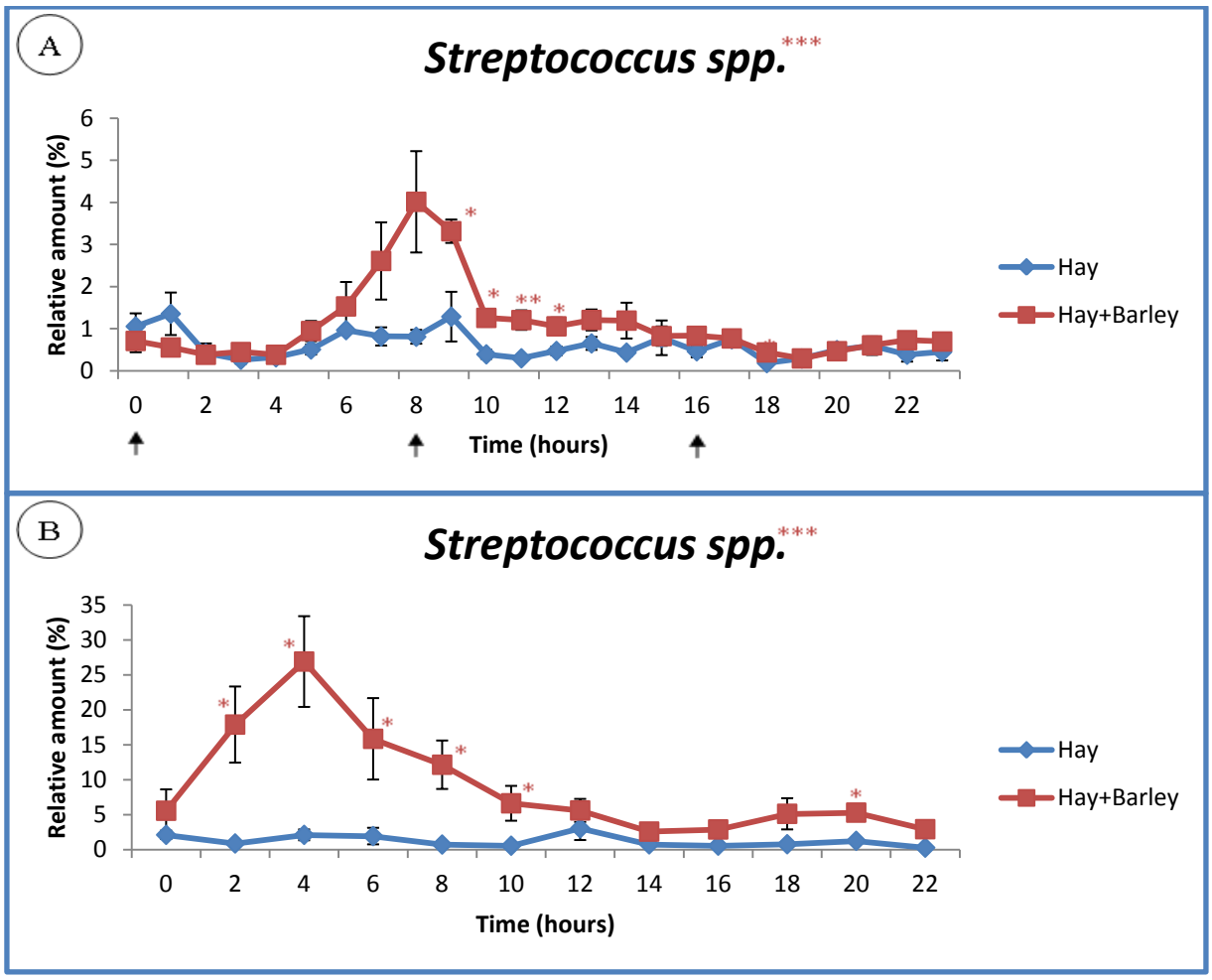


Figure 3-8: Relative amounts of *Streptococcus spp.* according to diet and different time measurements (mean ± SEM). Time is defined as hours after sampling start/morning meal. A) Cecum samples. Arrows indicate feeding time points in both diets. B) Feces samples. Asterisk associated with the diagram title indicate total significant diet difference, while asterisk associated with the graphs indicate significant diet difference at the specific time point (*0.010<p<0.050 **0.001p<0.010 ***p<0.001).

The *Treponema spp*. group showed significant diet difference with a p-value less than 0.001 in the cecum samples and the relative amount of *Treponema spp*. varied (0.7-2.7%), with three distinct peaks when the horses were fed the hay supplemented with barley diet (Fig. 3-9A). The minimum amount of this bacterium appeared to be associated with feeding time point. In contrast, the bacterial amount in hay diet remained stabile throughout the collection period. Larger amounts of the *Treponema spp*. were detected in feces samples (up to 4%) (Fig. 3-9B). Still, this bacteria group showed only significant diet difference at time point 2 (p=0.011) in feces and no correlating pattern with cecum samples. However, ordinary least squares (OLS) regression analysis confirmed linear relationship between these bacteria and *F. succinogenes* (p=0.046) (Appendix G).

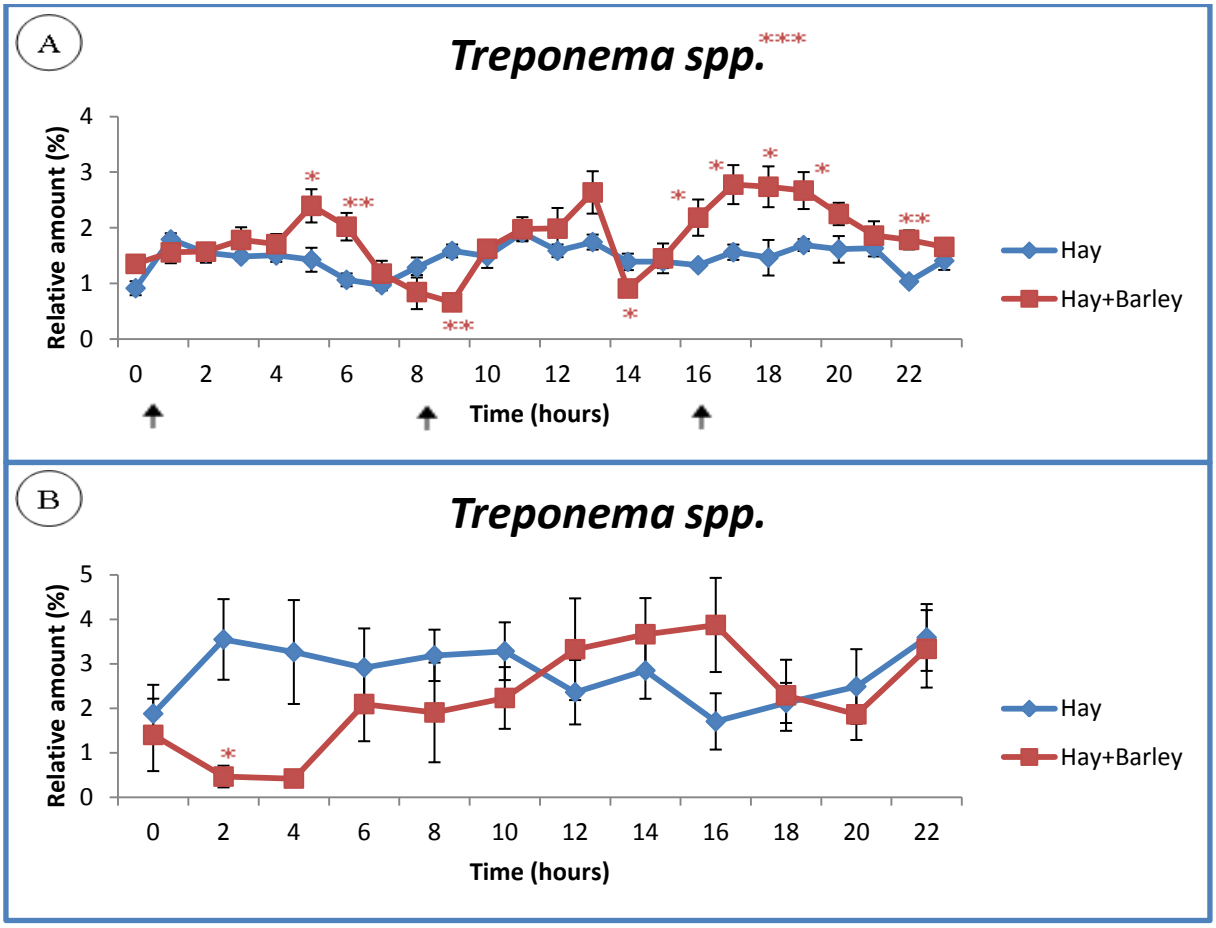


Figure 3-9: Relative amounts of *Treponema spp.* according to diet and different time measurements (mean ± SEM). Time is defined as hours after sampling start/morning meal. A) Cecum samples. Arrows indicate feeding time points in both diets. B) Feces samples. Asterisk associated with the diagram title indicate total significant diet difference, while asterisk associated with the graphs indicate significant diet difference at the specific time point (*0.010<p<0.050 **0.001<p<0.010 ***p<0.001).

The *Anaerovibrio spp.* group showed significant diet difference with a p-values less than 0.001, in both cecum and feces samples. This bacteria group was detected in minute amounts in the cecum, less than 0.1% of total bacteria amount, when the horses were fed only hay (Fig. 3-10A). In contrast, a considerable quantity of the *Anaerovibrio spp*. group was observed trough a large peak between 7 and 17 hours, with a maximum above 6% of total bacteria amount at 14 hours after the morning meal with hay and barley. An increased amount of *Anaerovibrio spp*. in hay plus barley diet were observed as well in the feces samples after 4 hours but, with amounts considered non-dominating (<0.1%) (Fig. 3-10B).

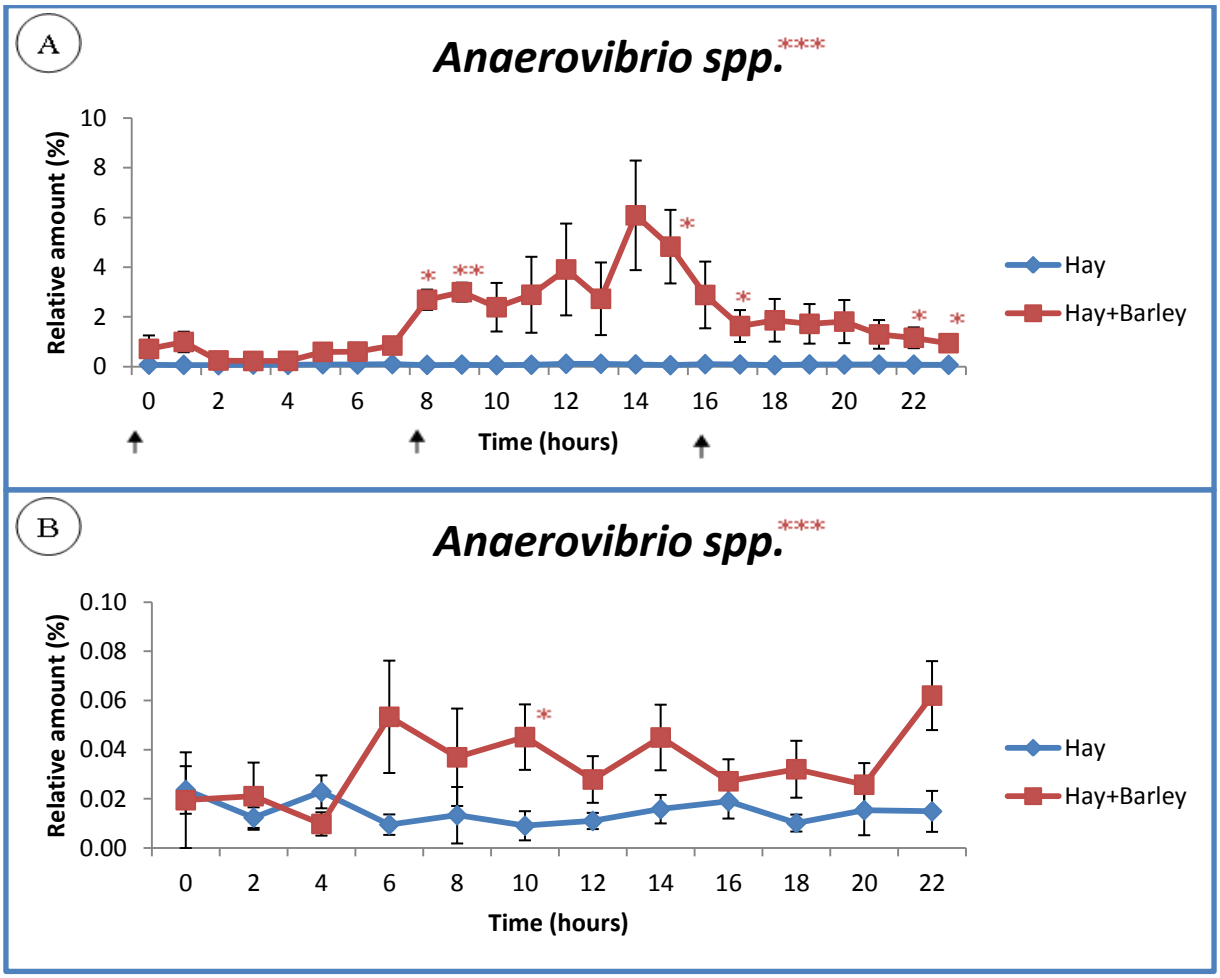


Figure 3-10: Relative amounts of *Anaerovibrio spp.* according to diet and different time measurements (mean ± SEM). Time is defined as hours after sampling start/morning meal. A) Cecum samples. Arrows indicate feeding time points in both diets. B) Feces samples. Asterisk associated with the diagram title indicate total significant diet difference, while asterisk associated with the graphs indicate significant diet difference at the specific time point (*0.010<p<0.050 **p<0.010 ***p<0.001).

The *Lachnospiraceae* family group comprised of up to 24% of the total bacterial load in both cecum and feces samples. The *Lachnospiraceae* family showed a relatively stable trend in both diets (Fig. 3-10), but the *Lachnospiraceae* family amounts, in the two different diets, were significantly different from each other in the cecum (p=0.004). A peak in the time period 2 to 10 hours after the morning meal with hay supplemented with barley was identified in cecum (Fig. 3-11A). In contrast, a lower abundance in the time range 10 to 18 hours was detected in the hay plus barley diet in feces (Fig. 3-11B) with time point 14 and 16 showing significant diet difference (p=0.006 and 0.029 respectively). However, no significant total diet difference was detected in the feces samples and the growth pattern did not correlate to those of the cecum samples.

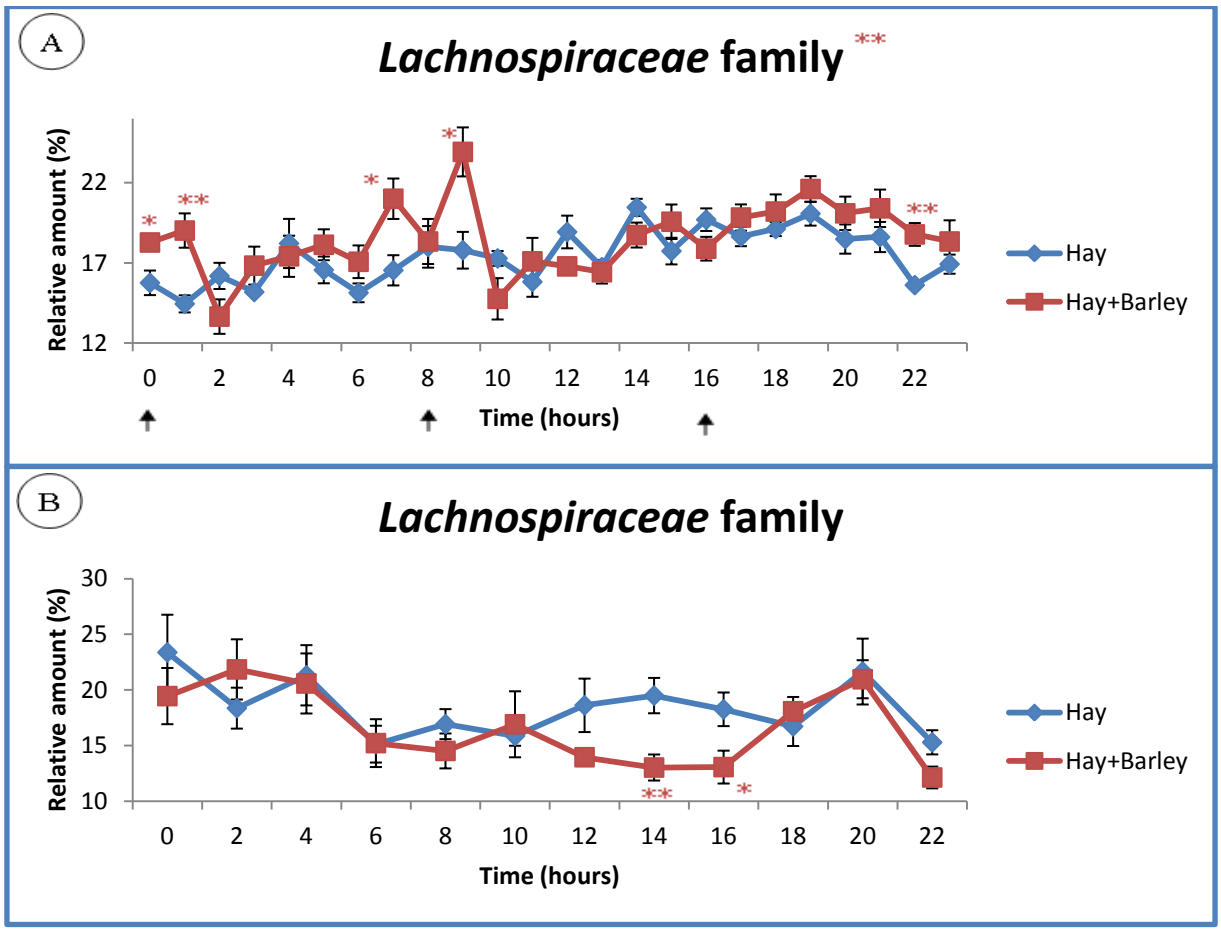


Figure 3-11: Relative amounts of the *Lachnospiraceae* family according to diet and different time measurements (mean ± SEM). Time is defined as hours after sampling start/morning meal. A) Cecum samples. Arrows indicate feeding time points in both diets. B) Feces samples. Asterisk associated with the diagram title indicate total significant diet difference, while asterisk associated with the graphs indicate significant diet difference at the specific time point (*0.010<p<0.050 **0.001<p<0.010 ***p<0.001).

Significantly more *Anaerostipes spp*. was detected in the cecum when the horses ate only hay (0.5-2.2%) compared to when they ate both hay and barley (0.1-0.9%) ($p<0.001$). The detected amount showed relatively continuous curves for both diets in cecum (Fig. 3-12A). However, a decline before feeding with further increases in the proportion of *Anaerostipes spp*. in the cecum after hay-feeding period occurred. *Anaerostipes spp*. still showed significant diet difference ($p=0.030$) in the feces samples although the graphical representation (Fig. 3-12B) shows no major diet difference. However, these bacteria were, in feces, detected in minimal amounts (<0.2%) and were, therefore, considered non-dominating.

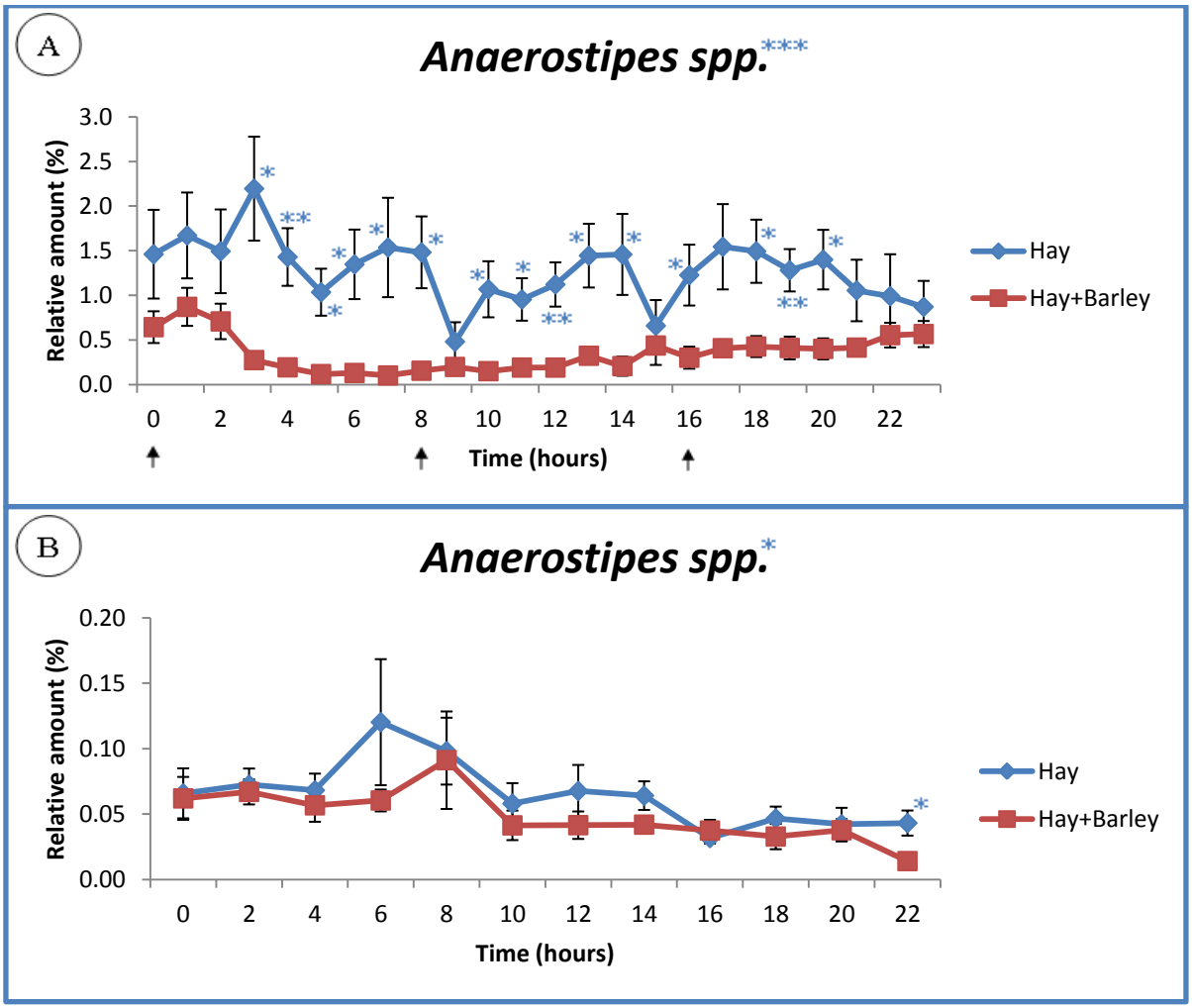


Figure 3-12: Relative amounts of *Anaerostipes spp.* according to diet and different time measurements (mean ± SEM). Time is defined as hours after sampling start/morning meal. A) Cecum samples. Arrows indicate feeding time points in both diets. B) Feces samples. Asterisk associated with the diagram title indicate total significant diet difference, while asterisk associated with the graphs indicate significant diet difference at the specific time point (*$0.010<p<0.050$ **$0.001<p<0.010$ ***$p<0.001$).

3.3.2 Shotgun metagenome analyses

24 cecum samples were selected for shotgun metagenomic sequencing and the sequencing data were analyzed in MG-RAST. Combined, the 24 samples contained a total of 8 507 537 sequences with a mean sequence count of about 350 000 sequence reads including an average read length of 188 bp. Approximately 5% of the sequence reads failed to pass the quality control pipeline of MG-RAST. Of the quality control passed sequences, 0.2-0.8% contained rRNA genes, about 38% contained predicted proteins with known function and about 47% contained predicted proteins with unknown function. Further, roughly 9% of the quality control passed sequence reads had no rRNA gene or predicted proteins.

Taxonomic profile

The taxonomic assignment was carried out by comparing the sequencing data to the M5 non-redundant (M5NR) protein database, which comprise of several sequence databases, in MG-RAST.

The Bacteria (88.82% ± 0.13) [mean ± SEM] dominated the cecal microbiota, while the Eukaryota (0.68% ± 0.03), Archaea (0.46% ± 0.01) and Viruses (0.08% ± 0.01) only were represented in low amounts. 0.01% of the sequences were assigned as other sequences, 9.80% unassigned and 0.15% were unclassified sequences. The dominating bacterial phyla were *Bacteroidetes* (50%) and *Firmicutes* (30%), followed by *Proteobacteria* (7%), *Actinobacteria* (2%), *Verrucomicrobia* (1%), *Fibrobacteres* (1%) (Appendix H). The phyla *Euryarchaeota* dominated the Archaea domain with abundance over 93% of the total detected Archaea. Further, fungi comprised of about 20% of the assigned eukaryotic microorganisms.

Functional profile

Functional profiles were constructed using the Subsystem database which compared homology of functional genes in the sequencing data against the database and displayed annotated genes within the samples.

A total of 28 subsystems were detected, of which 10 subgroups comprised of less than 1% and were thus assembled in the group Other (Sulfur Metabolism, Nitrogen Metabolism, Iron acquisition and metabolism, Metabolism of Aromatic Compounds, Motility and Chemotaxis, Dormancy and sporulation, Secondary Metabolism, Phosphorus Metabolism, Potassium Metabolism, Photosynthesis). Clustering-based subsystems, Carbohydrates and Protein Metabolism were the three most dominating subsystems (Fig. 3-13). The distribution of the functional subsystems was relatively similar in all samples, with no standard deviation of more than 0.3%. However, Membrane Transport and Clustering-based subsystems were significantly different in the two diets (p-values were 0.014 and 0.040 respectively).

Carbohydrate metabolism, the second most abundant subsystem, was divided into 12 subgroups where 8 of them showed significant difference either in diet or in time (Fig. 3-14) (Appendix I: Table A-8). Central carbohydrate metabolism and monosaccharide metabolism dominated the functional subgroups, with both over 15% of the total carbohydrate metabolic features. These two subgroups showed more distinct peak at time point 2 in the hay and barley diet, where central carbohydrate metabolism show a negative trend whereas monosaccharide metabolism show a positive trend. Both CO_2 fixation and fermentation comprised of about 5% of the total carbohydrate metabolic features, and fermentation was the only subgroup that showed significant difference in diet independent of time ($p=0.011$). Both subgroups showed a positive peak in the hay diet while a negative drop in the hay and barley diet. Further, glycoside hydrolases comprised of less than 1% of the total carbohydrate metabolic features. This subgroup showed similar trend in both diets with an increase at time point 2.

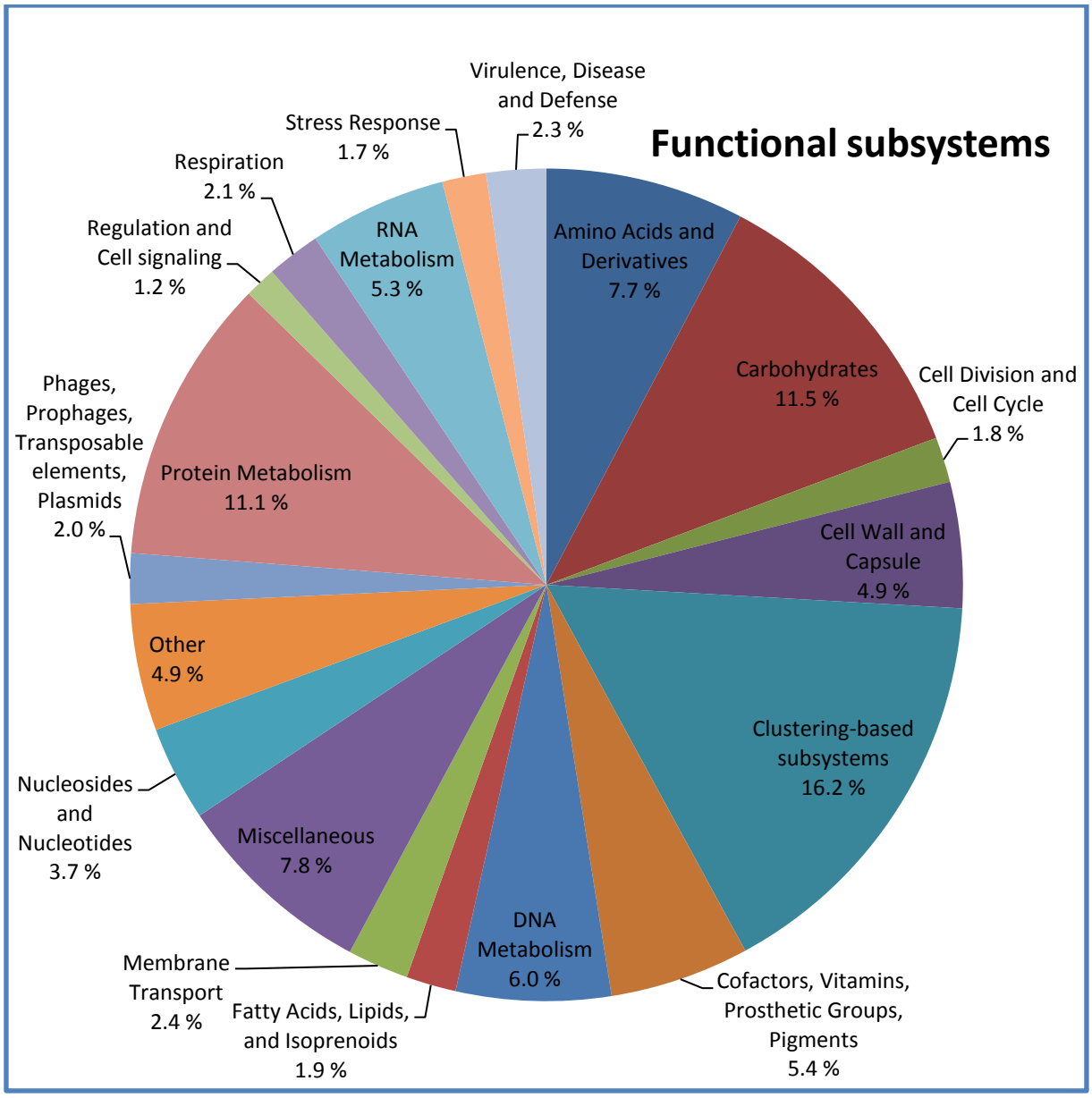


Figure 3-13: Relative distribution of the functional subsystems annotated in MG-RAST (mean).

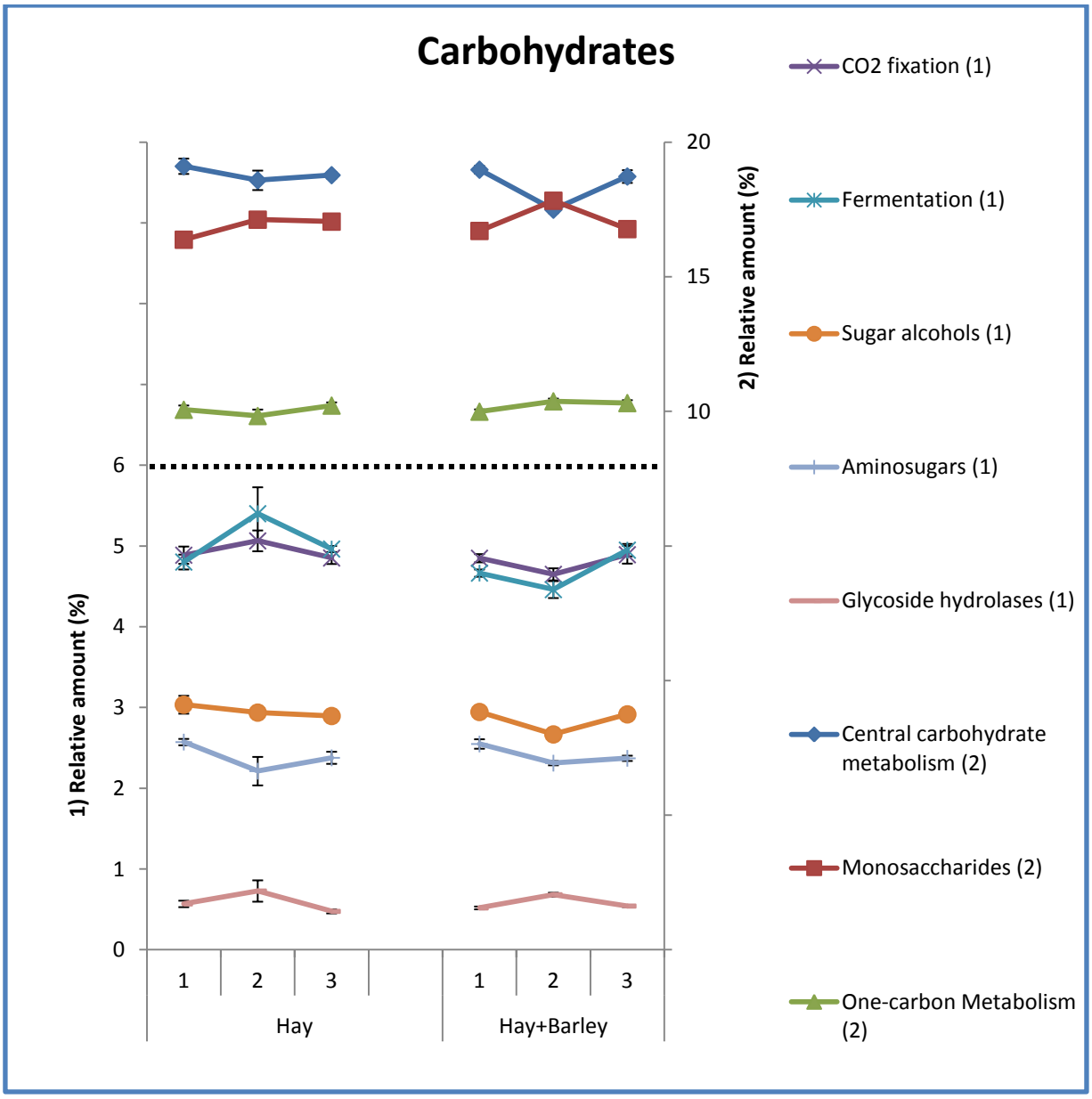


Figure 3-14: Relative sequence amount (%) belonging to carbohydrate related functions (mean ± SEM). CO_2 fixation, Fermentation and Sugar alcohols are plotted against the primary y-axis and, Central carbohydrate metabolism and Monosaccharides are plotted against the secondary y-axis.

3.4 Controls and parallels

3.4.1 qPCR

qPCR with PRK primers was performed primarily to verify the DNA isolation step. The Ct-values were quite similar in all the samples and showed no typical pattern in the different diets or horses. Most of the samples showed Ct-values between 15 and 20. In addition to the samples there were also added a negative isolation control and a negative qPCR control, and both controls showed Ct-values over 30.

3.4.2 16S rRNA gene metagenome sequencing

The PRK and PRKIllumina PCR products showed the expected size (ca. 500 bp) by gel electrophoresis and the pool concentration before dilution for sequencing, measured by Qubit, were 2.7 ng/µl.

Positive and negative controls were added in the PCR reaction and none of the negative PCR controls were sequenced due to any band registration by gel electrophoresis. The species *Escherichia coli* was used as a positive PCR control. The positive controls showed low weighted UniFrac distance between each other (Fig. 3-15), but as much as 379 different OTUs were assigned the two sequenced PCR positive controls. No OTUs had taxonomy as specific as *E. coli*, but 135 OTUs associated with the family *Enterobacteriaceae* accounted for 95.65% and 97.32% of total bacterial distribution respectively. No other OTUs were detected over 0.2% of the total bacterial load in the positive controls. Regression analysis showed that the positive PCR controls had a linear relationship ($R^2 > 0.99$) (Appendix J: Fig. A-8A).

Positive and negative controls were added to every DNA isolation run in order to determine technical errors in the DNA purification and to trace potential contaminations. Positive controls consisted of a sample mix with five randomly selected samples. Cecum mix was used for DNA isolation from the cecum samples and feces mix was used for DNA isolation from feces samples. No DNA was detected in the negative controls. Weighted UniFrac diversity calculations showed short distance between PCR controls plus shorter distance between cecum than of feces controls (Fig. 3-15). Greater distance between cecum (or feces) controls and PCR controls was detected

(>0.8). Furthermore, comparison of within groups and between groups showed significant difference ($p<0.001$). Additionally, regression analysis showed a linear relationship between the sample controls with no cecum controls having lower R^2-value than 0.87 and the minimum R^2-value associated with feces was 0.78 (Appendix J: Fig. A-8B and A-8C).

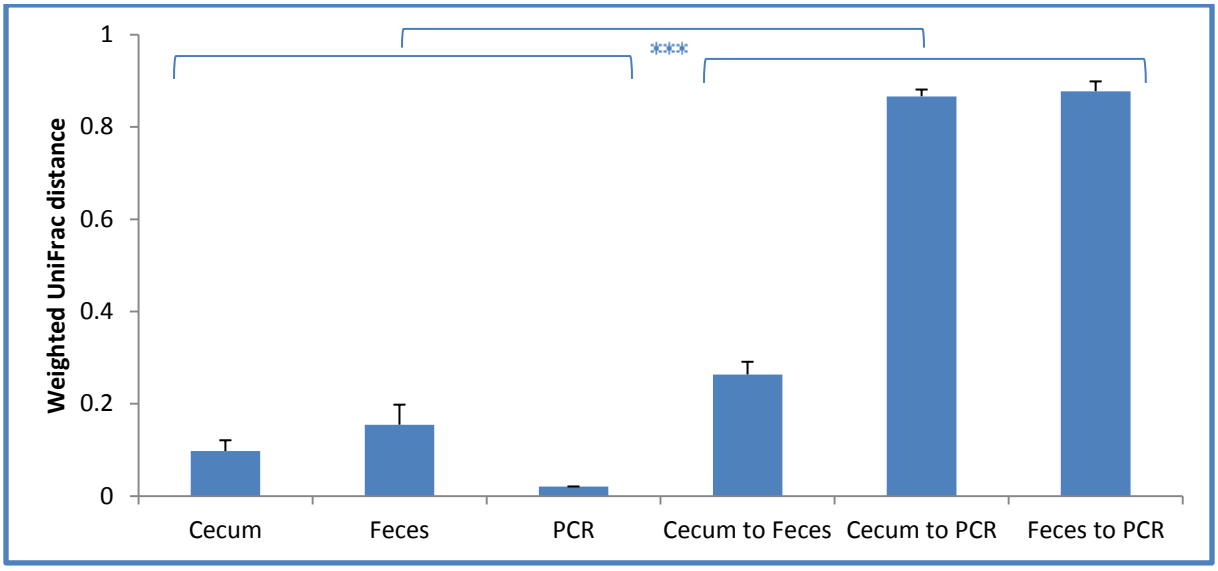


Figure 3-15: Weighted UniFrac diversity index diagram for positive controls (mean + STD). ***p<0.001

Random samples were selected to examine OTU abundance profiles using DNA isolation parallels. The parallels were plotted against each other and the regression analysis showed linear correlation between the parallels ($R^2 > 0.85$).

Through weighted UniFrac plot, the β-diversity according to sampling day showed random sample distribution. Thus, no daily sampling result variation was detected (Appendix D: Fig. A-1B).

3.4.3 16S rRNA gene sequencing and shotgun sequencing comparison

Great variety and significant difference between the methods ($p<0.001$) was detected at bacterial phylum level, except the phylum *Fibrobacteres* (Fig. 3-16). However, the phyla which were detected over 1% of the total bacterial content by shotgun sequencing, was also share of the phyla represented in more than 1% by 16S rRNA gene sequencing.

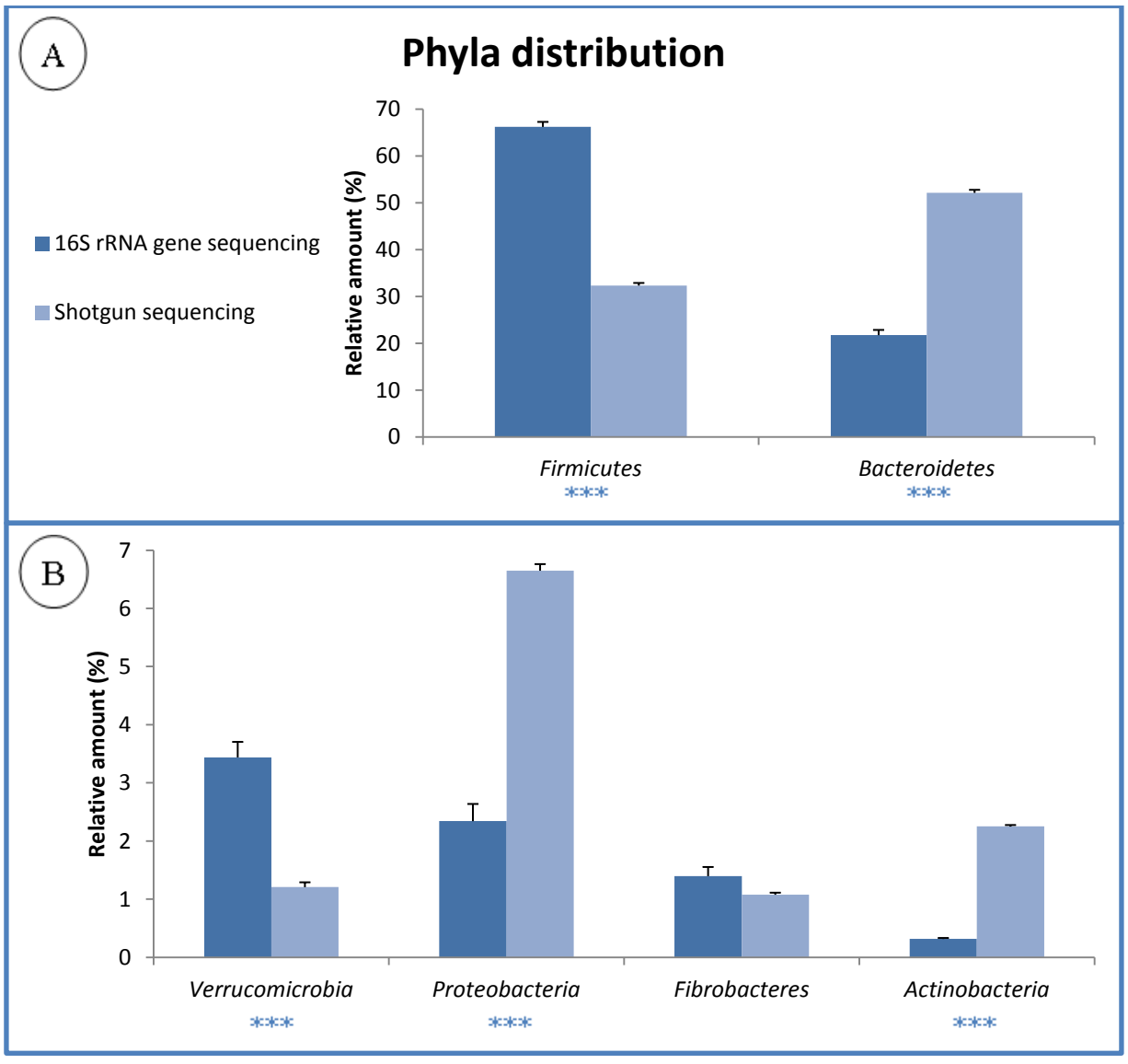


Figure 3-16: Phyla distribution comparison between 16S rRNA gene sequencing and shotgun sequencing in respective samples (mean + SEM). A) Most dominating phyla *Firmicutes* and *Bacteroidetes*. **B)** Remaining common phyla representing more than 1% of the total bacterial content with shotgun sequencing and respective phyla in 16S rRNA gene sequencing. ***Significant difference between the two methods at the specific phyla ($p<0.001$).

4. Discussion

4.1 Feces vs. cecum and horse individuality

Mainly three methods are used to obtain equine intestinal sample material: collection from intestinal cannulated horses, postmortem collection of intestinal content and feces sampling. Fecal samples are often used to investigate the gut microbiota (Schoster et al. 2013) due to difficulty in obtaining other types of intestinal samples (Dougal et al. 2012). This study therefore compared fecal and cecal samples to address whether feces are good representatives for describing the temporal dynamics of the equine cecal microbiota.

Cecum showed distinct difference from feces. The feces showed no clear dietary pH influence, in contrast to cecum. Fecal pH varied, however, between sampling time points and thus the dietary effects might be misinterpreted in studies that only utilize fecal samples. Furthermore, as demonstrated in previous study (Dougal et al. 2012), the microbiota diversity were higher in feces than cecum. Because of the microbiota diversity difference, overestimation of the cecal microbiota diversity will occur when using fecal samples as representatives. Additionally, calculations of which feces and meal that correlates are a complex issue in dietary studies, since passage rate varies in different intestinal compartments depending on a number of feed and animal related factors (Van Weyenberg et al. 2006).

Moreover, the bacterial temporal trends detected in cecum were not supported by the feces findings. Several bacteria detected over 1% of the total cecal microbiota were almost absent in feces. So, when only analyzing feces, these bacteria would not be considered dominant. Interestingly, two of the bacteria showing correlating cecum growth patterns also showed correlating growth patterns in feces, suggesting a symbiotic relationship between these bacteria. The drastic increase of these bacteria observed in cecum however, was absent in feces. Studying different hindgut compartments of five horses, Schoster et al. (2013) concluded that cecum showed most resemblance to feces compared to the other hindgut compartments. Nevertheless, feces bacterial findings in this study did not correspond to findings in cecum. The temporal bacterial growth patterns could to some extent describe some of the findings in this study; however the collection time of the samples must be given great focus.

Individual variation of the gut microbiota composition affects the results of dietary studies. Similar living conditions are beneficial in terms of minimizing individual variation. Despite that the horses used in this study were stabled together, large differences in microbiota composition between the horses were observed. Horses from different stables are then believed to show even greater individual variation. Unlike fecal sampling, postmortem sampling provides limited information about digestive process dynamics and difficulties arise with result comparison due to variation in sampling time after feeding (Julliand et al. 2006). In addition, the large individual differences cause questions to studies using euthanized horses for diet comparison. Basically, the observed significant dietary effects may be caused by individual variation, when comparing diets based on different horses. In this respect, cannulated horses are a great advantage for studying phenomenon that should be unconnected with individuals.

4.2 Taxonomical composition of cecal microbiota

Dougal et al. (2013) reported the gut microbiota in grass/hay fed horses to be dominated by the phyla *Firmicutes* (46%), *Bacteroidetes* (43%) followed by *Fibrobacteres*, *Spirochaetes* and *Proteobacteria*, where all phyla comprised by less than 4% of the total microbiota. Although the *Firmicutes* and *Bacteroidetes* amount in the cecum, in this study according to hay diet, did not correspond to the amounts in the study by Dougal et al. (2013), the same phyla were detected as dominant in both studies. Additionally, this study suggests that *Verrucomicrobia*, *Tenericutes* and *Cyanobacteria* should be added as the dominant microbiota in the cecum as they also comprised of more than 1% of the total microbiota. Strong selection pressure in the gut, where well-adapted microorganisms benefit from regular carbohydrate digestion, provided by the host, and in return give pathogen protection and available nutrients to the host, result in few phyla comprising the gut compared to e.g. different soil ecosystems (Claus et al. 2011). *Firmicutes* and *Bacteroidetes*, the most dominant phyla in the hindgut, were also shown to dominate the gut microbiota of numerous vertebrates including human (Ley et al. 2008). These phyla seem to only grow in the gut and are probably transferred from parents to offspring (Ley et al. 2006). All other detected phyla were also seen in the human intestine with exception of *Tenericutes* and *Fibrobacter* (Ley et al. 2006). *F. succinogenes* (phylum *Fibrobacteres*), one of the three dominant rumen fibrolytic

bacteria, were also previously reported present in the equine hindgut (Julliand et al. 1999; Lin & Stahl 1995). Only *Tenericutes*, *Spirochaetes*, *Cyanobacteria* and *Fibrobacter* showed significant diet effect where *Tenericutes* and *Cyanobacteria* showed highest amounts in the hay diet while *Spirochaetes* and *Fibrobacteres* showed highest amounts in hay and barley diet. Daly et al. (2012) reported a *Fibrobacteres* decrease in concentrate fed horses (0.4%) compared to grass fed horses (2.7%) when studying colonic microbiota of euthanized horses. When considering their colonic samples collection time (between 12 and 16 hours after the last meal), our hay and barley findings correspond although the *Fibrobacteres* amount, when the horses were fed only hay, did not reach their detected amount. Lin and Stahl (1995) detected 12% of the total microbial content in the cecum belonging to the phylum *Fibrobacteres*. The study only comprised one single horse and the high *Fibrobacteres* amount detected were probably due to this particular individual, since this phylum have not shown such dominance in other studies (Daly et al. 2012; Dougal et al. 2013; Julliand et al. 1999).

4.3 Functional characteristics of cecal microbiota

The cecal samples selected for shotgun metagenome analyses, to investigating functional characteristics of the cecal microbiota, showed relatively stable functional subsystem composition. Together with carbohydrates and protein metabolism, clustering-based subsystems dominated the functional subsystems found in the equine hindgut. Clustering-based subsystems and membrane transport were the only first level functional subsystems that showed significant diet difference. The functional characteristics of clustering-based subsystems are not clearly understood (Delmont et al. 2012). However, membrane transport involves all forms of molecular transport between the bacteria and the environment. The membrane transport significant diet difference may indicate an altered ability to excrete e.g. SCFA, which in turn can affect the horse's SCFA uptake.

Carbohydrates, the second most abundant subsystem, involve different features connected with carbohydrate metabolism. Central carbohydrate metabolism includes the main glucose catabolic pathways. Conversely, monosaccharide metabolism includes functions related to degradation and utilization of various monosaccharides, such as xylose, mannose and others. Central carbohydrate metabolism related genes decreased, while monosaccharide metabolism related genes increased

when the horses were fed hay and barley, suggesting that type of carbohydrate fed reflects the microbiota and its functions. Due to higher hemicelluloses and lower cellulose content in the hay and barley diet, these findings may indicate an increased possibility to degrade and utilize hemicelluloses that contain various monomers, instead of cellulose that only consist of glucose monomers linked together.

CO_2 fixation is often connected with photosynthesis in plants. The amounts of CO_2 fixation microbial genes registered can probably be chloroplast derived, since the cecal content also contain plant material. However, these genes may also be linked to the *Cyanobacteria* observed with 16S rRNA gene sequencing. The *Cyanobacteria* in the gut have probably evolved to adapt life in the gastrointestinal tract (Ley et al. 2005).

Fermentation occurs during anaerobic conditions where pyruvate, through substrate phosphorylation, is converted into various end products. Genes connected with carbohydrate fermentation showed decrease when the horses were fed hay and barley, indicating reduced possibility to produce SCFA which was in contrast with the observed result of total SCFA detected. Then again, regulation of these genes may have a major impact on the observed SCFA production.

Glycoside hydrolases are enzymes involved in polysaccharide degradation by hydrolyzing the glycosidic bonds that links the monomers together (Qi et al. 2007; Williams et al. 1984). Genes connected with glycoside hydrolases showed peak at time point 2 in both diets, indicating equal microbiota potential to break down polysaccharides in both diets in addition to increased potential in correspondence with meals.

4.4 Model for cecum bacterial succession

Diet effects on the short-term temporal dynamics of the hindgut microbiota were investigated through 16S rRNA gene metagenome sequencing of cecal samples. The temporal growth patterns of all bacteria remained stable when the horses were fed hay. In contrast, large variations in pH, SCFA and microbiota were detected when the horses additionally ate barley (Fig. 4-1).

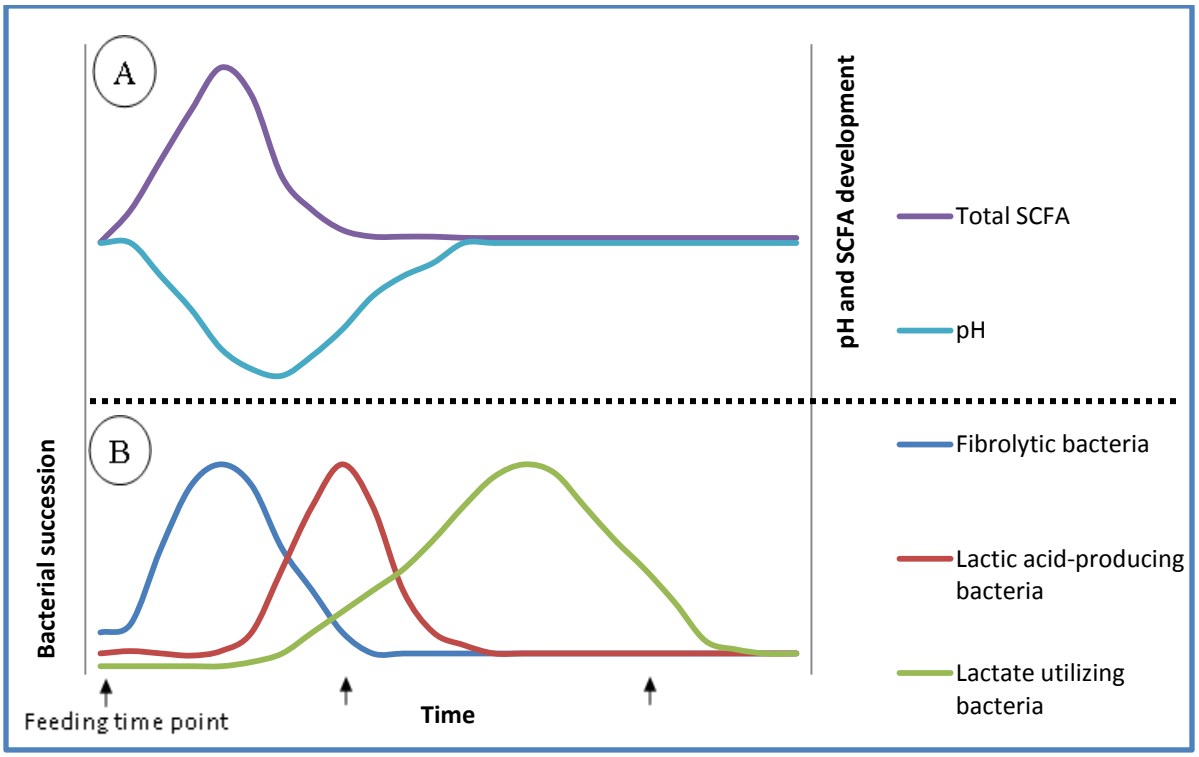


Figure 4-1: Model of the temporal changes of fibrolytic, lactic acid-producing and lactate utilizing bacteria (B) in correspondence with pH and production of SCFA (A) in hay and barley diet. Rapid growth of fibrolytic bacteria cause increase in SCFA production leading to decreased cecum pH. Acidic environment favor growth of lactic acid producing bacteria that result in further pH decrease. Finally, accumulation of lactate facilitates growth of lactate utilizing bacteria that reduce lactate amounts in the cecum and stabilize the pH.

The cecal microbiota showed great variation during the 24 hours after hay and barley feeding. *F. succinogenes* ferment cellulose and produce SCFA (Daly et al. 2012; Stewart & Flint 1989). The drastic observed increase of *F. succinogenes* may have led to rapid fermentation and increased total SCFA production, which in turn led to pH decrease (Al Jassim et al. 2005). Acidic environment may have favored rapid growth of lactic acid producing bacteria, like *Streptococcus spp*. that ferment starch (Daly et al. 2012), which increased lactate amounts in the cecum and led to further pH reduction (Al Jassim et al. 2005). Acidic cecum environment has previously been shown to cause disease, like laminitis (Katz & Bailey 2012), indicating that pH drops are undesirable. The pH drop was absent when the horses only eat hay. However, the main reason for concentrate feeding is the horse's energy requirement. Feeding order like concentrates before hay, and vice versa has shown to provide no significant pH effect. Thus, to avoid pH drop, the starch content of the feed ration must be reduced (Jensen et al. 2012).

F. succinogenes has an optimum growth pH at 6.8 (Miyazaki et al. 1992) and due to acid intolerance (Stewart & Flint 1989), growth may have been suppressed in compliance with pH decrease. It is however unclear why the quantity of *F. succinogenes* was not detected at higher level in hay diet, since cecum pH during the entire process remained around 6.8 and the total fiber content was higher than that of hay and barley diet. Since hemicelluloses are considered to be more easily degradable than cellulose, higher hemicellulose-content of the hay and barley diet could be suggested as a probable explanation for the drastic growth of *F. succinogenes.* Previous *in vitro* studies have shown that *F. succinogenes* can only break down hemicelluloses in order to access more cellulose (Suen et al. 2011), but it cannot utilize them (Osborne & Dehority 1989; Suen et al. 2011). However, it cannot be ruled out that this bacterium behaves differently *in vivo* than *in vitro*.

At the same time, the cause of the large increase in *F. succinogenes* may be caused by a symbiotic relationship promoting growth and cellulose utilization. *Bacteroides ruminicola* H8a, a hemicellulolytic bacteria, have previously shown to increase the total cellulose digestion when co-cultured with *F. succinogenes* (Dehority & Scott 1967). The *Spirochaetaceae* family produces acetate by using H_2/CO_2 as substrate for acetogenesis (Santos et al. 2011) and are expected be involved in SCFA production in the equine cecum and colon (Steelman et al. 2012). The *Treponema spp*. showed similar growth pattern as *F. succinogenes* with relatively stable growth in hay diet and three peaks, after each feeding, in hay and barley diet. In a *Treponema bryantii* and *F. succinogenes* co-culturing study, Stanton and Canale-Parola (1980) showed that the saccharolytic spirochete grew in cellulose containing media, though *T. bryantii* had not shown any cellulolytic activity. They also reported that *Treponema* enhanced cellulose breakdown by *F. succinogenes* suggesting a symbiotic relationship between these bacteria. They proposed that *Treponema spp*. in the rumen are directed, by chemotaxis, to soluble carbohydrates released from cellulose breakdown by *F. succinogenes* and use these sugars as essential fermentable substrates. On the other hand, *F. succinogenes* are non-motile and *Treponema*'s high motility may randomly push *F. succinogenes* to new cellulose substrates, leading to increased cellulolytic activity. It is likely that this symbiotic event also may occur in the equine cecum and may explain the similar growth patterns for these bacteria.

Cellulolysis in the rumen have been shown to be reduced when rumen pH decrease due to rapid starch fermentation (Stewart & Flint 1989) which may also be true in the equine hindgut. In addition, increased amount of lactic acid inhibits absorption of SCFA in the equine intestines (Daly et al. 2001). *Anaerovibrio spp*., belonging to the *Veillonellaceae* family, utilizes lactate and are acid tolerant bacteria that can survive pH drops (Biddle et al. 2013). The large presence of these bacteria in the hay and barley diet, and absence in the hay diet, suggests a high lactic acid concentration, due to *Streptococcus spp.,* in the cecum at the hay and barley diet. Such lactate utilizers are needed to prevent lactate accumulation over prolonged periods of time in the cecum during starch fermentation (Mackie & Gilchrist 1979).

Species within the genus *Anaerostipes* may produce butyrate (Schwiertz et al. 2002) through pathways including butyryl-CoA:acetate CoA-transferase (Meehan & Beiko 2014), which by the enzyme name indicate the need of available acetate (Duncan et al. 2002). The increase in the proportion of *Anaerostipes spp*. after hay-feeding period may indicate that the bacterial load is affected by the new meals. The peaks observed in the hay diet may be caused by utilization of acetate produced by other bacteria in the cecum. No correlation between the growth of these bacteria and butyrate measurements were observed, suggesting that these bacteria contribute only to maintain butyrate production in general and the observed butyrate production increase, when the horses were fed hay and barley, was caused by other butyrate-producing bacteria. The *Lachnospiraceae* family contains several species that produce butyrate (Meehan & Beiko 2014). The relative amount of this family increased in the time range 2-10 hours after the horses were fed hay and barley, which correspond with the measured butyrate peak, suggesting that the increased butyrate production were caused by other family members of the *Anaerostipes*.

4.5 Technical evaluation

The next-generation sequencing process has revolutionized biological research by producing large sequencing data sets within short period of time (Dohm et al. 2008). However, these sequencing methods are not error free. Dephasing may occur during the sequencing process by incomplete extension or addition of multiple nucleotides. As a result, further incorporated nucleotides will no longer be in synchronized position (Metzker 2010) leading to increased

fluorescence noise, base-calling errors and shorter reads (Erlich et al. 2008). For Illumina, substitution is the most common error, often in connection with guanine incorporation (Dohm et al. 2008). Additionally, an underrepresentation of AT-rich and GC-rich regions has been observed, possibly occurring during template preparation by the amplification step (Metzker 2010).

Fibrolytic bacteria are often closely associated with the substrate surfaces (Flint et al. 2008; Krause et al. 2003; Suen et al. 2011). Although all samples in this study were vortexed before pipetting in connection with DNA isolation, consideration must be given to this phenomenon and an underestimation of these fibrolytic bacteria may have occurred.

Differences in the taxonomical microbiota distribution were observed by the two sequencing methods used in this thesis. Possible phyla discrimination of the PRK primers used for 16S rRNA gene sequencing cannot be excluded, since they were initially produced to only examine methanogenic bacteria and archaea. Previous testing of these primers showed a matching efficiency of about 87% (Yu et al. 2005), which suggests that certain bacterial and archaeal groups are not captured by these primers. The two sequenced positive PCR controls, *E. coli,* were assigned as much as 379 different OTUs and none as specific as *E. coli*. However, 95.65% and 97.32% of the OTUs belonged to the *Enterobacteriaceae* family in respective PCR controls. The positive controls should in principle be pure bacterial cultures and the large amount of nonspecific OTUs assigned the positive controls may indicate an overestimation of OTUs from QIIME (Edgar 2013). Regression analysis showed a linear relationship between the two parallels and they additionally show short weighted UniFrac distance between each other, which gives reason to believe in a systematic selection of OTUs by QIIME. The phyla distribution designed from 16S rRNA gene sequencing and shotgun sequencing demonstrated different percentage division of the various phyla. Small differences in the taxonomic profile were found in the metagenomic shotgun data, probably due to non-specific classification of organisms by MG-RAST. Many sequences were unclassified and will hence not show the difference in the same manner as 16S rRNA gene sequencing. However, most phyla demonstrated dominance by both methods. In addition, previous studies have shown that MG-RAST performs poorer than QIIME according to bacterial classification (D'Argenio et al. 2014). Therefore, the 16S rRNA gene

sequencing results concludes to be more trust worthy. 16S rRNA gene sequencing data analyzed in QIIME are predicted by comparing this specific gene to the Greengenes database (Caporaso et al. 2010), while the shotgun metagenomic sequencing data analyzed in MG-RAST are annotated to proteins and wherefrom predicted to taxonomical classification (Meyer et al. 2008). The data were generated in different manner and may, thus, result in different taxonomic distribution. However, all sample origin, horse and diet comparisons were conducted using the same method, which means that we still believe in the observed significant effects.

4.6 Future work

The hindgut microbial community is probably more complex than what has been described in this thesis and collaboration between different bacteria cannot be excluded. Additionally, despite low abundance of specific bacterial phylum they may be essential, thus further investigation of these bacterial groups may also be appropriate. The cannulated horses, at IHA, provide a golden opportunity to further work toward gaining more understanding about the bacterial symbiotic relationship of the equine hindgut microbiota and how this microbiota affects its host.

Concentrate feeding lowered the microbiota diversity. Reduced gut microbiota diversity has previously shown connection to different human diseases (Lozupone et al. 2012; Rook 2013) due to poor inflammation control by the immune system (Rook 2013). Probably, the microbiota diversity may affect the horse in the same way. Microbiota diversity seems to coincide with pH, since the horse with largest pH drop had the least diverse cecal microbiota, whereas the horse with the smallest drop in pH showed highest diversity. Additionally, the horse showing the smallest pH drop ate the barley much slower than the other horses. Consequently, this indicates that prolonged concentrate ingestion may help maintaining a more stable pH as well as microbiota. However, this observation concerned only one single horse and prolonged ingestion time by feeding machines may be a suggested study design in the future. Additionally, in future studies investigating diet effects on the temporal bacterial development, lactate measurements should be included in order to gain more validation of the course of bacterial events.

During altered types of carbohydrate available in the cecum, several bacteria have the ability to change carbohydrate utilization pathways and due to their adaptability, their growth patterns will be more stable. To investigate these bacteria and their active genes in a greater extent, metatranscriptomics may be suggested to provide a more complete understanding of the actual events of this complex community in the equine cecum. RNA analyzes often provide difficulties in practical work because of RNA's rapid degradation time. However, because the samples are collected straight from the cecum, the samples can be directly preserved allowing opportunities for metatranscriptomics.

5. Conclusion

This study addressed the diet effects on the equine hindgut microbiota, where the microbiota composition showed clear difference between the two conducted diets. The microbiota composition remained stable when the horses were fed only hay, whereas showed great variation when the horses were fed barley in addition to hay. However, the functional traits of the hindgut microbiota remained relatively stable in both diets. This study also demonstrated the temporal changes of the equine hindgut microbiota and suggests a course of events where fibrolytic microorganisms that produce SCFA leads to pH decrease in the cecum and the revival of lactic acid producing bacteria, where ultimately lactate utilizing bacteria prevents accumulation of lactate in the cecum over prolonged periods of time when the horse is fed hay and barley. The equine cecum samples compared to feces samples showed few cecum bacterial findings corresponding to the findings in the feces samples. The temporal bacterial growth patterns could to some extent describe some of the findings in this study; however the collection time of the samples must be given great focus. Ultimately, with this study we claim that feces provide little description of the cecal microbiota temporal dynamics, and we therefore do not recommend using fecal samples as a proxy for cecum. The understanding of the horse’s hindgut microbiota is still not complete, but this study provides a good foundation for further studies to gain a more complete understanding of this complex bacterial community.

References

Ahn, J. Y., Yang, L. Y., Paster, B. J., Ganly, I., Morris, L., Pei, Z. H. & Hayes, R. B. (2011). Oral Microbiome Profiles: 16S rRNA Pyrosequencing and Microarray Assay Comparison. *Plos One*, 6 (7).

Al Jassim, R. A., Scott, P. T., Trebbin, A. L., Trott, D. & Pollitt, C. C. (2005). The genetic diversity of lactic acid producing bacteria in the equine gastrointestinal tract. *FEMS Microbiol Lett*, 248 (1): 75-81.

Ansorge, W. J. (2009). Next-generation DNA sequencing techniques. *New Biotechnology*, 25 (4): 195-203.

Austbø, D. & Volden, H. (2006). Influence of passage model and caecal cannulation on estimated passage kinetics of roughage and concentrate in the gastrointestinal tract of horses. *Livestock Science*, 100 (1): 33-43.

Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., et al. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456 (7218): 53-59.

Biddle, A. S., Black, S. J. & Blanchard, J. L. (2013). An In Vitro Model of the Horse Gut Microbiome Enables Identification of Lactate-Utilizing Bacteria That Differentially Respond to Starch Induction. *Plos One*, 8 (10).

Brøkner, C., Austbø, D., Næsset, J. A., Knudsen, K. E. B. & Tauson, A. H. (2012). Equine pre-caecal and total tract digestibility of individual carbohydrate fractions and their effect on caecal pH response. *Archives of Animal Nutrition*, 66 (6): 490-506.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Pena, A. G., Goodrich, J. K., Gordon, J. I., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7 (5): 335-336.

Claus, S. P., Ellero, S. L., Berger, B., Krause, L., Bruttin, A., Molina, J., Paris, A., Want, E. J., de Waziers, I., Cloarec, O., et al. (2011). Colonization-induced host-gut microbial metabolic interaction. *MBio*, 2 (2): e00271-10.

Costa, M. C. & Weese, J. S. (2012). The equine intestinal microbiome. *Animal Health Research Reviews*, 13 (01): 121-128.

D'Argenio, V., Casaburi, G., Precone, V. & Salvatore, F. (2014). Comparative metagenomic analysis of human gut microbiome composition using two different bioinformatic pipelines. *Biomed Res Int*, 2014: 325340.

Daly, K., Stewart, C. S., Flint, H. J. & Shirazi-Beechey, S. P. (2001). Bacterial diversity within the equine large intestine as revealed by molecular analysis of cloned 16S rRNA genes. *FEMS Microbiology Ecology*, 38 (2-3): 141-151.

Daly, K., Proudman, C. J., Duncan, S. H., Flint, H. J., Dyer, J. & Shirazi-Beechey, S. P. (2012). Alterations in microbiota and fermentation products in equine large intestine in response to dietary variation and intestinal disease. *British Journal of Nutrition*, 107 (7): 989-995.

Davenport, C. F. & Tummler, B. (2013). Advances in computational analysis of metagenome sequences. *Environ Microbiol*, 15 (1): 1-5.

Dehority, B. A. & Scott, H. W. (1967). Extent of Cellulose and Hemicellulose Digestion in Various Forages by Pure Cultures of Rumen Bacteria. *Journal of Dairy Science*, 50 (7): 1136-&.

Delmont, T. O., Prestat, E., Keegan, K. P., Faubladier, M., Robe, P., Clark, I. M., Pelletier, E., Hirsch, P. R., Meyer, F., Gilbert, J. A., et al. (2012). Structure, fluctuation and magnitude of a natural grassland soil metagenome. *ISME J*, 6 (9): 1677-87.

Dohm, J. C., Lottaz, C., Borodina, T. & Himmelbauer, H. (2008). Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Research*, 36 (16).

Dougal, K., Harris, P. A., Edwards, A., Pachebat, J. A., Blackmore, T. M., Worgan, H. J. & Newbold, C. J. (2012). A comparison of the microbiome and the metabolome of different regions of the equine hindgut. *FEMS Microbiol Ecol*, 82 (3): 642-52.

Dougal, K., de la Fuente, G., Harris, P. A., Girdwood, S. E., Pinloche, E. & Newbold, C. J. (2013). Identification of a Core Bacterial Community within the Large Intestine of the Horse. *Plos One*, 8 (10).

Duncan, S. H., Barcenilla, A., Stewart, C. S., Pryde, S. E. & Flint, H. J. (2002). Acetate utilization and butyryl coenzyme A (CoA):acetate-CoA transferase in butyrate-producing bacteria from the human large intestine. *Appl Environ Microbiol*, 68 (10): 5186-90.

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods*, 10 (10): 996-8.

Erlich, Y., Mitra, P. P., delaBastide, M., McCombie, W. R. & Hannon, G. J. (2008). Alta-Cyclic: a self-optimizing base caller for next-generation sequencing. *Nat Methods*, 5 (8): 679-82.

Espy, M. J., Uhl, J. R., Sloan, L. M., Buckwalter, S. P., Jones, M. F., Vetter, E. A., Yao, J. D., Wengenack, N. L., Rosenblatt, J. E., Cockerill, F. R., 3rd, et al. (2006). Real-time PCR in clinical microbiology: applications for routine laboratory testing. *Clin Microbiol Rev*, 19 (1): 165-256.

Flint, H. J., Bayer, E. A., Rincon, M. T., Lamed, R. & White, B. A. (2008). Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nature Reviews: Microbiology*, 6 (2): 121-131.

Forde, B. M. & O'Toole, P. W. (2013). Next-generation sequencing technologies and their impact on microbial genomics. *Briefings in Functional Genomics*, 12 (5): 440-453.

Hagen, L. H. (2012). *Gassproduksjon og struktur av det mikrobielle samfunnet i labskala biogassreaktorer - effekt av storfegjødsel og ensilert fiskeavfall*: Universitetet for miljø- og biovitenskap, IKBM.

Harris, T. D., Buzby, P. R., Babcock, H., Beer, E., Bowers, J., Braslavsky, I., Causey, M., Colonell, J., DiMeo, J., Efcavitch, J. W., et al. (2008). Single-Molecule DNA Sequencing of a Viral Genome. *Science*, 320 (5872): 106-109.

Illumina. (2012). *Nextera® XT DNA Sample Preparation Guide*. Part # 15031942 Rev. C.

Jansson, A. & Lindberg, J. E. (2012). A forage-only diet alters the metabolic response of horses in training. *Animal*, 6 (12): 1939-46.

Jensen, R. B., Austbø, D. & Tauson, A. H. (2012). Feeding forage before or after oats affects caecum pH profiles of the horse. In Saastamoinen, M., Fradinho, M., Santos, A. & Miraglia, N. (eds) Forages and grazing in horse nutrition, vol. 132 *Forages and grazing in horse nutrition*, pp. 327-330: Wageningen Academic Publishers.

Jordhøy, F. (2012). *Det mikrobielle samfunnet ved biogassproduksjon. Et taksanomisk studie av mikroorganismer tilstede ved anaerob nedbrytning av storfegjødsel og fiskeensilasje*: Universitetet for miljø- og biovitenskap, IKBM.

Julliand, V., de Vaux, A., Millet, L. & Fonty, G. (1999). Identification of Ruminococcus flavefaciens as the predominant cellulolytic bacterial species of the equine cecum. *Appl Environ Microbiol*, 65 (8): 3738-41.

Julliand, V., De Fombelle, A. & Varloud, M. (2006). Starch digestion in horses: The impact of feed processing. *Livestock Science*, 100 (1): 44-52.

Katz, L. M. & Bailey, S. R. (2012). A review of recent advances and current hypotheses on the pathogenesis of acute laminitis. *Equine Vet J*, 44 (6): 752-61.

Krause, D. O., Denman, S. E., Mackie, R. I., Morrison, M., Rae, A. L., Attwood, G. T. & McSweeney, C. S. (2003). Opportunities to improve fiber degradation in the rumen: microbiology, ecology, and genomics. *FEMS Microbiol Rev*, 27 (5): 663-93.

Kubista, M., Andrade, J. M., Bengtsson, M., Forootan, A., Jonák, J., Lind, K., Sindelka, R., Sjöback, R., Sjögreen, B., Strömbom, L., et al. (2006). The real-time polymerase chain reaction. *Molecular Aspects of Medicine*, 27 (2–3): 95-125.

Ley, R. E., Backhed, F., Turnbaugh, P., Lozupone, C. A., Knight, R. D. & Gordon, J. I. (2005). Obesity alters gut microbial ecology. *Proc Natl Acad Sci U S A*, 102 (31): 11070-5.

Ley, R. E., Peterson, D. A. & Gordon, J. I. (2006). Ecological and Evolutionary Forces Shaping Microbial Diversity in the Human Intestine. *Cell*, 124 (4): 837-848.

Ley, R. E., Lozupone, C. A., Hamady, M., Knight, R. & Gordon, J. I. (2008). Worlds within worlds: evolution of the vertebrate gut microbiota. *Nature Reviews Microbiology*, 6 (10): 776-788.

Lin, C. Z. & Stahl, D. A. (1995). Taxon-Specific Probes for the Cellulolytic Genus Fibrobacter Reveal Abundant and Novel Equine-Associated Populations. *Applied and Environmental Microbiology*, 61 (4): 1348-1351.

Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature*, 489 (7415): 220-230.

Macfarlane, S. & Macfarlane, G. T. (2003). Regulation of short-chain fatty acid production. *Proc Nutr Soc*, 62 (1): 67-72.

Mackay, I. M., Arden, K. E. & Nitsche, A. (2002). Real-time PCR in virology. *Nucleic Acids Research*, 30 (6): 1292-1305.

Mackie, R. I. & Gilchrist, F. M. (1979). Changes in Lactate-Producing and Lactate-Utilizing Bacteria in Relation to pH in the Rumen of Sheep During Stepwise Adaptation to a High-Concentrate Diet. *Appl Environ Microbiol*, 38 (3): 422-30.

Mackie, R. I. & Wilkins, C. A. (1988). Enumeration of anaerobic bacterial microflora of the equine gastrointestinal tract. *Appl Environ Microbiol*, 54 (9): 2155-60.

Meaburn, E. & Schulz, R. (2012). Next generation sequencing in epigenetics: Insights and challenges. *Seminars in Cell & Developmental Biology*, 23 (2): 192-199.

Meehan, C. J. & Beiko, R. G. (2014). A Phylogenomic View of Ecological Specialization in the Lachnospiraceae, a Family of Digestive Tract-associated Bacteria. *Genome Biology and Evolution*.

Metzker, M. L. (2005). Emerging technologies in DNA sequencing. *Genome Res*, 15 (12): 1767-76.

Metzker, M. L. (2010). Sequencing Technologies - the Next Generation. *Nature Reviews Genetics*, 11 (1): 31-46.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., et al. (2008). The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*, 9: 386.

Miyazaki, K., Hino, T. & Itabashi, H. (1992). Effects of Extracellular Ph on the Intracellular Ph and Membrane-Potential of Cellulolytic Ruminal Bacteria, Ruminococcus-Albus,

Ruminococcus-Flavefaciens, and Fibrobacter-Succinogenes. *Journal of General and Applied Microbiology*, 38 (6): 567-573.

Nelson, M. C., Morrison, H. G., Benjamino, J., Grim, S. L. & Graf, J. (2014). Analysis, Optimization and Verification of Illumina-Generated 16S rRNA Gene Amplicon Surveys. *PLoS One*, 9 (4): e94249.

O' Donnell, M. M., Harris, H. M. B., Jeffery, I. B., Claesson, M. J., Younge, B., O' Toole, P. W. & Ross, R. P. (2013). The core faecal bacterial microbiome of Irish Thoroughbred racehorses. *Letters in Applied Microbiology*.

Olsen, G. J. & Woese, C. R. (1993). Ribosomal-RNA - a key to phylogeny. *Faseb Journal*, 7 (1): 113-123.

Osborne, J. M. & Dehority, B. A. (1989). Synergism in degradation and utilization of intact forage cellulose, hemicellulose, and pectin by three pure cultures of ruminal bacteria. *Appl Environ Microbiol*, 55 (9): 2247-50.

Petrosino, J. F., Highlander, S., Luna, R. A., Gibbs, R. A. & Versalovic, J. (2009). Metagenomic Pyrosequencing and Microbial Identification. *Clinical Chemistry*, 55 (5): 856-866.

Qi, M., Jun, H. S. & Forsberg, C. W. (2007). Characterization and synergistic interactions of Fibrobacter succinogenes glycoside hydrolases. *Appl Environ Microbiol*, 73 (19): 6098-105.

Rook, G. A. (2013). Regulation of the immune system by biodiversity from the natural environment: an ecosystem service essential to health. *Proc Natl Acad Sci U S A*, 110 (46): 18360-7.

Rosenfeld, I., Austbø, D. & Volden, H. (2006). Models for estimating digesta passage kinetics in the gastrointestinal tract of the horse. *Journal of Animal Science*, 84 (12): 3321-3328.

Ruijter, J. M., Ramakers, C., Hoogaars, W. M., Karlen, Y., Bakker, O., van den Hoff, M. J. & Moorman, A. F. (2009). Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res*, 37 (6): e45.

Santos, A. S., Rodrigues, M. A. M., Bessa, R. J. B., Ferreira, L. M. & Martin-Rosset, W. (2011). Understanding the equine cecum-colon ecosystem: current knowledge and future perspectives. *Animal*, 5 (1): 48-56.

Schoster, A., Arroyo, L. G., Staempfli, H. R. & Weese, J. S. (2013). Comparison of microbial populations in the small intestine, large intestine and feces of healthy horses using terminal restriction fragment length polymorphism. *BMC Res Notes*, 6: 91.

Schwiertz, A., Hold, G. L., Duncan, S. H., Gruhl, B., Collins, M. D., Lawson, P. A., Flint, H. J. & Blaut, M. (2002). Anaerostipes caccae gen. nov., sp. nov., a new saccharolytic, acetate-utilising, butyrate-producing bacterium from human faeces. *Syst Appl Microbiol*, 25 (1): 46-51.

Sloet van Oldruitenborgh-Oosterbaan, M. M. (1999). Laminitis in the horse: a review. *Vet Q*, 21 (4): 121-7.

Stanton, T. B. & Canale-Parola, E. (1980). Treponema bryantii sp. nov., a rumen spirochete that interacts with cellulolytic bacteria. *Archives of Microbiology*, 127 (2): 145-156.

Steelman, S. M., Chowdhary, B. P., Dowd, S., Suchodolski, J. & Janecka, J. E. (2012). Pyrosequencing of 16S rRNA genes in fecal samples reveals high diversity of hindgut microflora in horses and potential links to chronic laminitis. *BMC Vet Res*, 8: 231.

Stewart, C. S. & Flint, H. J. (1989). Bacteroides (Fibrobacter) Succinogenes, a Cellulolytic Anaerobic Bacterium from the Gastrointestinal-Tract. *Applied Microbiology and Biotechnology*, 30 (5): 433-439.

Suen, G., Weimer, P. J., Stevenson, D. M., Aylward, F. O., Boyum, J., Deneke, J., Drinkwater, C., Ivanova, N. N., Mikhailova, N., Chertkov, O., et al. (2011). The Complete Genome Sequence of Fibrobacter succinogenes S85 Reveals a Cellulolytic and Metabolic Specialist. *Plos One*, 6 (4).

Valones, M. A. A., Guimaraes, R. L., Brandao, L. A. C., de Souza, P. R. E., Carvalho, A. D. T. & Crovela, S. (2009). Principles and applications of polymerase chain reaction in medical diagnostic fields: A review. *Brazilian Journal of Microbiology*, 40 (1): 1-11.

van Dijk, E. L., Jaszczyszyn, Y. & Thermes, C. (2014). Library preparation methods for next-generation sequencing: Tone down the bias. *Experimental Cell Research*, 322 (1): 12-20.

Van Weyenberg, S., Sales, J. & Janssens, G. P. J. (2006). Passage rate of digesta through the equine gastrointestinal tract: A review. *Livestock Science*, 99 (1): 3-12.

Williams, A. G., Withers, S. E. & Coleman, G. S. (1984). Glycoside hydrolases of rumen bacteria and protozoa. *Current Microbiology*, 10 (5): 287-293.

Wylie, C. E., Collins, S. N., Verheyen, K. L. & Richard Newton, J. (2011). Frequency of equine laminitis: a systematic review with quality appraisal of published evidence. *Vet J*, 189 (3): 248-56.

Yu, Y., Lee, C., Kim, J. & Hwang, S. (2005). Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction. *Biotechnology and Bioengineering*, 89 (6): 670-679.

Zhang, T. & Fang, H. H. P. (2006). Applications of real-time polymerase chain reaction for quantification of microorganisms in environmental samples. *Applied Microbiology and Biotechnology*, 70 (3): 281-289.

Appendix

Appendix A: Primer sequences

PRK primers targeting prokaryotic 16S rRNA gene:

Forward (PRK341F): CCTACGGGRBGCASCAG

Reverse (PRK806): GGACTACYVGGGTATCTAAT

PRKillumina primers:

Forward (PRKi F):

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

Reverse (PRKi R):

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
31. caagcagaagacggcatacgagatATCGTGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT
32. caagcagaagacggcatacgagatTGAGTGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT
33. caagcagaagacggcatacgagatCGCCTGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT
34. caagcagaagacggcatacgagatGCCATGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT
35. caagcagaagacggcatacgagatAAAATGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT
36. caagcagaagacggcatacgagatTGTTGGgtgactggagttcagacgtgtgctcttccgatctGGACTACYVGGGTATCTAAT

Appendix B: QIIME manuscript for analyzing 16S rRNA gene metagenome data

1. Attach a volume to a folder in which all the data will be
   1.1 Go to https://console.aws.amazon.com/ec2/home?region=us-east-1#s=Volumes and attach one of the volumes (or create a new one).
   Horse data: vol-05106872 (/dev/xvdb1)
   1.2 in Putty, run: 'sudo fdisk -l' to identify volume ID
   1.3 then mount it to a directory:
   mkdir /home/ubuntu/data_horse
   sudo mount /dev/xvdb1 /home/ubuntu/data_horse
2. Copy the files into the folder; make sure to copy sequencing data, mapping files, make_split_libr_command_R1R2.py, convert_all_fastq_files.py, uc_fast_params.txt, mybashscript.sh
3. Unzip the files through
   tar –xvf filename.tar
   gzip –d *.fastq.gz
4. Convert all fastq files into fasta and qual files
   screen –S convert
   python convert_all_fastqfiles.py /home/ubuntu/data_horse/131029_M01132.Project_Rudi-Horse300-2013-09-27
   (to come back to screen type 'screen –r convert'; to terminate the screen type 'screen –S convert –X quit')
5. Check all the mapping files
   check_id_map.py –m ./checked_map_files/Mapping_file_R1.txt –o ./checked_map_files/qiime_otuput
6. Terminate the screen after it is finished
   screen –S convert –X quit
7. Split sequences in each sample file (don't forget to make changes in the code, specifying the mapping files pathway and names of files if different)
   screen –S split_libr

python make_split_libr_command_R1R2.py ./data/131029_M01132.Project_Rudi-Horse300-2013-09-27/fastaqual Rudi-MiSeq300-Knutrudi210813- ./data/131029_M01132.Project_Rudi-Horse300-2013-09-27/split_output

8. Concatenate all the sequences in one file

   (First move all the files from split_output* (forward direction only) into another folder)

   mkdir ./data/131029_M01132.Project_Rudi-Horse300-2013-09-27/split_all

   cat ./data/131029_M01132.Project_Rudi-Horse300-2013-09-27 /forward/split_output*/seqs.fna > ./data/131029_M01132.Project_Rudi-Horse300-2013-09-27/split_all

9. Install mpich2 if it is not installed yet

   sudo apt-get install libcr-dev mpich2 mpich2-doc

10. Edit the StarCluster config file

    vi ~/.starcluster/config

    use 'insert' or 'i' to edit the file, add the information on the volume (check volume id in console aws webpage, see p.1)

    ```
    [cluster qiime-horse]
    node_image_id = ami-64d0af0d
    cluster_user = ubuntu
    keyname=monikakey
    cluster_size=8
    node_instance_type=m2.4xlarge
    plugins=tmux, mpich2
    volumes=qiime-horseData

    [volume qiime-horseData]
    VOLUME_ID = vol-05106872
    MOUNT_PATH=/home/ubuntu/data_horse
    ```

    To save and quit: Esc; Shift+ZZ

    To quit without saving: Esc; :q!

11. Unmount and deattach the folder from the FileZilla

    sudo umount /home/ubuntu/data_horse

then detach it on the webpage https://console.aws.amazon.com/ec2/home?region=us-east-1#s=Volumes

12. Check the spotprice history of m2.4xlarge

    starcluster spothistory m2.4xlarge –d 60

13. Launch the cluster, bidding the price a bit over the average one

    starcluster start -c qiime-horse spotclusterQiime --bid 1.05

14. Login to starcluster

    starcluster sshmaster spotclusterQiime –u ubuntu

15. Check whether the folder /data_horse/temp exists, if not, make new (mkdir)
16. Edit qiime config file on the starcluster

    vi /home/ubuntu/qiime_software/qiime_config

    i (to insert text)

    a) cluster_jobs_fp start_parallel_jobs_sc.py
    b) temp_dir /home/ubuntu/data_horse/temp

    Esc; Shift+ZZ to quit and save

17. Check if changes have been made

    print_qiime_config.py

18. Check uc_fast_params.txt file

    vi uc_fast_params.txt

    i

    pick_otus:enable_rev_strand_match True

    pick_otus:max_accepts 1

    pick_otus:max_rejects 8

    pick_otus:stepwords 8

    pick_otus:word_length 8

    Esc; Shift + ZZ

19. Check mybashscript.sh file

    export

    reference_seqs=/home/ubuntu/qiime_software/greengenes/gg_13_5_otus/rep_set/99_otus.fasta;

    export

    reference_tree=/home/ubuntu/qiime_software/greengenes/gg_13_5_otus/trees/99_otus.tree;

    export

reference_tax=/home/ubuntu/qiime_software/greengenes/gg_13_5_otus/taxonomy/99_otu_taxonomy.txt; pick_closed_reference_otus.py -o /home/ubuntu/data_horse/131029_M01132.Project_Rudi-Horse300-2013-09-27/myotus99 -i /home/ubuntu/data_horse/131029_M01132.Project_Rudi-Horse300-2013-09-27/split_all.seqs.fna -r $reference_seqs -a --parallel -O 80 -p /home/ubuntu/data_horse/uc_fast_params.txt -f

20. Run parallel analysis of out picking

    start_parallel_jobs_sc.py -ms /home/ubuntu/data_horse/mybashscript.sh my_job_

21. Check whether the job has finished

    qstat

    If not output is given, then the job has finished.

22. Check if you had any errors by listing all files

    ls –all

    Check the size of .o and .e files (.e-files give information on errors). If their size is > 0, use cat file.e to read the file

23. Get an overview of sequence reads which were assigned to otu table

    print_biom_table_summary.py –i /home/ubuntu/data_horse/ 130924_M01132.Project_rudi-MiSeq300-2013-08-27/myotus99/out_table.biom

24. Choose sequencing depth to use in core diversity analysis (for example 4000 per sequence)

25. Convert the biom table to text format

    convert_biom.py -i /home/ubuntu/data_horse/ 131029_M01132.Project_Rudi-Horse300-2013-09-27/myotus99/out_table.biom -o /home/ubuntu/data_horse/ 131029_M01132.Project_Rudi-Horse300-2013-09-27/myotus99/out_table.txt -b --header_key taxonomy

26. Check the my_core_diversity_job.sh script

    vi my_core_diversity_job.sh

    i

    core_diversity_analyses.py -o /home/ubuntu/data_horse/131029_M01132.Project_Rudi-Horse300-2013-09-27/myotus99/e4000 -i /home/ubuntu/data_horse/131029_M01132.Project_Rudi-Horse300-2013-09-27/myotus99/out_table.biom -m /home/ubuntu/data_horse/checked_map_files/Mapping_file_all.txt -e 4000 -c “AgeCategory” -a --parallel -O 80 -t /home/ubuntu/qiime_software/greengenes/gg_13_5_otus/trees/99_otus.tree;

    Esc; Shift+ZZ to save changes and quit

27. Run the script

    start_parallel_jobs_sc.py –ms /home/ubuntu/data_horse/my_core_diversity_job.sh my_core_job_

28. Check whether the job has finished

    qstat

    If not output is given, then the job has finished.

29. Check if you had any errors by listing all files

    ls –all

    Check the size of .o and .e files (.e-files give information on errors). If their size is > 0, use cat file.e to read the file

30. If you want to delete all the jobs started by the user

    qdel -u ubuntu

31. Terminate starcluster (first log out from master and do it from the instance where starcluster was started)

    starcluster terminate -c spotclusterQiime

32. Reattach the volume to the folder

    Horse data: vol-05106872 (/dev/sdg)

    32.1 in Putty, run: 'sudo fdisk -l' to identify volume ID

    32.2 then mount it to a directory:

    sudo mount /dev/xvdb1 /home/ubuntu/data_horse

33. Transfer the data to PC through FileZilla

Appendix C: Full nutritional content of the feed

Nutritional content analysis results of the feed used in this feeding experiment is shown in Table A-1. The analysis were performed at IHA and dry matter, ash, crude protein, NDF, ADF, stringy, crude fat and starch were analyzed.

Table A-1: Full nutritional content of the hay and barley used in this feeding experiment.

	Dry matter	Ash	Crude protein	NDF	ADF	Stringy	Crude fat	Starch
Sample ID	**g/kg**	**g/kg**	**g/kg**	**g/kg**	**g/kg**	**g/kg**	**g/kg**	**g/kg**
Hay	889	49	83	553	307	275	21	0
Barley	870	21	84	142	36	26	12	499

Appendix D: Weighted UniFrac summary

Table A-2: Compare diversities within feces to those within cecum and those between these two groups.

	Feces	Cecum	Feces to Cecum
Mean	0.280445	0.198106	0.34915
STD	0.118203	0.066195	0.08672
p-values	**Feces to Cecum**	**Feces to Feces_Cecum**	**Cecum to Feces_Cecum**
Kruskal-wallis	0.00E+00	5.12E-122	0.00E+00
ANOVA	0.00E+00	1.48E-89	0.00E+00

Table A-3: Compare diversities between various horses.

	Feces				Cecum			
	Horse 1	**Horse 2**	**Horse 3**	**Horse 4**	**Horse 1**	**Horse 2**	**Horse 3**	**Horse 4**
Mean	0.236323	0.216539	0.307835	0.258328	0.155814	0.167317	0.124889	0.142379
STD	0.097815	0.100992	0.132071	0.126650	0.052592	0.046022	0.033857	0.038154

Kruskal-Wallis p-value	**Horse 1 to 2**	**Horse 1 to 3**	**Horse 1 to 4**	**Horse 2 to 3**	**Horse 2 to 4**	**Horse 3 to 4**
Feces	4.55E-06	4.39E-27	0.041325	4.85E-48	7.01E-12	2.18E-18
Cecum	3.28E-34	2.41E-122	2.19E-12	2.46E-279	3.35E-98	9.38E-75

	Horse 1	**Horse 2**	**Horse 3**	**Horse 4**
Feces to Cecum	5.34E-100	4.13E-23	0.00E+00	5.00E-191

ANOVA p-value	**Horse 1 to 2**	**Horse 1 to 3**	**Horse 1 to 4**	**Horse 2 to 3**	**Horse 2 to 4**	**Horse 3 to 4**
Feces	1.37E-04	6.43E-32	0.000117	3.72E-52	8.56E-14	1.62E-16
Cecum	6.35E-17	3.18E-157	7.05E-28	0.00E+00	5.91E-105	7.96E-82

	Horse 1	**Horse 2**	**Horse 3**	**Horse 4**
Feces to Cecum	2.76E-165	7.51E-78	0.00E+00	0.00E+00

Table A-4: Compare diversities between different diets.

	Feces		Cecum	
	Hay+Barley	**Hay**	**Hay+Barley**	**Hay**
Mean	0.232407	0.309847	0.175168	0.209879
STD	0.090033	0.124757	0.057821	0.077188
Kruskal-Wallis p-value			**Feces**	**Cecum**
Diet 1 to Diet 2			4.73E-24	4.42E-98
	Hay+Barley		**Hay**	
Feces to Cecum	2.78E-247		0.00E+00	
ANOVA p-value	**Feces**		**Cecum**	
Diet 1 to Diet 2	1.40E-26		1.94E-104	
	Hay+Barley		**Hay**	
Feces to Cecum	0.00E+00		0.00E+00	

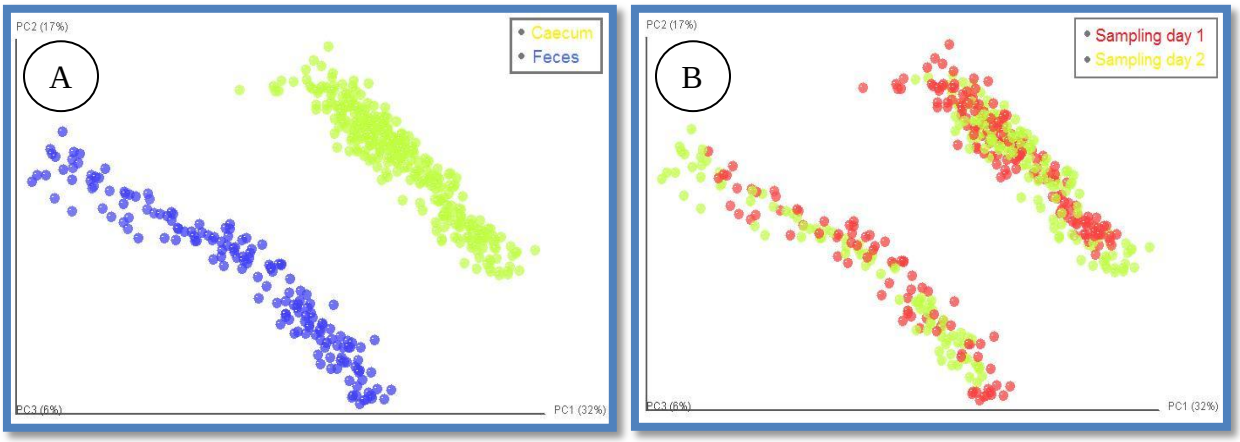


Figure A-1: Weighted UniFrac plot colored by sample origin (A) and sampling day (B).

Appendix E: PCA plots

PCA analysis of OTU abundance was used in order to assess an effect of diet, time and sample origin.

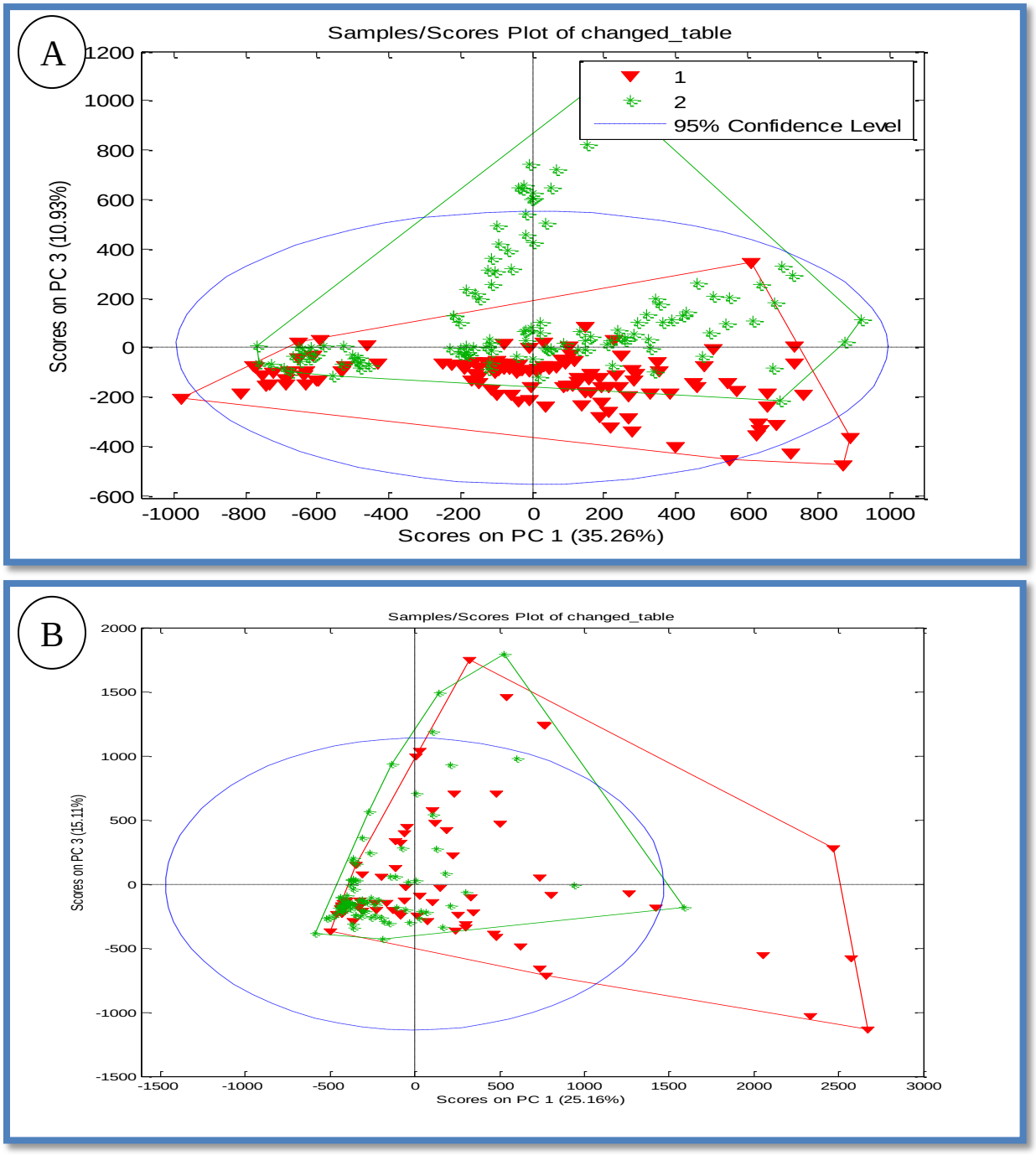


Figure A-2: PCA plot by diet (PC1 vs. PC3). Diet 1 (hay+barley) marked in red and Diet 2 (hay) marked in green. A) Cecum samples. B) Feces samples.

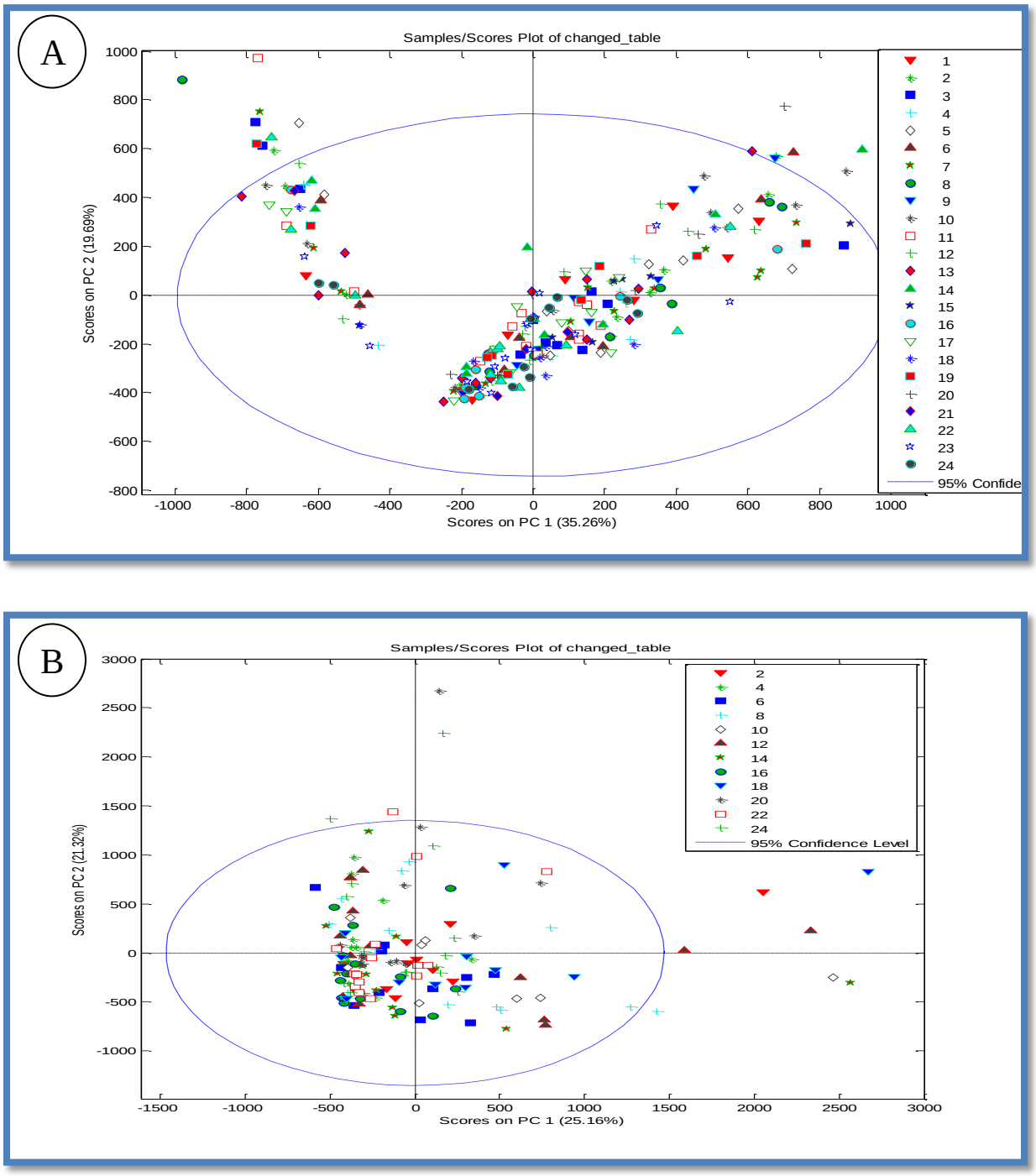


Figure A-3: PCA plot by time (PC1 vs. PC2). The numbers indicate different time points. A) Cecum samples B) Feces samples

Appendix F: OTU loadings

Through PCA analysis in MATLAB, loadings for both the cecum and feces samples were generated.

Table A-5: PC scores of different OTUs in cecum samples. The table only shows OTUs with loadings above 0.1 or below -0.1 in one of the three PC.

OTU_id	PC1	PC2	PC3	Taxonomy[1)]
102910	0.05	0.08	0.15	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_; s_
670167	0.27	0.28	0.31	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_; s_
297613	0.65	0.29	-0.44	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_;s_
49817	0.09	-0.03	-0.19	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_;s_
348828	0.10	0.10	0.39	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_;s_
289958	0.10	0.04	-0.03	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_;s_
4438136	-0.10	0.11	-0.03	k_Bacteria;p_Bacteroidetes c_Bacteroidia;o_Bacteroidales;f_;g_;s_
325743	-0.53	0.71	-0.21	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_;g_;s_
295015	-0.03	0.15	-0.02	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_;g_;s_
340727	0.09	0.23	0.11	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_;g_;s_
290027	-0.05	0.15	-0.03	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_RF16;g_;s_
325340	0.09	0.18	0.02	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_[Paraprevotellaceae];g_YRC22;s_
320615	0.15	0.05	0.02	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_[Paraprevotellaceae];g_CF231;s_
337167	0.06	0.11	0.53	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_;g_;s_
812596	0.11	-0.02	-0.21	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Veillonellaceae;g_Anaerovibrio;s_
541394	0.09	0.06	0.19	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_Anaerostipes;s_
288448	0.14	0.04	-0.02	k_Bacteria;p_Verrucomicrobia;c_Verruco-5;o_WCHB1-41;f_RFP12;g_;s_

1) k = kingdom, p = phylum, c = class, o = order, f = family, g = genus, s = species

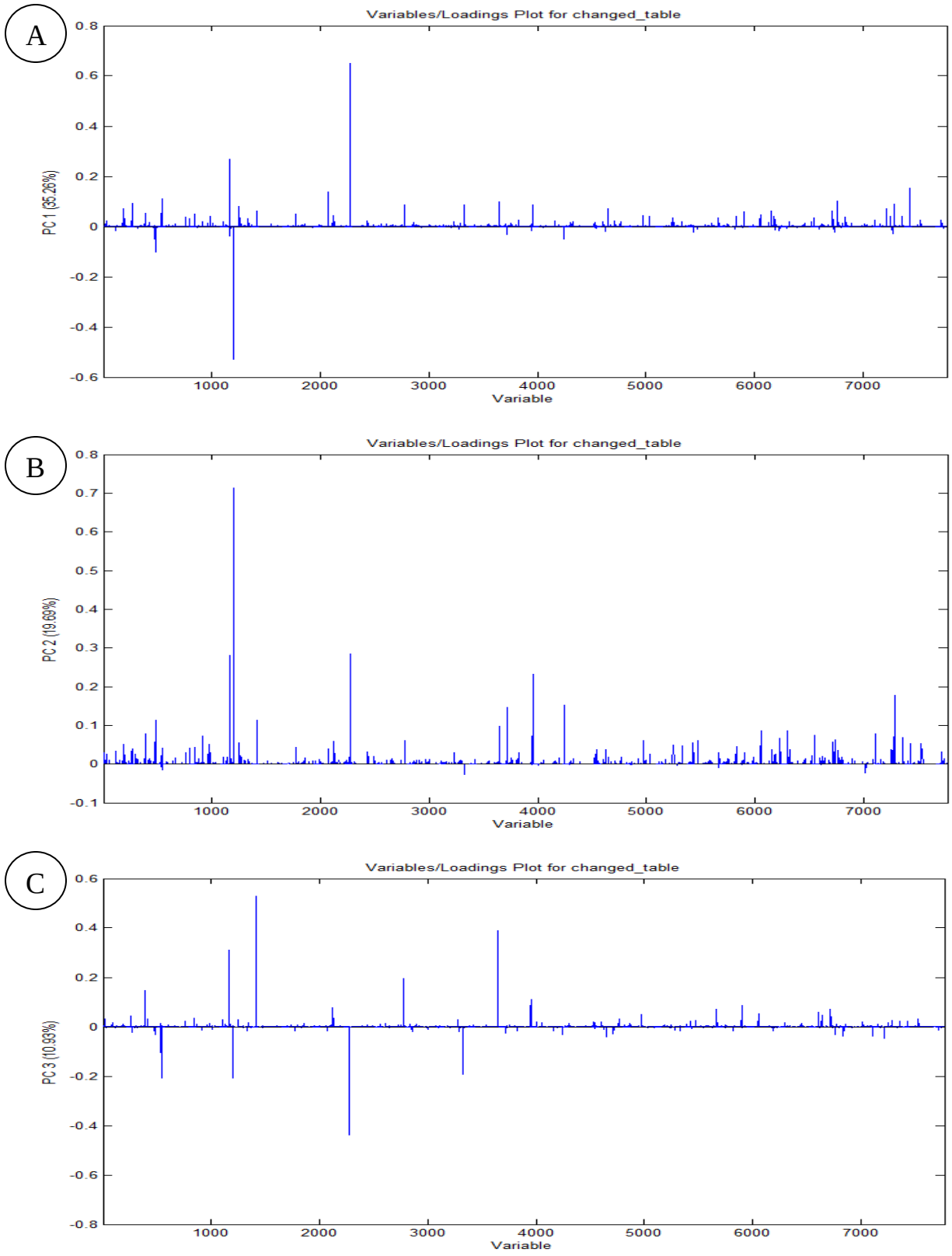


Figure A-4: Loadings plot according to PC1 (A), PC2 (B) and PC3 (C) for cecum samples.

Table A-6: PC scores of different OTUs in feces samples. The table only shows OTUs with loadings above 0.1 or below -0.1 in one of the three PC.

OTU_id	PC1	PC2	PC3	Taxonomy[1)]
126	0.06	-0.03	0.10	k_Archaea;p_Euryarchaeota;c_Methanobacteria;o_Methanobacteriales;f_Methanobacteriaceae;g_Methanobrevibacter;s_
114	0.11	-0.02	0.23	k_Archaea;p_Euryarchaeota;c_Methanobacteria;o_Methanobacteriales;f_Methanobacteriaceae;g_Methanobrevibacter;s_
301555	0.13	-0.08	0.34	k_Bacteria;p_Actinobacteria;c_Coriobacteriia;o_Coriobacteriales;f_Coriobacteriaceae;g_;s_
292150	-0.01	0.15	0.11	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_;g_;s_
290980	0.02	0.11	-0.08	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_;g_;s_
297555	-0.14	0.21	-0.09	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_;g_;s_
295015	-0.04	0.34	0.06	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_;g_;s_
340727	0.03	0.21	0.03	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_;g_; s_
330276	0.00	0.16	0.02	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_;g_;s_
330831	0.00	0.11	0.13	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_BS11;g_;s_
346659	0.00	0.30	0.28	k_Bacteria;p_Bacteroidetes;c_Bacteroidia;o_Bacteroidales;f_BS11;g_;s_
290276	0.08	0.13	-0.13	k_Bacteria;p_Fibrobacteres;c_Fibrobacteria;o_Fibrobacterales;f_Fibrobacteraceae;g_Fibrobacter;s_succinogenes
300658	0.18	0.00	-0.06	k_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g_Streptococcus;s_
299918	0.87	-0.02	-0.28	k_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g_Streptococcus;s_luteciae
300139	0.07	0.15	0.03	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_;g_;s_
313602	0.02	-0.01	0.12	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_;g_;s_
292128	0.13	-0.07	0.39	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_[Mogibacteriaceae];g_Mogibacterium;s_
291013	0.02	0.15	-0.03	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_;s_
353085	0.08	0.00	0.13	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_;s_
356061	0.04	0.05	0.28	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Lachnospiraceae;g_;s_
318278	0.07	-0.04	0.14	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_;s_
101501	0.08	-0.02	0.14	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_;s_
111019	0.03	0.12	0.00	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_;s_
314743	-0.01	0.11	0.00	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_Ruminococcus;s_
299609	0.03	0.20	0.00	k_Bacteria;p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_Ruminococcus;s_
293999	-0.05	0.31	-0.10	k_Bacteria;p_Firmicutes;c_Erysipelotrichi;o_Erysipelotrichales;f_Erysipelotrichaceae;g_RFN20;s_
292458	0.00	0.12	-0.03	k_Bacteria;p_Spirochaetes;c_Spirochaetes;o_Spirochaetales;f_Spirochaetaceae;g_Treponema;s_
293538	0.01	0.10	-0.08	k_Bacteria;p_Spirochaetes;c_Spirochaetes;o_Spirochaetales;f_Spirochaetaceae;g_Treponema;s_
297140	0.03	0.11	-0.04	k_Bacteria;p_Spirochaetes;c_Spirochaetes;o_Spirochaetales;f_Spirochaetaceae;g_Treponema;s_

1) k = kingdom, p = phylum, c = class, o = order, f = family, g = genus, s = species

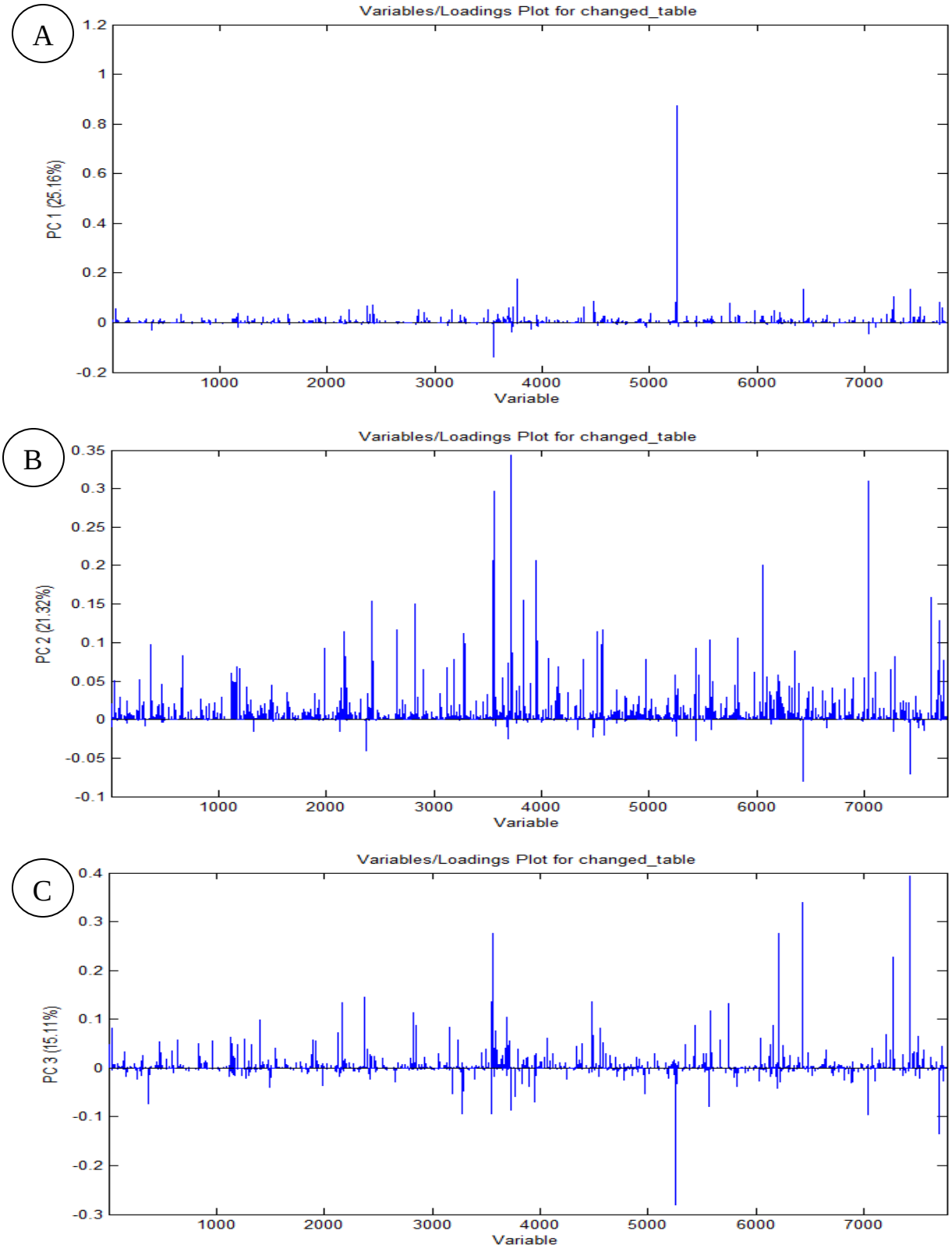


Figure A-5: Loadings plot according to PC1 (A), PC2 (B) and PC3 (C) for feces samples.

Appendix G: OLS regression analysis of *F. succinogenes* and *Treponema spp.*

OLS regression analysis of *F. succinogenes* and *Treponema spp*. were performed to address whether these bacteria showed correlating growth patterns.

Table A-7: OLS regression analysis summary.

Dependent Variable	VAR(1)
N	24
Multiple R	0.410
Squared Multiple R	0.168
Adjusted Squared Multiple R	0.131
Standard Error of Estimate	0.799

Regression Coefficients B = $(X'X)^{-1}X'Y$

Effect	Coefficient	Standard Error	Std. Coefficient	Tolerance	t	p-Value
CONSTANT	0.544	0.539	0.000	.	1.010	0.323
VAR(11)	0.596	0.283	0.410	1.000	2.110	0.046

Analysis of Variance

Source	SS	df	Mean Squares	F-Ratio	p-Value
Regression	2.843	1	2.843	4.454	0.046
Residual	14.046	22	0.638		

Durbin-Watson D-Statistic	0.526
First Order Autocorrelation	0.731

Information Criteria

AIC	61.253
AIC (Corrected)	62.453
Schwarz's BIC	64.787

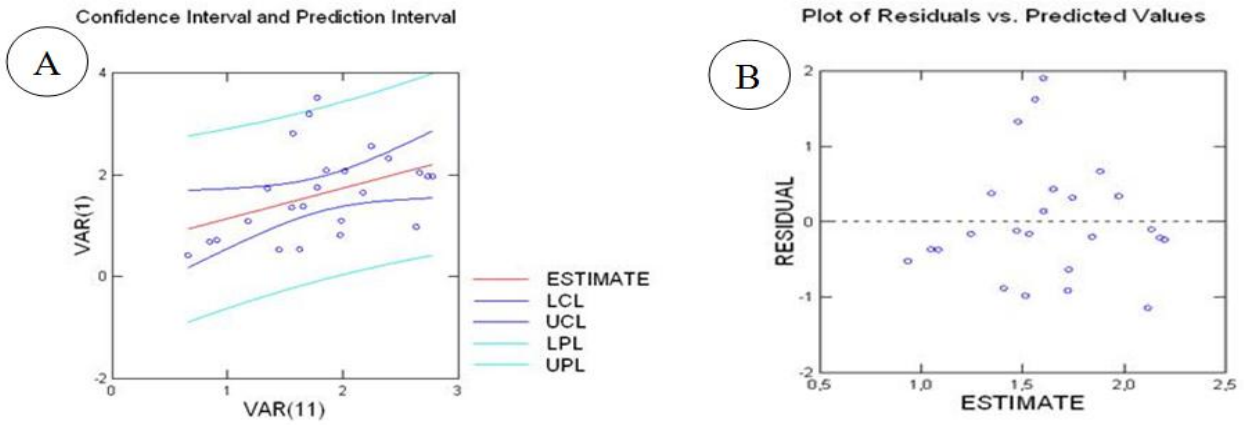


Figure A-6: OLS regression analysis. A) Confidence and prediction intervals. B) Residual plot.

Appendix H: Phyla distribution generated from MG-RAST

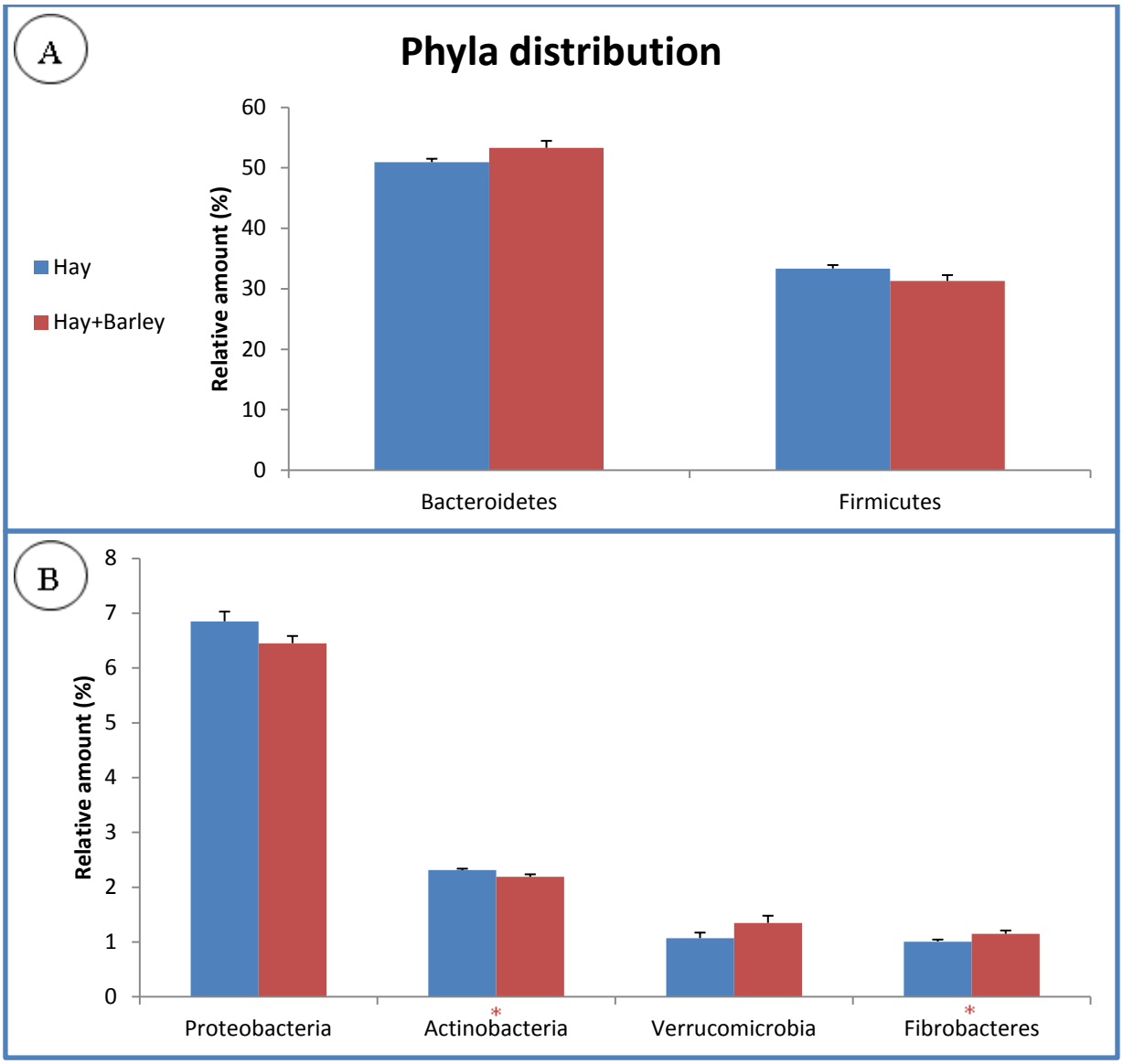


Figure A-7: Shotgun sequencing, Phyla distribution in cecum according to diet (mean ± SEM). A) Most dominant phyla: *Firmicutes* and *Bacteroidetes*. B) Remaining phyla over 1% in either diet. *Significant diet difference ($0.010<p<0.050$).

Appendix I: P-values of the carbohydrate metabolic features

Diet and time difference in carbohydrate metabolism related genes were addressed through t-test.

Table A-8: t-Test calculated p-values.

		Diet comparison at different time points			Time point comparison in hay diet			Time point comparison in hay+barely diet			Total diet comparison
Metabolic feature	**Time points**	**1**	**2**	**3**	**1--2**	**2--3**	**1--3**	**1--2**	**2--3**	**1--3**	
Aminosugars		0.746	0.594	0.940	0.083	0.410	0.043	0.004	0.181	0.023	0.762
CO2 fixation		0.756	0.017	0.766	0.309	0.181	0.793	0.047	0.092	0.729	0.091
Central carbohydrate metabolism		0.696	0.020	0.854	0.277	0.628	0.312	0.000	0.001	0.381	0.072
Fermentation		0.210	0.025	0.877	0.116	0.225	0.140	0.105	0.002	0.008	0.011
Glycoside hydrolases		0.281	0.741	0.036	0.280	0.100	0.073	0.000	0.000	0.288	0.853
Monosaccharides		0.295	0.006	0.153	0.032	0.746	0.043	0.000	0.000	0.721	0.176
One-carbon Metabolism		0.701	0.064	0.569	0.437	0.178	0.444	0.011	0.664	0.036	0.131
Sugar alcohols		0.467	0.004	0.807	0.453	0.659	0.309	0.001	0.001	0.621	0.063

-- Between time points.

Appendix J: Regression analysis plots of controls

Regression analysis were performed in Microsoft excel to generate R^2-values to describe the linear relationship between the positive controls.

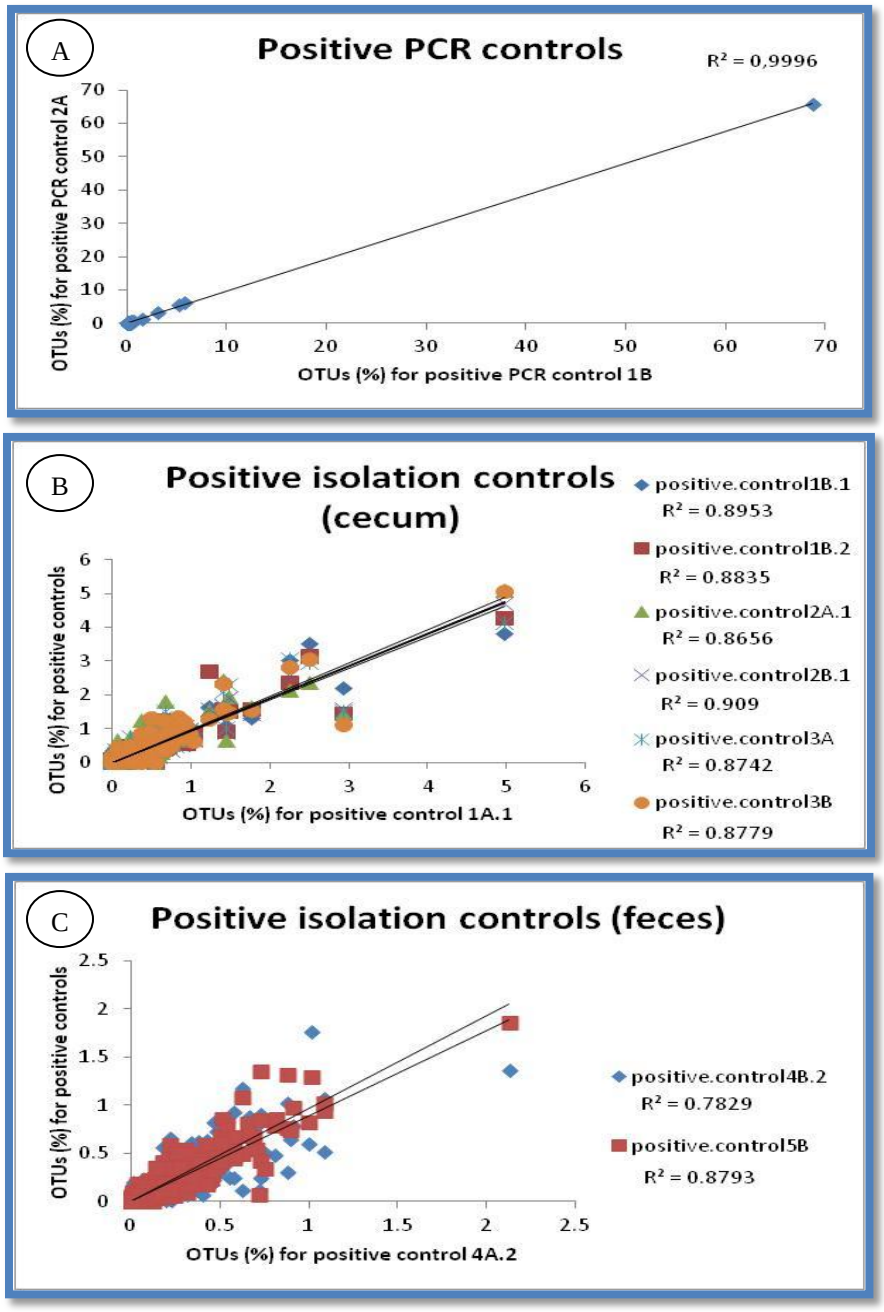
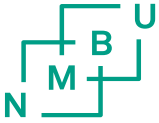


Figure A-8: Regression plots of positive controls. A) Positive PCR controls. **B)** Positive isolation controls from cecum samples**. C)** Positive Isolation controls from feces samples.



Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no