TURUN
YLIOPISTO
Matemaattis-luonnontieteellinen
tiedekunta

Stress-associated changes in the gut microbiome of Asian elephants

Minka Ovaska

Master's Degree Program in Physiology and Genetics
Master's thesis
Credits 30 op

Supervisors:
Virpi Lummaa, Prof.
Eero Vesterinen, Ass. prof.

3.2.2023
Turku

Master's thesis

Subject: Biology
Author: Minka Ovaska
Title: Stress associated changes the gut microbiome of Asian elephants
Supervisors: Prof. Virpi Lummaa and Ass. prof. Eero Vesterinen
Number of pages: 53 pages and 5 appendices
Date: 3.2.2023

Background: The intestinal tract of practically all animals is inhabited by a diverse micro-organism referred to as gut microbiome. The importance of gut microbiome to its host is well known and an increasing number of studies show that the microbes are associated with many body functions, such as the gut-brain axis function. Stress, in particular, can impact the gut-brain axis at all stages of life. Stress can reshape the gut bacterial composition through glucocorticoid secretion, inflammation, or autonomic alterations. This often leads to gut bacterial imbalance (dysbiosis), as well as low microbial diversity. Although studies on gut microbiome are increasing, the mechanisms of gut-brain axis and how microbes associate with stress still remain poorly understood. Also, the gut microbiome studies have focused mostly on humans or laboratory animals in controlled environments. It is important to widen the studies on mammals in their natural environment to better understand the complex development of gut microbiome. The aim of this thesis was to (1) determine the composition of gut microbiome of Asian elephants living in their natural environment and (2) test the hypothesis that increased stress alters the gut microbiome composition.
Methods: The gut microbiome was determined from 94 semi-captive Asian elephants *(Elephas maximus)* in Myanmar, from which both fecal sample collection and blood sampling was possible. The gut microbiome was analyzed from the fecal samples using 16S rRNA metabarcoding approach. Stress levels were determined from the blood samples as: (I) serum cortisol (SC), (II) heterophil to lymphocyte ratio in blood samples (H:L) and (III) fecal glucocorticoid metabolites (FGM).
Results: The overall gut microbe composition in Asian elephants was similar to previous studies, dominated by *Firmicutes* (55%) and *Bacteroidetes* (25%) followed by *Spirochaetaes* (8%). One specific genus *Solibacillus* was found to be significantly more abundant in this thesis compared to previous studies. There were also differences in the microbiome between elephants. The age and location of the elephant had significant effects on the gut microbe composition. The stress measure H:L ratio was also associated with gut microbiome by reducing the alpha diversity. All three stress measures were also associated with compositional changes in the gut microbiome.
Conclusions: Gut microbe composition in Asian elephants is diverse and similar to other fermenters. Stress, particularly long-term stress, can shape the composition of the microbiome. This thesis provides in my knowledge the most comprehensive picture of gut microbiome in Asian elephants in their natural living environment. In addition, this thesis adds important knowledge of the microbiota-gut-brain axis and how stress levels are associated with it.

Key words: gut microbiome, stress, Asian elephant, alpha diversity, beta diversity, community comparison, natural living environment

Table of Contents

1. Introduction

1.1 Gut microbes

Microbes are a diverse group of microscopic organisms consisting of bacteria, archaea, algae, viruses and fungi. Through the development of high-throughput sequencing technologies, the knowledge of the importance of microbes, especially gut microbes, is increasing. The use of the term "superorganism" has been increasing in scientific literature since 2000 to describe the symbiotic relationship between host and its microbiome. Microbes colonize different parts of the body including intestinal tract, genital area, skin, and mucus membranes both in the mouth and in the respiratory system, with different niches in each body cavity, forming different microbiomes. The term microbiome is used to refer to the complex microbial community, with all its genetic material, in a specific body cavity (Ursell et al., 2012; Wu et al., 2013). Microbes colonizing the intestinal tract form the largest microbiome in the body.

Recent studies on the microbiome have increased the knowledge about the importance of the gut microbes. Studies in both humans and other mammals have implicated the role of gut microbiome in a range of physiological processes vital to health and overall quality of life: energy homeostasis, metabolism, gut epithelial health, immunologic activity, and neurobehavioral development are all affected by gut microbes (Barko et al., 2018). In order to understand the impact that gut microbes have on health, it is necessary to decipher the content, diversity and functioning of the microbial gut community.

Currently there are two main DNA-sequencing based approaches to analyze gut microbiome: DNA amplicon sequencing and shotgun metagenomic sequencing approaches (Xia et al., 2018). In amplicon sequencing only a fragment of one gene is amplified and sequenced depending on what is studied. In bacterial community analysis the 16S rRNA gene is often targeted for its suitability for bacterial species identification. The 16S rRNA gene encodes the RNA component of the prokaryotic ribosome, with a slow rate of evolution. The gene contains both conserved regions and variable regions. There are together nine hypervariable regions (V1-V9) in the 16S rRNA gene, and they are used in the identification of different prokaryotic taxa. The conserved domains serve

as a universal primer binding site for the PCR amplification of gene fragments (Boers et al., 2019). The shotgun metagenomic sequencing seeks to understand the microbial community beyond the taxonomic identification, as well as adding resolution to the species determination. Metagenomics refers to sequencing of the DNA fragments from whole genomic DNA in a specific environment like the digestive system of humans and other mammals. The difference to amplicon sequencing is the ability to directly analyze the biological functions of specific bacterial taxa (Sharpton et al., 2014).

1.1.1 Composition of gut microbiome

It is challenging to create an overview of the whole gut microbiome. The microbial composition varies between host species and there are a vast majority of factors affecting the composition and diversity of the microbial gut community. The gut microbiome composition also varies between individuals, making it even harder to distinguish the normal composition. There are two main diversity measures that can be used to describe the microbiome: alpha and beta diversity. Alpha diversity is a measure for species diversity in one individual and beta diversity quantifies similarities or dis-similarities in the microbiome between individuals. The human gastrointestinal tract is inhabited by approximately 1014 microorganisms, with over 1000 species, largely from 14 “core” genera of bacteria. (Wiley et al., 2017) Despite the variation, it is possible to determine the core gut microbiome in different species and individuals.

The gastrointestinal tract of mammals consists of esophagus, stomach (and rumen in ruminants), small intestine and large intestine. The microbiome composition is reflective of the physiological properties in each region. Colon as a part of the large intestine provides a suitable environment for anaerobic bacteria to grow and reproduce, and for that reason the microbial community is the most diverse and abundant there. Dominating taxa in the colon are *Prevotellaceae*, *Lachnospiraceae* and *Rikenellaceae* (Thursby & Juge, 2017). The second most diverse microbial community exists in the small intestine. The environmental differences between these two intestinal tracts drive the separation between the microbial communities. For example, the small intestine is susceptible for digestive enzymes and oxygen, the peristaltic of the bowl and short transit time, making it a more difficult environment for the microbes to colonize. The small bowel is largely dominated by *Lactobacillaceae* (Thursby & Juge, 2017). Depending on the host species

the gastrointestinal tract has different adaptations like ruminant or cecal. The two adaptations are recognized in herbivorous animal species (Savage, 1977).

The core gut microbiome consists of two most abundant phyla *Firmicutes* and *Bacteroidetes. Firmicutes* make up for 50 – 80 % and *Bacteroidetes* 20 – 40 % of the gut microbiome in various species from humans to different herbivorous species (Million et al., 2013; Rojas et al., 2021). Gut microbiome in humans is classified into 12 different phyla, *Firmicutes* and *Bacteroidetes* being the most abundant, followed by *Proteobacteria*, *Actinobacteria*, *Verrucomicrobia*, *Fusobacteria* and *Cyanobacteria* (Molina-Torres et al., 2019; Thursby & Juge, 2017). In herbivores the most abundant bacterial families in the gut microbiome are *Ruminococcaceae, Rikenellaceae*, *Lachnospiraceae*, and *Prevotellaceae*. Members of these bacterial families participate in digestion of lignin, cellulose, hemicellulose, and protein found in vegetation (Rojas et al., 2021).

1.1.2 Function of gut microbiome

Gut microbes play important roles in many different functions in the mammalian body, having both systemic and local effects (Hill & Round, 2021). One of the main functions of the gut microbiome is the fermentation of indigestible food ingredients. It is a process which increases energy intake and nutrient use, amongst other things, by producing short-chain fatty acids (SCFAs). In herbivorous species members of bacterial families *Ruminococcaceae, Rikenellaceae*, *Lachnospiraceae*, and *Prevotellaceae* participate in digestion of lignin, cellulose, hemicellulose, and protein found in vegetation, turning these into a SCFAs (Rojas et al., 2021). Fermentation occurs in the colon where carbohydrates, fats and proteins are fermented.

In addition to fermentation gut microbes are involved in the development of many systems including the brain (Cryan et al., 2019) and immune system (Zheng et al., 2020). Specific microbes from the gut microbiome can produce antimicrobial agents like metabolites or express molecules on their surface that can affect the immune system pathways. In addition to that, the microbes can prevent the colonization of pathogenic organisms by competing from resources and producing toxins or detrimental metabolites.

Studies also suggest a connection between gut microbes and brain function and also mental health issues (Molina-Torres et al., 2019; Peirce & Alviña, 2019). The development of a diverse gut microbiome, especially in the early life, is believed to be vital for multiple features of behavior and physiology. If the composition or diversity is impaired, it can lead to neuropsychiatric illnesses such as anxiety and depression (Wiley et al., 2017). Stress is often a common feature in many neuropsychiatric disorders (Musazzi & Marrocco, 2016).

1.1.3 Development of gut microbiome

The early life microbial exposure plays an important role in determining the composition of gut microbiome in humans and in other mammals. The initial microbial exposure occurs 10 weeks after conception as the fetus begins to swallow the amniotic fluid. There is evidence that the microorganisms could start to colonize the gastrointestinal tract already after that (Golofast & Vales, 2020). The next microbes that the fetus encounters depend on the mode of the birth, which plays an important role in determining the composition of an individual's gut microbiome (Odamaki et al., 2016). In C-section deliveries the lack of exposure to vaginal microbes and microbes from mother's skin alters the type and diversity of the gut microbial community. The total amount of microbes in infants' gut is fewer and the composition is different from infants delivered vaginally. The C-sectionally delivered infants are more likely to be colonized by *Clostridium difficile* and less likely by *Bifidobacterium* and *Bacteroides* in the first months (Francino, 2018). It is also shown that the fecal microbiome of babies delivered vaginally resembles the fecal microbiome of mothers' in 72% of the cases. In babies delivered C-sectionally the percentage is reduced to only 41% (Thursby & Juge, 2017).

The establishment of a stable gut microbiome is usually followed by two remarkable transitions in infancy. The first transition occurs during lactation when the gut microbiome becomes dominated by *Bifidobacterium.* The undigested oligosaccharides from breast milk pass through infants' gastrointestinal tract and provide favorable conditions for the growth of *Bifidobacterium* genus (Marcobal et al., 2011). The second transition occurs by introduction to solid foods. By consuming solid food, the microbiome composition changes towards adult-like complex microbiome dominated by the phyla *Bacteroidetes* and *Firmicutes.* In humans the stable gut microbiome composition is

acquired during the first two years of life and after that the composition depends largely on lifestyle (Tanaka & Nakayama, 2017).

The gut microbiome develops gradually, meaning that the bacterial species that already colonize the gastrointestinal tract create the conditions for the subsequent species. After the stable gut microbiome has been formed in early life, the microbiome composition can be altered by lifestyle factors such as diet, environmental changes, antibiotics, and exercise, but also different genetic factors, aging and life events affect the composition (Odamaki et al., 2016; Willing et al., 2011; Wu et al., 2011). Stress is also known to affect the composition of gut microbes. Prolonged physical or mental stress can alter the gut microbiome by for example by reducing the diversity of the microbes.

Food is an important modifier of the adult gut microbiome in humans and other mammals. Low fiber and high fat diets are positively associated with phyla *Bacteroidetes* and *Actinobacteria* and negatively associated with *Firmicutes* and *Proteobacteria.* In human studies the Bacteroides enterotype was highly associated with animal protein and saturated fats, suggesting that meat consumption, as in Western diet, characterizes this enterotype. Prevotella enterotype in contrast is associated with carbohydrate-based diet, more typical of agrarian societies (De Filippo et al.,2010; Wu et al., 2011). Besides humans, the dietary effects of other mammals have been studied also (Moustafa et al., 2021; Rojas et al., 2021; Zhang et al., 2019). In their study Rojas et al. (2021) showed that in herbivores the gut microbiome is highly species-specific, and that the host taxonomy (30 %) accounted for more variation than the dietary guild (10 %) in the gut microbiome.

Antibiotics have been developed to treat bacterial infections, however through their use, lasting alterations are made for the composition of gut microbiome. Antibiotics use is one of the most dramatic factors affecting the gut microbiome in adults as well as in infants. The gut microbiome is a complex network of co-dependent microbes and by affecting one population, other populations that exchange for example secondary metabolites with the targeted population are indirectly affected. (Willing et al., 2011)

Genetic factors also play an important role in determining the gut microbiome. Host genetic variation is used to elucidate the variation with the microbiome in microbiome genome-wide association studies. Studies have shown how the genetic variation is linked

to the microbiome in humans. Goodrich et al. (2014) showed in their study how the microbiome of monozygotic twins is more similar compared to their other siblings. They also showed that siblings have a more similar microbiome than unrelated individuals. The most highly heritable taxon in the study were family *Christensenellaceae* and environmental factors shaped mostly the members of *Bacteroidetes* phylum because most of them were not heritable. Similar results have been found while studying the gut microbiome of monozygotic twins, marital partners, and unrelated individuals. Zoetendal et al. (2001) show in their study that the microbiome between monozygotic twins is significantly more similar than those of unrelated individuals. Similarities in the gut microbiome were not observed between marital partners in comparison to unrelated individuals. This study emphasizes the importance of genetic inheritance at the expense of environmental factors.

Lastly, aging affects the composition of microbes in the intestinal tract and the other way around. Aging is a genetically determined process that leads to decline in physiological functions and deterioration of various homeostatic functions. These alterations, related to gastrointestinal tract, like degeneration of the enteric nervous system and alteration of intestinal motility causes the gut microbiome to change. (Mangiola et al., 2018) In humans the relative abundance of *Lactobacillus* and *Bifidobacterium* decreased with age, while *Bacteroides* increased with age (Barko et al., 2018). After the stable gut microbiome composition is acquired in childhood the composition stays relatively stable through adulthood until it deteriorates at old age (Odamaki et al., 2016).

1.2 Stress and gut microbiome

Stress has negative implications for the host gut microbiome at all ages and stages of life, and it has become a significant health issue. The major response to stress is the activation of the hypothalamic-pituitary-adrenal (HPA) axis. HPA axis activation results in the release of various hormones, leading to a range of different biological effects, such as the modulation of the gut microbiome composition (Wiley et al., 2017). In general, stress can be defined as the state of disturbed homeostasis. It is an organism's response to different stressors like environmental pressure as well as mental or physical distress. Evolution has

developed ways to restore homeostasis, which requires a complex activation of responses from endocrine, nervous, and immune systems collectively.

Stress responses can be divided into acute and chronic responses. Acute stress responses consist of physiological and behavioral changes that have developed to aid survival in the wild. In an acute stress response, the normal life-history functions are suspended, allowing more energy to go to counteract the stressor. The primary response to stressors is the activation of the HPA axis and the activation of the sympathetic nervous system. (Misiak et al., 2020; Palma et al., 2014) Fight-or-flight response is a well-known example of the activation of the sympathetic nervous system in acute stress response. The blood flow to brain and muscles are heightened throughout the activation of the cardiovascular system, providing more energy for the animal to respond to danger. (Sapolsky et al., 2000)

In chronic stress response the same pathways are activated as in acute stress reaction. The difference is that chronic stressors are often persistent, or they have altered intensities, which causes the stress response to last longer. In a chronic stress response, the long-term release of glucocorticoid hormone and dysregulation of the HPA axis causes chronic disturbances in reproductive behavior, immune system functions and in other normal life-history functions. (Dickens & Romero, 2013) Overall, chronic stress is more difficult to define and measure than acute stress.

1.2.1 HPA axis

The HPA axis works through the positive and negative feedback system. Physical or physiological stressors activate the HPA axis resulting first into the release of corticotropin-releasing factor (CRF) from the hypothalamus. CRF binds into a receptor localized on pituitary corticotrophs and induces the release of adrenocorticotropic hormone (ACTH) into a systemic circulation. Circulating adrenocorticotropic hormone stimulates glucocorticoid synthesis in adrenal cortex, the principal target for the ACTH in systemic circulation. In response to stress a glucocorticoid (GC) hormone, such as cortisol and corticosterone are released causing a physical stress response. By measuring hormones like cortisol or corticosterone the levels of stress can be assessed. (Misiak et al., 2020)

1.2.2 Gut-brain axis

In all its simplicity the gut-brain axis means a conversation between intestinal tract and brain. The brain can communicate with the intestinal microbiome directly by releasing signal molecules to the lumen, or indirectly by altering intestinal permeability, activation of the HPA axis, autonomic nervous system, or the neuroendocrine system. Conversely the microbes can communicate with the brain via epithelial cells, microbial metabolites, and neurotransmitters. (Wiley et al., 2017) Several studies also suggest that stress can affect this bidirectional conversation between intestinal microbiome and the brain (Misiak et al., 2020; Molina-Torres et al., 2019).

There are different pathways for the cross-talk between intestinal microbiome and the central nervous system (CNS). The HPA axis is one of the most important pathways in the gut-brain axis (Misiak et al., 2020). Gut microbiome and HPA axis have bidirectional communication. Activated HPA axis increases the gut permeability allowing bacteria and bacterial antigens to cross the epithelial barrier causing the activation of mucosal immune responses, which in turn alters the composition of gut microbiome, and enhances HPA drive. On one hand, the altered gut microbiome may enhance the release of small bioactive molecules and cytokines from the intestine to the bloodstream, where they can migrate to the brain, pass through the blood brain barrier and serve as potent activator of the HPA axis (Misiak et al., 2020).

Besides the HPA axis, the communication between intestinal microbiome and brain occurs through enteric nervous system (ENS) and vagus nerve. Cranial vagus nerve is the fastest and most direct way to connect the gut and the central nervous system. Several bacteria can synthesize neurotransmitters such as acetylcholine (*Lactobacillus plantarum*), dopamine (*Bacillus*, *Proteus vulgaris*, *Serratia marcescens*), norepinephrine (*Bacillus*, *E. coli* and *Saccharomyces*), GABA (*Lactobacillus* and *Bifidobacterium*), histamine (*Citrobacter*, and *Enterobacter*), and serotonin (*Candida*, *E. coli*, *Enterococcus* and *Streptococcus*). (Strandwitz, 2018) Neurotransmitters are transported to the brain through the vagus nerve, where they can influence brain function. Then again, signals from the brain may influence sensory, motor and secretory modalities of the GI tract (Molina-Torres et al., 2019).

The indirect communication of the gut-brain axis occurs through microbial waste products like SCFA. They enhance the integrity of the intestinal epithelium as well as modulate the gut motility, stimulate vagus nerve, and also exert anti-inflammatory effects such as promotion of regulatory T cells (Farzi et al., 2018). The exact mechanisms of different communication pathways of the microbiota-gut-brain axis remains unknown, and more research is needed to understand the significance of the symbiotic relationship between host and microbe in the context of brain and intestinal health.

1.2.3 Stress and the gut microbiome composition

Stress can be detected as the activation of the HPA axis or in other changes in the functioning of the gut-brain axis. Overall, it can lead to changes in the gut microbiome composition. For instance, stress activates the HPA axis and in this way induces CRF release, which directly affects the bowel causing bowel dysfunction. The HPA axis also influences the microbiome through the glucocorticoid hormone release. Elevated cortisol hormone levels in mammals are associated with increase in the relative abundances of Proteobacterial groups and lower relative abundances of the lactic acid bacteria. (Misiak et al., 2020) Different *Lactobacillus* and *Bifidobacterium* species (*Lactobacillus helveticus R0052* and *Bifidobacterium longum R0175*) have also been found to reduce cortisol hormone levels, demonstrating the cross-talk between gut and brain (Messaoudi et al., 2011). Studies have revealed how chronic stress reduces the diversity of the gut microbiome and causes similar alterations to the composition of gut microbes as do acute stress responses (Wiley et al., 2017).

The association between stress and gut microbiome has been studied for over forty years. Tannock and Savage (1974) revealed how stressed mice showed dramatic reductions in the relative abundance of *Lactobacilli* when the environment of the mice was changed, and the food and water removed. The negative implications of stress on many health issues through the gut microbiome are indisputable and therefore the study of microbiome, as a potential therapeutic target for modulating stress response is becoming more important (Wiley et al., 2017). However, despite the fact that the association between stress and gut microbiome has already been studied over forty years, there are still a number of challenges that must be addressed before microbe manipulation can be used as a treatment for stress-related disorders.

Most importantly, many findings on how stress affects gut microbiome have come from studies with germ free mice. These results are not a realistic representation of the narrowing of gut microbiome diversity, which is the most frequently reported consequence of stress in clinical populations. Also, a controlled laboratory environment is not the most favorable environment of studying the complex and changing gut microbiome. In addition to laboratory mice, the focus has been on human gut microbiome and in the literature most of the gut microbiome studies revolve around humans. This may lead to severe bias, because microbiome and stress need to be studied also in natural conditions without access to advanced medical care and in a range of taxa in order to fully understand the complex mechanisms of gut-brain axis in other species as well.

To fully utilize the gut microbiome to improve health and use it as a treatment for stress-related disorders, it is important to know which microbes are present and what are their functions. Currently, despite the advances in sequencing technologies in the 2000s, the ideal "healthy" microbiome remains to be established and many microbes are still yet to be characterized. Humans are good targets for studying gut microbiome in long-lived species because age is a significant factor in the development of the gut microbe composition. One of the longest living terrestrial mammals alongside humans, are Asian elephants (*Elephas maximus*). Asian elephants could serve as a good model species to increase the diverse knowledge of gut microbes. Asian elephants differ from humans for example being herbivorous species and hindgut fermenters, but they also share similarities, like social lifestyle, reproductive pace and longevity.

1.3 Objectives and hypothesis

The scientific objective of this study is to assess the interaction between stress and gut microbes in Asian elephants. To complete this objective, my first sub-aim is to characterize the microbiome composition of the Asian elephant, second sub-aim is to measure the stress levels of the study elephants and determine whether there are elephants suffering from short-term or long-term stress, and finally I will compare the gut microbiome composition of individuals exposed to varying levels of stress to see the potential associations.

Previous studies on Asian elephants have provided information about their core microbiome as well as the effects of anthropogenic interferences on the Asian elephant gut microbiome (Moustafa et al. 2021), but the sample sizes have been small ranging between 3 and 30 individuals (Ilmberger et al., 2014; Zhang et al., 2019). This study will be one of the first to provide a comprehensive picture of the composition of the Asian elephant gut microbiome and how stress associates with it. The gut microbiome will be determined from 94 semi-captive Asian elephants living and foraging in their natural habitat in Myanmar, together with three different physiological stress measures used to assess stress levels; cortisol hormone measures from blood serum (SC), faecal glucocorticoid metabolites (FGM) and leucocyte profiles, more specifically the ratio between heterophils and lymphocytes (H:L).

I expect the gut microbe composition of Asian elephants to be similar to other fermenters, as far as the core species are concerned. I expect phyla *Bacteroidetes* and *Firmicutes* to be the dominant, but I also expect to detect phyla like *Proteobacteria* and *Clostridia* (Ilmberger et al., 2014). I also expect there to be variation in the composition of gut microbes inside the study population and that the variation is coming from the age, sex and origin (captive born, wild caught) of the elephant, as well as from the location of the working camp. For the second sub-aim I expect to find natural variation in the stress levels of Asian elephants. I also expect that the three different stress measures give differing pictures of the stress levels in elephants, because the measures are chosen to represent the overall stress as comprehensively as possible, with fecal glucocorticoid metabolites and leucocyte profiles representing long term stress and serum cortisol short term stress. As measures of stress, fecal glucocorticoid metabolites and leucocyte profiles also differ from each other considerably.

For the main objective about the connection between stress and microbiome I expect to find a decrease in the microbial diversity in the gut in stressed elephants. I will also expect that the relative abundance of various bacteria groups will alter in elephants suffering from stress. Reduction in genus *Lactobacillus* and increase in *Proteobacteria* have been previously reported as a stress associated change in the gut microbiome composition (Wiley et al., 2017).

My hypothesis is that stress has an effect on the gut microbiome through various mechanisms. I will focus my research on this line of discussion of the gut-brain axis, and

not assess how the microbes might affect the brain function. The knowledge of the different mechanisms is valuable, but in my study, I will be looking at the correlation between gut microbiome and different physical parameters known to associate with stress. Studying the causality of stress and the gut microbes would require a more interventional approach beyond the scope of a study such as mine focused on an endangered species roaming in its natural habitat and exposed to a wide range of stress causes.

2. Material and methods

2.1 Study population

Asian elephants are used for various human purposes throughout their range countries. Significant proportion of the remaining population of Asian elephants are working as draft and transport animals in the timber industry in Myanmar. Myanmar has the biggest population of semi-captive Asian elephants worldwide with a population of 5000 individuals, which is more than the rapidly shrinking wild population (Sukumar, 2006). More than half of this semi-captive population is state-owned through Myanma Timber Enterprise (MTE) (Seltmann et al., 2020).

This study focuses on MTE elephants. These elephants are a good model animal for studying the association between stress and gut microbiome. MTE elephants live in semi-captive conditions meaning that the elephants are working during daytime in the timber industry and at night they are released into the forest to forage naturally and interact with each other as well as occasionally wild conspecifics. This amount of freedom leads to the effects of natural environmental variation such as food availability, disease prevalence, complex climate patterns and host behavior affecting the physiology of the elephant and development of their gut microbiome. These elephants receive basic veterinary care to assess their working condition regularly, and the MTE keeps detailed records of all their elephants and their health information. (Crawley et al., 2020)

The diet plays an important role in determining the composition of gut microbes. This is especially the case with Asian elephants, which are considered one of the few extant megaherbivores and which can spend 12-18 hours a day feeding. Asian elephants are hind-gut fermenters and the plant cellulose is digested by the gut microbes in the large caecum and colon. (Sukumar, 2006) In Myanmar there are three seasons: 1) cool season from late October to March, 2) hot season from March until May, 3) wet/monsoon season from June to late October and the food availability is highly dependent on the season. In hot season over 70 % of the diet is browse while in the wet season grasses comprise the majority of the diet. In tropical forests, such as rainforests, the diet is mainly fruits and browse. (Sukumar, 2006)

The MTE elephants live in logging camps across the country. Camps are divided by age, health, or work status of the elephants. For example, the working elephants are located in different camps than the tamed, pregnant or lactating elephants. Age is connected to the elephant's physiology, composition of gut microbiome, and in MTE elephants to the particular stages of life. Elephants from age 20 until 50 are considered working adults. Over 50-year-old elephants are retired, and elephants under 20 are considered juveniles. Calves are tamed at the age of 4-5 years. In this thesis the age is categorized in these four groups.

In this study population one third of the elephants were born wild, captured, and tamed for working purposes. Two thirds of the elephants were born in captivity. Capture of wild animals may have long-term consequences on life-history for many reasons, one of them being altered behavior, physiology or immunity through chronic stress and sustained injuries. It has also been demonstrated that the capturing of the wild elephants increases the mortality compared to the captive-born elephants (Lähdenperä et al., 2018). The wellbeing and growth of the Asian elephant population worldwide is an important matter, as the population of Asian elephants has halved since 1950 and it is classified as an endangered species. The population growth is most limited by low birth rates and high juvenile mortality, with over 25% of calves being reported to die before the age of five. (Crawley et al., 2020)

2.2 Material

The gut microbe composition was determined from fecal samples collected from 94 elephants (males: 25 females: 69) during field sampling in May 2020, between the hot and the wet season. The samples were collected fresh after defecation from elephants in two different working camps in Sagaing region: Kawlin (n= 53) and West Katha (n=38). The age of the elephants varied between 1- to 67-year-old. Also, the life-history stage of the elephants varied between reproductive and non-reproductive as well as the background of the elephants, 68 born in captivity and 26 born in the wild. This kind of variety in the samples enables one to obtain a comprehensive picture of the gut microbiome as well as a wide range of stress experiences.

All the fecal samples were preserved in Myanmar, using two different preservation methods: ethanol (70 %) and by drying the feces in the low temperature oven 50 °C. The ethanol preserved samples were used for the gut microbiome analysis and the dried fecal samples for the stress analysis. Roughly at the same time as the fecal sample collection, blood samples for the stress analysis were taken by veterinarians in charge of the regular health care of the animals, as part of their health monitoring routine. Blood samples were collected early in the morning from an ear vein into vacuettes containing either ethylenediaminetetraacetic acid (EDTA), for differential white blood cell counts for H:L measure, or serum separator tubes for SC measure. The blood samples were collected from 68 elephants (males: 18 females: 50) corresponding the collected fecal samples. Sample size varied due to difficulties in obtaining blood samples for young untamed individuals.

2.3 Gut microbiome

2.3.1 DNA extraction

The ethanol-preserved fecal samples were stored at room temperature in Myanmar until transporting them to Finland. The samples were then stored in – 80°C, prior to DNA extractions. For the DNA extraction Qiagen´s PowerSoil Pro Kit (Ref nr: 47014) was used, following the manufacturer's protocol (version 05/2019). 120 mg of ethanol-preserved feces were dried and loaded to Powerbead tubes, deviating from manufactures recommendation of 250 mg. The optimization of the 120 mg sample weight was conducted prior to the extractions using two test samples, 100 mg, and 250 mg. The quality of the extracted DNA from the test samples were compared using Nanodrop (ThermoFisher Scientific). 800 µl of solution CD1 was added to the Powerbead tubes and vortexed briefly to start the cell lysis. Sample homogenization was conducted using TissueLyser II (QIAGEN). Powerbead tubes were placed on the TissueLyser Adapter Set, shaken 5 minutes at the speed of 25 Hz, re-orientated and shaken again. Powerbead tubes were centrifuged at 15,000 x g for 1 minute and the supernatants were transferred to a 2 ml Microcentrifuge tube. 200 µl of solution CD2 was added to the Microcentrifuge tube and vortexed for 5 seconds. The Microcentrifuge tube was then centrifuged at 15,000 x g for 1 minute and the supernatant transferred to a new Microcentrifuge tube, avoiding the

pellet. In addition to supernatant, 600 µl of solution CD3 was added into the Microcentrifuge tube and vortexed briefly. 650 µl of the lysate was loaded onto MB Spin Column and centrifuged at 15,000 x g for 1 minute. The flow-through the MB Spin Column was discarded and the rest of the lysate was loaded onto the same MB Spin Column and centrifuged again.

The DNA was bound to the silica filter membrane of MB Spin Column and the next phases included DNA purification. MB Spin Column was placed into 2 ml Collection tube and 500 µl of wash buffer EA was added and centrifuged at 15,000 x g for 1 minute. The flow-through was discarded and MB Spin Column placed back into the Collection tube. 500 µl of ethanol-based wash solution C5 was added to the MB Spin Column and centrifuged again at 15,000 x g for 1 minute. The MB Spin Column was placed into a new 2 ml Collection tube and centrifuged at 16,000 x g for 2 minutes for the ethanol to come off the membrane. MB Spin Column was placed into a 1,5 ml Elution Tube and 100 µl of solution C6 was added to the center of the filter membrane in MB Spin Column and centrifuged at 15,000 x g for 1 minute. The MB Spin Column was discarded, and the DNA was diluted in 100 µl of elution buffer (C6), in the Elution Tube. Prior to the PCR reactions the DNA was stored in -20°C.

2.3.2 PCR

To obtain gut microbiome composition DNA amplicon sequencing method was used in this study. Polymerase chain reaction (PCR) was conducted to the extracted DNA. PCR was targeted on prokaryotic 16S rRNA gene regions V3 to V4, containing species-specific signature sequences, useful for identification of different bacteria. The length of the amplified region was 464 bp. Targeting of PCR was conducted using universal primers Bakt-341F (CCTACGGGNGGCWGCAG) and Bakt-805R (GACTACHVGGGTATCTAATCC) (Herlemann et al., 2011). Like the DNA extractions, also the PCR protocol were optimized, and the reagents tested before actual analyses.

Altogether three PCR replicates were conducted on all the DNA samples. Negative controls, and a positive control sample ”mock” (ZymoBIOMICS Microbial Community

DNA Standard) was also used in the PCR. The reaction volume used in the PCR was 10 µl and the reaction components and volumes are presented in Table 1.

Table 1. PCR reagents and volumes used to amplify bacterial 16S rRNA gene regions V3 to V4.

Reagents	**Concentration in reaction**	**Volume (µl)**	**96 x volume (µl)**
Sterile H_2O	-	2.6 µl	250 µl
Bakt-341 F	0.2 µM	0.2 µl	19.2 µl
Bakt-805R R	0.2 µM	0.2 µl	19.2 µl
DNA polymerase KAPA Hifi	1x	5 µl	480 µl
DNA	-	2 µl	-

Master mix was prepared by adding sterile H_2O, DNA polymerase KAPA Hifi (Roche) and both forward and reverse primers into a sterile Eppendorf tube (Table 1). After that 8 µl of the master mix and 2 µl of the extracted DNA was pipetted into the 96 well plate and the reagents were spun down. The PCR reaction was conducted on a Bio-rad Thermocycler. The amplification program started by initial denaturation (95 °C, 3 min) and followed by 25 cycles of denaturation (95 °C, 30 s), annealing (55 °C, 30 s) and extension (72 °C, 30 s). The program finished on a 10-minute final extension on a 72 °C temperature. The PCR products were kept at 4 °C after the PCR.

The agarose gel electrophoresis was used to evaluate the PCR products. The reaction was conducted on 1,5 % agarose gel, prepared by mixing 100 ml of 0,5 x TBE Buffer, 1,5 g of agarose and 5 µl of Midori green advanced DNA Stain marker (Nippon Genetics). GeneRuler 100 bp (ThermoFisher Scientific) molecular weight marker was used in the reaction. The PCR products and the molecular weight marker were loaded into the agarose gel and run 30 minutes with a voltage of 120 V. DNA was visualized using ChemiDoc (BioRad).

2.3.3 Index PCR

Following the protocol for Illumina 16S sequencing library preparation, two-step PCR reaction was performed to the extracted DNA. The indexing PCR reaction was conducted in order to tag each DNA sample with specific indexes. The DNA from the first PCR reaction and other reagents (Table 2) were used in indexing PCR. Different sets of primers, with individual indexes for each of the samples were used in the second PCR reaction.

Table 2. Reagents and volumes used in the index PCR.

	plate 1.1	**plate 1.2**	**plate 2.1**	**plate 2.2**	**plate 3.1**	**plate 3.2**
Reagents	Conc.	Volume	Conc.	Volume	Conc.	Volume
DNA polymerase	1x	5 µl	1x	5 µl	1x	5 µl
i7_14R GAGATCTG	0.5 µM	1 µl	-	-	-	-
i7_58R TGACAGAG	0.5 µM	1 µl	-	-	-	-
i7_7R AGATCTG	-	-	0.5 µM	1 µl	-	-
i7_11R GGCTACAG	-	-	0.5 µM	1 µl	-	-
i7_5R ACCACTGT	-	-	-	-	0.5 µM	1 µl
i7_98R TGTGCACT	-	-	-	-	0.5 µM	1 µl
DNA 1:st PCR	-	3 µl	-	3 µl	-	3 µl

Altogether three index PCR reactions were conducted, one on each replicate plate. Six master mixes were prepared, one for each i7 reverse primer, by adding 250 µl of DNA polymerase and 50 µl of i7 reverse primer. PCR primers i7_14R and i7_58R were used in plate 1 and both master mixes were pipetted for one 96 well plate. The volume of master mix per well was 6 µl. Half of the plate contained reverse primer i7_14R and another half i7_58R. After that, 1 µl of i5 forward primer were added into the wells. Each well was added with a different i5 forward primer (supplement 1). DNA from the first PCR reaction was added into the wells (V= 3 µl), making the total volume of the reaction 10 µl. The PCR program used in the index PCR was shorter and had fewer rounds than the first PCR program. Index PCR program started with initial denaturation (95 °C, 4min) and followed by 10 cycles of denaturation (98 °C, 20 s), annealing (60 °C, 15 s) and extension (72 °C, 30 s). The program finished on a 3-minute final extension at a temperature of 72 °C.

In the other two index PCR reactions, different reverse primers were used to balance the nucleotides of the primers for the Illumina sequencing. The reverse primers used in plate 2 and plate 3 are presented in Table 2. The i5 forward primers were the same as in plate 1 (Supplement 1).

2.3.4 DNA purification and pooling

Before sequencing, the DNA samples had to be purified. Purification of the DNA was conducted using SPRI-bead purification method following the protocol of Vesterinen et al. (2016). The SPRI-bead solution was prepared as described in the Supplement 2. The purification method was based on the concentration of the SPRI-bead solution and was conducted in two parts. The length of the amplified DNA fragments was 427 bp and in the first SPRI-bead purification, fragments longer than 800 pb were discarded. In the first part of the purification, the DNA from one PCR plate at a time was combined and 100 µl of the DNA was placed into an Eppendorf tube. After that 100 µl of SPRI-bead solution and 100 µl on ddH_2O was added into the tube. After five-minute incubation the tube was placed on a magnet and 300 µl of the supernatant was transferred into a second Eppendorf tube. The long DNA fragments had combined with the magnetic beads and constructed a pellet purifying the sample from unwanted long DNA fragments.

In the second part of the DNA purification 300 µl of the supernatant from the first purification and 60 µl of SPRI-bead solution was added into the second Eppendorf tube. In the second purification all the DNA fragments longer than 400 bp were combined with the SPRI-beads. The supernatant was discarded after it was cleared from the SPRI-bead solution, and fresh 80 % ethanol was added into the Eppendorf tube. The one-minute ethanol wash was repeated keeping the tube still on a magnet. After the ethanol wash, the DNA was diluted into a 150 µl of ddH_2O, vortexed and incubated before placing it back on a magnet again. The clear solution, containing the purified DNA, was then pipetted into a new Eppendorf tube and the purity of the DNA was ensured using Qubit 2.0 Fluorometer (ThermoFisher Scientific). The same SPRI-bead purification protocol was conducted on all the replicate plates.

Before the samples were sent to Finnish Functional Genomics Centre (FFGC) for sequencing, the quality of the purified DNA was also analyzed on 2100 Expert Bioanalyzer (Agilent Technologies) and E-gel. The DNA pools were analyzed separately, and the concentrations are presented in Table 3, the Bioanalyzer results are presented in the Supplement 3. The purified DNA from three replicate plates were pooled into one DNA library based on the Qubit concentrations presented in Table 3.

Table 3. DNA concentrations measured in Qubit 2.0 Fluorometer. DNA volumes in the final sequencing library based on Qubit concentrations.

	Concentration Qubit	**DNA Volume**
Pool 1	4.96 ng/ µl	38.8 µl
Pool 2	45.3 ng/ µl	2.3 µl
Pool 3	27.6 ng/ µl	6.8 µl

2.3.5 Bioinformatics

To obtain the gut microbe composition the sequencing data needed to be analyzed using different bioinformatics methods. First the Illumina MiSeq (2 X 300 bp) sequenced fastq files was uploaded to CSC (IT Center for Science) for data trimming. The data handling was carried out in a batch job in CSC (Supplement 4). In short, the script contained (i) modified VSERCH (Rognes et al., 2016) algorithm to filter the forward and reverse reeds and merge them together for one fasta file. (ii) cutadapt (Martin, 2011) to create a cleaned fasta file by removing adapter sequences, primers and poly A-tails. (iii) VSERCH to dereplicate sequences (iiii) usearch (Edgar, 2010) and VSERCH algorithms to make zero-radius OTU (Zotu) and taxonomy assignments for the sequences. Taxonomy assignations was made by comparing sequences to known bacterial sequences in the SILVA database (Quast et al., 2013).

The total amount of reads Illumina MiSeq sequencing produced were 18 274 368 and after merging and quality-filtering 13 166 863 reads remained for the further analysis. The sequences were pooled and collapsed into 1 165 438 unique haplotypes where the singleton reeds were removed, and clustered into 8 809 Zotus. The 8 809 Zotus formed a Zotu table which together with the taxonomy table were moved into R software for further analysis.

2.4 Stress analyses

Glucocorticoid hormones, secreted by the adrenal cortex in response to stressors can be measured directly in blood, as an immediate response to stress, or they can be measured through glucocorticoid hormone metabolites excreted in feces (Crawley et al., 2021). Stress can also be measured from blood samples using alternative stress measures to hormones, such as characterizing and calculating the white blood cell counts. Heterophils (or neutrophils, depending on the species) and lymphocytes are both involved in immunological processes and studies suggest that corticosteroids can drive the changes in heterophils and lymphocytes numbers by increasing the number of heterophils and decreasing the number of lymphocytes (Davis et al., 2008). To obtain a comprehensive measure of the stress variation in my study subjects, I used these three methods combined. The sample sizes varied between different stress measures because of the difficulties to collect the blood samples.

2.4.1 Serum cortisol

The cortisol hormone levels were analyzed from the serum to measure the circulating stress hormone levels in the blood at the time of sampling. The serum from the elephant blood samples were obtained by centrifuging serum separator tubes for 20 minutes at 3400 rpm. 64 serum samples (males: 17 females: 47) were analyzed for cortisol at the University of Turku, using a species independent Enzyme Immunoassay (Arbor Assays) following the manufacturers protocol. Monoclonal mouse antibody was used following the manufacturer's protocol.

Serum samples were prepared by mixing 10 µl dissociation agent and 10 µl blood serum. Samples were diluted 1:32 in assay buffer and 50 µl of the sample were pipetted into a well of the plate (coated with goat anti-mouse IgG) along with 25µl DetectX® Cortisol Conjugate and 25µl Cortisol Antibody. Two parallel samples were used in each plate. After that the plate was shaken at room temperature for one hour. Before adding 100µl TMB Substrate and incubating 30 minutes the plate was tapped dry and aspirated 4 times with 300 µl of wash buffer. The incubation was terminated by adding 50µl of Stop Solution and the optical density was red of each well at 450 nm. The optical density analyses were conducted on MyAssays software that yielded the cortisol concentrations.

The accepted intra-assay coefficient of variation was <10% and all samples with duplicate intra-assay coefficients of variation >10% were reanalyzed.

2.4.2 Fecal glucocorticoid metabolites

The Fecal glucocorticoid metabolites were analyzed to obtain a more integrated level of circulating glucocorticoids that reflects an individual's stress exposure over a longer time (Harper & Austad, 2000). The FGM analyses were carried out for 78 elephants (males: 18 females: 60) at Veterinary Diagnostic Laboratory, Chiang Mai University, Thailand. Glucocorticoid metabolite concentrations were obtained via boiling extraction and measured using a double-antibody enzyme immunoassay with a polyclonal rabbit antibody (CJM006). The protocol followed a validated enzyme immunoassay (Watson et al., 2013). Absorbance was measured at 450 nm, and the accepted intra-assay coefficients of variation for duplicate samples were <10%. FGM concentrations varied between (29,7 to 103,5 ng/g/faeces) a range that has been previously observed in Asian elephants (Seltmann et al., 2020).

2.4.3 Heterophil and lymphocyte ratio

Heterophil and lymphocyte ratio was analyzed from the whole blood samples in a laboratory in Myanmar. To carry out the analysis blood cells were counted manually using a blood smear stained with Romanowsky stain and optical microscope with an amplification of 1000 x. The analysis was conducted within 12 hours of the sample collection. In total 68 blood samples were analyzed (males: 18 females: 50). The ratios were calculated based on the amounts of heterophils and lymphocytes in the blood sample slide. The H:L ratio varied between (0,2 to 3,9) which is considered as a normal range for elephants (Seltmann et al., 2020).

2.5 Statistical analysis

Statistical analyses of the gut microbiome and the stress measures were conducted using R Software (version 4.2.1). Statistical analysis focused on the three basic microbiome

analysis methods; analyzing alpha diversity (Shannon diversity), beta diversity, and differential abundance analysis (DAA). Also, bar plot analysis was used to visualize the microbiome composition. Microbiome alpha diversity, sometimes used with the term species diversity, summarizes the distribution of species abundances in one sample into a single number that depends on species richness and evenness. The species richness refers to the total number of species in one sample and the evenness focuses on species abundances. Then again, beta diversity quantifies similarities and dis-similarities between all of the samples. To focus particularly on the differences in the gut microbiome between groups, DA analysis was used.

2.5.1 Gut microbiome composition

First to answer the question about the overall composition of gut microbes in Asian elephants, the microbiome composition in males and females was visualized at the phylum level using bar plot. The analysis was conducted using the phyloseq R package where the composition was visualized using relative abundances (McMurdie & Holmes, 2013). Also, in order to understand what biologically meaningful explanatory factors might cause variation in the gut microbiome community, permutational multivariate analysis of variance (PERMANOVA) was conducted using adonis2 function on a vegan package (Oksanen et al., 2022). Adonis2 function runs an analysis of variance using distance matrices. The Bray-Curtis distance metric was chosen for the analysis based on the literature (Lahti et al., 2021) The independent variables in the model were camp (Kawlin, West Katha), age (fixed factor with four categories calf, juvenile, adult and senior), sex (male, female), and origin (captive or wild). The distance between samples, calculated by the phyloseq distance function, was the dependent variable. Before running PERMANOVA the homogeneous group dispersions (variances) were checked by plotting the dependent variables on R.

Age and camp effects on the microbiome were further analyzed in order to create the overall picture of the microbiome composition. The composition of gut microbes in different age groups was visualized on a genus level using bar plot and the analysis was conducted using the phyloseq R package (McMurdie & Holmes, 2013). The age and alpha diversity were analyzed later. The effect of the location to the composition of gut microbes was tested using differential abundance analysis. In order to determine which

species differ in abundance between the two camps, DESeq2 package was used (Love et al., 2014). DESeq2 uses shrinkage estimation for dispersions and log2 Fold-Changes to perform the quantitative analysis of differential expression of sequences corresponding to bacterial species.

Lastly the Shannon alpha diversity of the Asian elephants was assessed. Depending on the diversity index, alpha diversity can describe species richness (number of different species) or diversity (species richness and species evenness) in a single sample. Shannon alpha diversity summarizes the distribution of species abundances in a single sample into a number based on both species' richness and evenness. First the Shannon alpha diversity was calculated and visualized using the phyloseq R package. The Shannon alpha diversity was visualized by the camp and age. Second, the factors affecting Shannon alpha diversity were tested using a general linear model (GLM) from R software's own stats package. In the model the dependent variable was the Shannon alpha diversity. The Shannon alpha diversity for each individual elephant was calculated using the phyloseq estimate_richness function. The independent variables used in the model were age (calf, juvenile, adult and senior), sex (male, female), and camp (Kawlin, West Katha). The residuals were tested for over/underdispersion, outliers and for normal distribution (Kolmogorov–Smirnov test) using DHARMa package. The best model was chosen based on a lower AIC value between two biologically meaningful models. The analysis of the deviance table (type III test) was done by using the car R package (Fox & Weisberg, 2019). Post hoc analysis was conducted to the age groups using the emmeans package in R (Lenth, 2023).

2.5.2 Stress and gut microbiome composition

Next, I investigated how stress was connected to the gut microbiome composition. The measured stress values (FGM, H:L ratio, SC) were grouped in three categories: "low" "medium" and "high" stress. The categories were formed by dividing all of the values in three equal size groups. The stress groups and the average stress values are presented in Table 4 and the whole list of the stress values can be found in Supplement 5.

Table 4. The number of studied elephants (n), the average stress value, and the cut-off values for different measures (SC, FGM, H:L ratio). Stress grouped in three categories (low-stress, medium-stress and high-stress)

	SC	FGM	H:L ratio
n elephants	n=64	n=78	n=68
average	31.18 ng/ml	66.75 ng/g/feces	1.18
cut-off values	4.42- 86.24 ng/ml	29.73-103.56 ng/g/feces	0.22-3.92
Low	n=21	n=26	n=22
	15.26 ng/ml	48.15 ng/g/feces	0.665
	4.42-22.02 ng/ml	29.73-60.48 ng/g/feces	0.22-0.84
Medium	n=22	n=26	n=22
	26.41 ng/ml	66.56 ng/g/feces	1.05
	22.16-33.99 ng/ml	60.62-73.57ng/g/feces	0.84-1.27
High	n=21	n=25	n=21
	51.14 ng/ml	85.45 ng/g/feces	1.80
	34.48-86.24 ng/ml	73.89-103.56 ng/g/feces	1.31-3.92

First, I tested whether stress as a categorical variable or as a continuous variable affected the microbiome beta diversity. I tested this by using the same PERMANOVA analysis as previously and added different stress measures to the model one at a time. In the model the distance between samples, calculated by the phyloseq distance function, were the dependent variable and camp, age, sex, origin and FGM levels (high, medium, low) were the independent variables. After that the FGM levels were replaced in the model with continuous FGM values. The same was done for each stress measures (FGM, H:L ratio, SC).

After that I tested whether the stress levels were connected to the microbiome alpha diversity. The Shannon alpha diversity was calculated and visualized for all stress measures (FGM, H:L ratio, SC) and stress groups (low, medium, high) with the microbiome R package following the protocol of Lahti et al (2017). The Shannon alpha diversity and stress was analyzed using a general linear model. The same standard model was used as previously, where the dependent variable was the Shannon alpha diversity and independent variables were camp, age, sex and origin. Each stress measure (FGM, H:L ratio, SC) were added to the model one at a time using both categorical variables and continuous variables. The residuals were tested for over/underdispersion, outliers and for normal distribution (Kolmogorov–Smirnov test) using DHARMa package and the analysis of the deviance tables (type III test) were done by the car R package (Fox & Weisberg, 2019). Post hoc analysis was conducted to all of the stress groups using the emmeans package in R (Lenth, 2023).

The final analysis was conducted for all of the three stress measures separately. The DA analysis was conducted in order to see how different levels of stress and the composition of gut microbes might be connected, regarding specific microbial taxa. The analysis was conducted on all the stress measures FGM, H:L ratio, and SC comparing the low and high stress groups together. The covariates in the analysis were camp and age. The DESeq2 R package was used in the DA analysis, as previously.

3. Results

3.1 Gut microbiome composition

The first aim of this study was to determine the gut microbe composition of Asian elephants. The relative abundance of the most abundant phyla of the gut microbiome differed visually little between females and males (figure 1). The most abundant phylum was *Firmicutes* (55%) followed by *Bacteroidetes* (25%) These two phyla made up 80% of the total gut microbe composition in Asian elephants. The second most common phyla were *Spirochaetae* (8%), *Proteobacteria* (4.5%), *Actinobacteria* (2%) *and Lentisphaerae* (1.6%). One phylum of Archaea was also detected in elephant gut microbiome: *Euryarchaeota.*

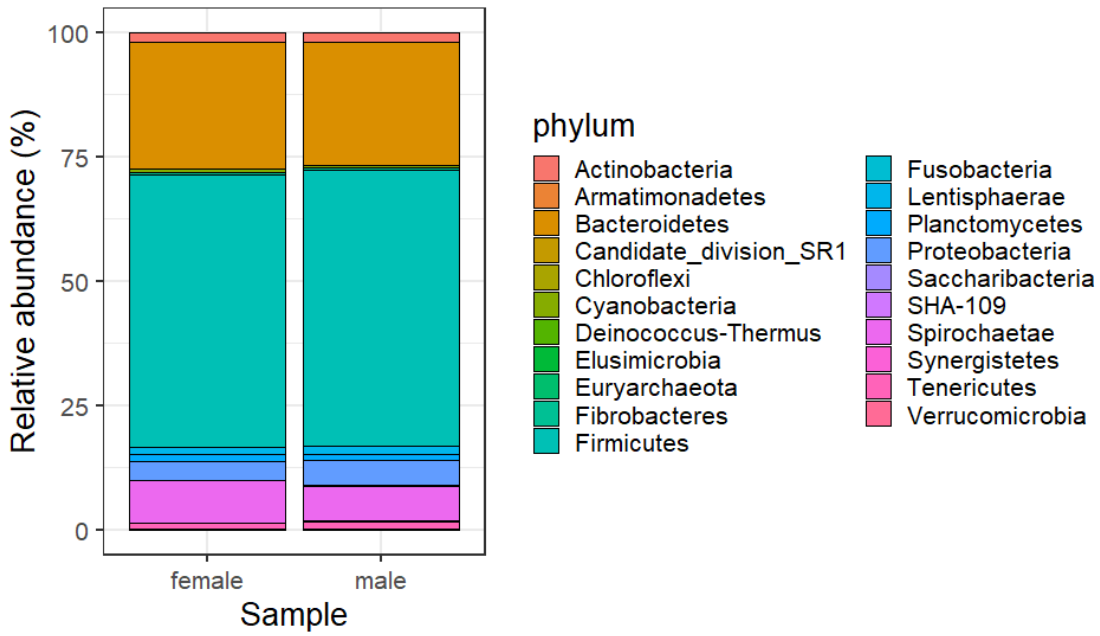


Figure 1. Gut microbiome composition of Asian elephants. The y-axis represents the relative abundance of the most abundant phyla of gut microbes. In the x-axis, samples are grouped for female (n=69) and male (n=25) elephants.

To determine which biologically meaningful factors such as sex, working location (logging camp), age, and origin (captive born or wild caught) formally associated with the composition of the gut microbiome, the beta diversity was tested using PERMANOVA (number of permutations: 999). The microbiome composition differed

significantly between logging camps (West Katha and Kawlin) (Table 5). Also, different age groups show statistically significant differences in the communities. Overall, camp explains 8 % of the total variation in the microbiome composition between individuals and age close to 5 %. Sex of the elephant and the origin does not statistically significantly influence the composition.

Table 5. The results of the PERMANOVA analysis. In the standard model the effects of camp, age, sex and origin to the microbiome beta diversity was assessed. Camp and age explain most of the variation in the microbiome community within the study population.

	Df	**SumOfSqs**	**R2**	**F**	**Pr(>F)**
Camp	1	1.497	0.083	8.215	0.001 ***
Age	3	0.891	0.049	1.629	0.003 **
Sex	1	0.242	0.013	1.329	0.111
Origin	1	0.173	0.010	0.951	0.493
Residual	84	15.309	0.845		
Total	90	18.112	1.000		

3.1.1 Age and gut microbiome

The most abundant bacteria in Asian elephants on a genus level is *Solibacillus,* followed by *Ruminococcaceae_UCG-005* and *Pseudobutyrivibrio* depending on the age group, all belonging to phylum *Firmicutes* (figure 2). Because age explained some of the variation in the microbiome composition, it was further visualized in the figure 2, where relative abundances of bacteria in different age groups (calves: 4–10 years old, juveniles: 11–20 years old, adults: 21–50 years old, seniors 51–72 years old) were visualized using bar blot. Depending on different age groups the following genus are *Rikenellaceae_rc9_gut_group*, *Trepomena_2* and *Eubacterium hallii group.* From the figure 2 can be seen that the genus *Rikenellaceae_rc9_gut_group* seems to decrease with age while *Ruminococcaceae_UCG-005* increases. Also, the gut microbe composition of calves and senior elephants to resemble each other the most, especially with the genus *Solibacillus* and *Pseudobutyrivibrio.* Whenever the genus *Solibacillus* is dominating the microbiome composition *Pseudobutyrivibrio* is present less than when *Solibacillus i*s not that dominant. In calves and seniors *Solibacillus* is dominating and *Pseudobutyrivibrio* is less present and in adults and juveniles *Solibacillus* is not dominating and the genus *Pseudobutyrivibrio* is almost as abundant as *Solibacillus.*

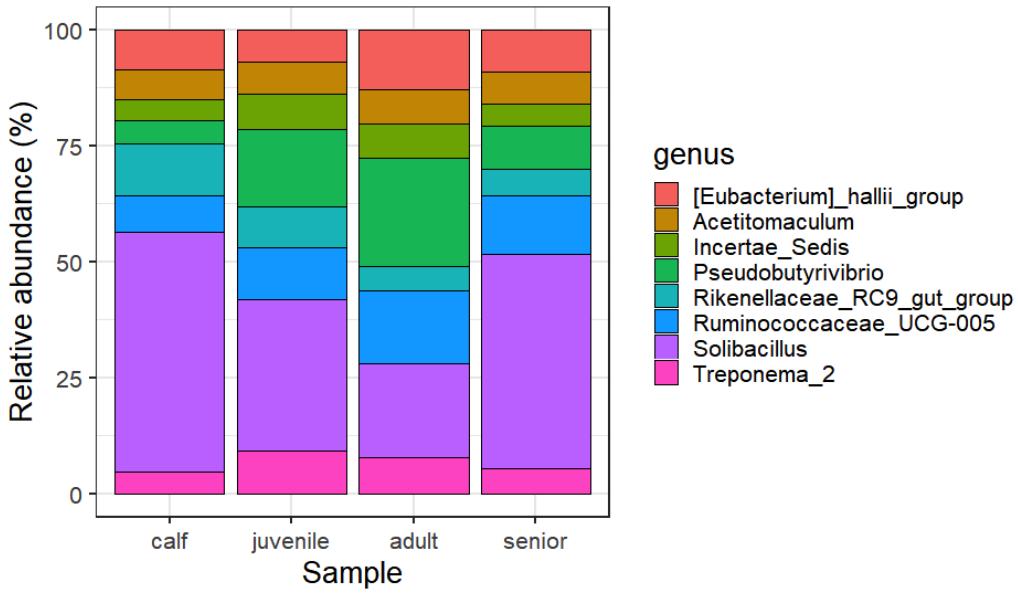


Figure 2. Gut microbe composition of Asian elephants on a genus level. The y-axis presents the relative abundance of the most abundant genus of gut microbes. In the x-axis, samples are grouped to age groups: calf (n=14), juvenile (n=13), adult (n=41), senior (n=26).

3.1.2 Camp and gut microbiome

In addition to age, the location where the elephant was working had an effect on the microbiome composition. The gut microbiome composition of elephants in Kawlin camp is more diverse than in elephants in West Katha camp (figure 3). The gut microbes that are more common in Kawlin are presented in the figure 3 y-axis as a negative log2 fold change, positive log2 fold change in the y-axis represents the microbes that are more common in West Katha. *Firmicutes* and *Bacteroidetes* are the most represented phyla in the figure and more common in Kawlin compared to West Katha. In phylum *Bacteroidetes* there are a lot of uncharacterized genera of bacteria, most of them belong to the family *Bacteroidales_S24-7_group* as well as class *Bacteroidales_BS11_gut_group* and *Porphyromonadaceae.* Phylum *Actinobacteria* is also more common in Kawlin. The uncharacterized genus of *Actinobacteria* belongs to the family *Coriobacteriaceae.* In addition to previous, phyla *Spirochaetae, Tenericutes and Proteobacteria* also have bacteria that are also more common in Kawlin.

The log2 fold change in family *Coriobacteriaceae* is the greatest in the figure, meaning that the difference between camps in that family of bacteria is also the greatest. From the

figure 3 it is also possible to see that there are only two genus of bacteria (*Lachnospiraceae_NK4A136_group* and *Rikenellaceae_RC9_gut_group*) that are found more often in West Katha camp.

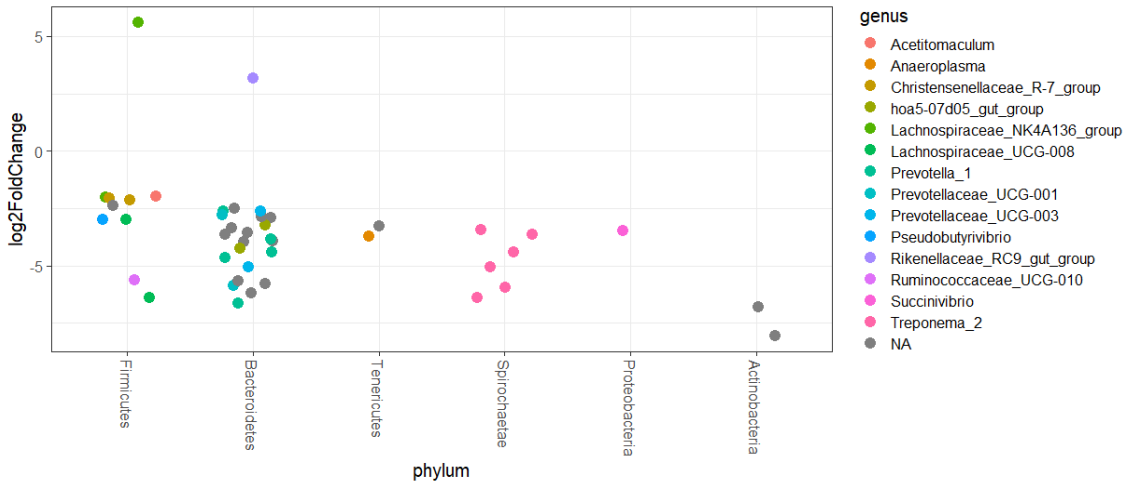


Figure 3. The difference in gut microbe composition between camps Kawlin (n= 53) and West Katha (n=38). Positive log2 fold change in y-axis represents microbes more common in West Katha and negative log2 fold change microbes more common in Kawlin. The greater the log2 fold change is the more likely it is that the genus is met on that camp compared to the other.

3.1.3 Alpha diversity and gut microbiome

Shannon alpha diversity for the study population is between 5.5 and 7.0 (figure 4). There are differences in the alpha diversity between individuals. In figure 4 the samples are grouped based on the age groups on the x-axis and the calculated Shannon diversity index is in the y-axis. The location of the camps is presented in the figure 4 in colors. The alpha diversity is higher in Kawlin than in West Katha. The alpha diversity changes with different age groups. The diversity is lower in calves and seniors and highest in the juveniles.

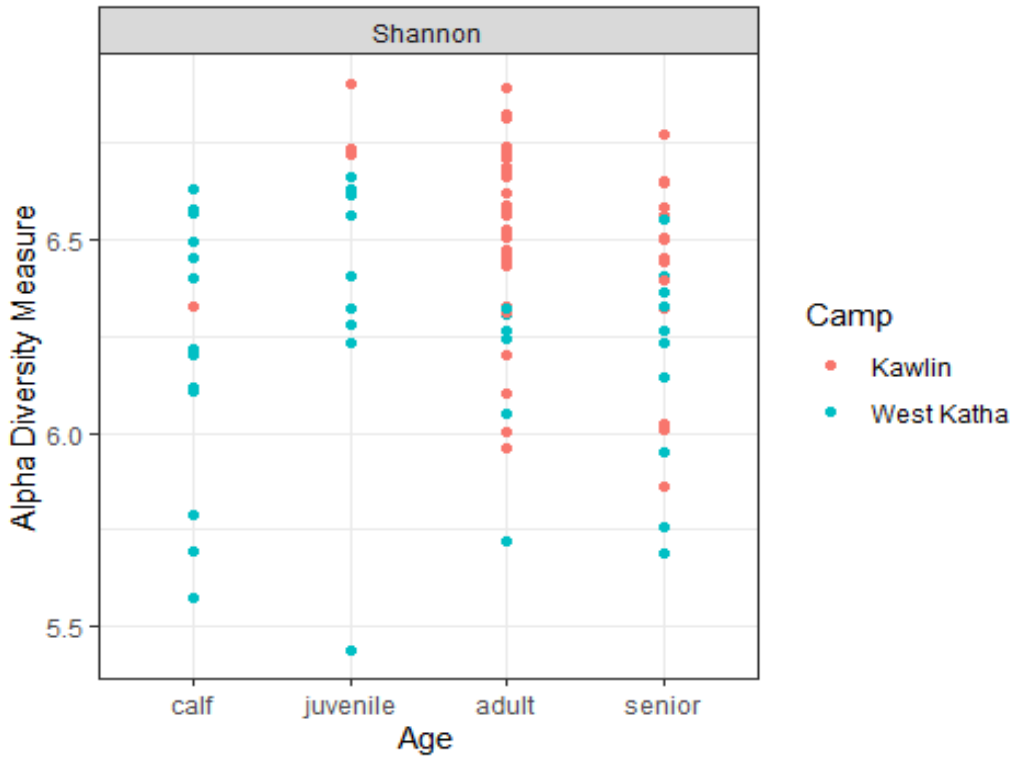


Figure 4. Shannon alpha diversity measure of Asian elephants by the age groups. Each dot in the picture represents the calculated alpha diversity of that elephant. Shannon alpha diversity measure is presented in the y-axis and the age groups of the elephants in the x-axis: calf (n=14), juvenile (n=13), adult (n=41), senior (n=26). The colors represent different camps.

The results of the GLM confirms the observed findings in the Figure 4 (Table 6). The model shows how the independent variables age, sex, camp and origin affect alpha diversity. From the independent variables, camp showed a statistically significant effect on the Shannon alpha diversity (GLM, p-value < 0.05, F-value= 18.2). Also, when the different age groups (calf, juvenile, adult, senior) were compared pairwise, averaged over the levels of camp, sex and origin, the comparisons between juvenile and senior age groups were statistically significant (p-value= 0,05) (Supplement 4). Sex or the origin of the elephant didn't affect the alpha diversity.

Table 6. The results of the general linear model (GLM). The standard model represents change in the alpha diversity in respect to age, camp, sex and origin.

	Df	**SumOfSqs**	**F value**	**Pr(>F)**
Camp	1	1.4068	18.2317	5.119e-05 ***
Age	3	0.3855	1.6652	0.1807
Sex	1	0.0001	0.0012	0.9726
Origin	1	0.0000	0.0005	0.9830
Residual	84	6.4817		

3.2 Stress and gut microbiome composition

There was no statistically significant association between stress measures and gut microbiome beta diversity (Table 7). The associations were tested on both categorical stress values (low, medium, high) (Table 7) and continuous variables (Supplement 4) and the stress measures were added to the standard model one by one.

Table 7. Results of three PERMANOVA analysis where stress measures (FGM, H:L ratio, SC) were added to the standard model one at a time. Stress groups (low, medium, high) did not significantly explain the variation in the gut microbiome in the population.

	Df	SumOfSqs	R2	F	Pr(>F)
Camp	1	1.19	0.084	6.88	0.001 ***
Age	3	0.863	0.061	1.67	0.001 ***
Sex	1	0.216	0.015	1.25	0.134
Origin	1	0.197	0.014	1.15	0.228
FGM	2	0.344	0.024	0.999	0.449
Residual	66	11.4	0.802		
Total	74	14.2	1.00		

	Df	SumOfSqs	R2	F	Pr(>F)
Camp	1	0.968	0.078	5.51	0.001 ***
Age	3	0.790	0.064	1.50	0.006 **
Sex	1	0.262	0.021	1.49	0.043 *
Origin	1	0.172	0.014	0.977	0.439
H:L	2	0.377	0.030	1.07	0.259
Residual	56	9.84	0.793		
Total	64	12.4	1.00		

	Df	SumOfSqs	R2	F	Pr(>F)
Camp	1	0.922	0.078	5.25	0.001 ***
Age	3	0.991	0.084	1.88	0.001 ***
Sex	1	0.209	0.018	1.19	0.166
Origin	1	0.180	0.015	1.02	0.369
SC	2	0.386	0.033	1.10	0.221
Residual	52	9.12	0.772		
Total	60	11.8	1.00		

Stress however was associated with the alpha diversity of the gut microbiome. Stress measure H:L ratios showed statistically significant association to the Shannon alpha

diversity (GLM, p-value < 0.05, F-value= 3.68, adjusted for age, camp, sex, origin) whereas FGM and SC measures did not (Table 8).

Table 8. Results of three GLM analyses where stress measures (FGM, H:L ratio, SC) were added to the standard model one at a time. Stress measure H:L ratio showed statistically significant association to the Shannon alpha diversity. FGM and SC measures were not associated to the Shannon alpha diversity.

	Df	SumOfSqs	F value	Pr(>F)
Camp	1	0.667	12.4	0.0008 ***
Age	3	0.396	2.44	0.072 .
Sex	1	0.022	0.410	0.524
Origin	1	0.008	0.148	0.701
FGM	2	0.017	0.156	0.855
Residual	66	3.57		

	Df	SumOfSqs	F value	Pr(>F)
Camp	1	0.768	13.6	0.0005 ***
Age	3	0.436	2.58	0.062 .
Sex	1	0.184	3.27	0.076 .
Origin	1	0.020	0.365	0.548
H:L	2	0.415	3.67	0.032 *
Residual	56	3.15		

	Df	SumOfSqs	F value	Pr(>F)
Camp	1	0.778	11.4	0.001 **
Age	3	0.767	3.75	0.016 *
Sex	1	0.201	2.95	0.092 .
Origin	1	0.087	1.28	0.263
SC	2	0.006	0.047	0.954
Residual	52	3.54		

From the Figure 5b it is possible to see that in the H:L ratio measure the low stress group is higher in the y-axis than the high stress group, indicating higher Shannon alpha diversity in the low stress group compared to the high. The same trend can be seen in the FGM measure (Figure 5a), but there is no statistical support for the difference. In the post hoc analysis of H:L ratio measure for stress groups (low, medium, high) the statistical significance was between the low and high stress group (p-value = 0.0008) (Supplement 4).

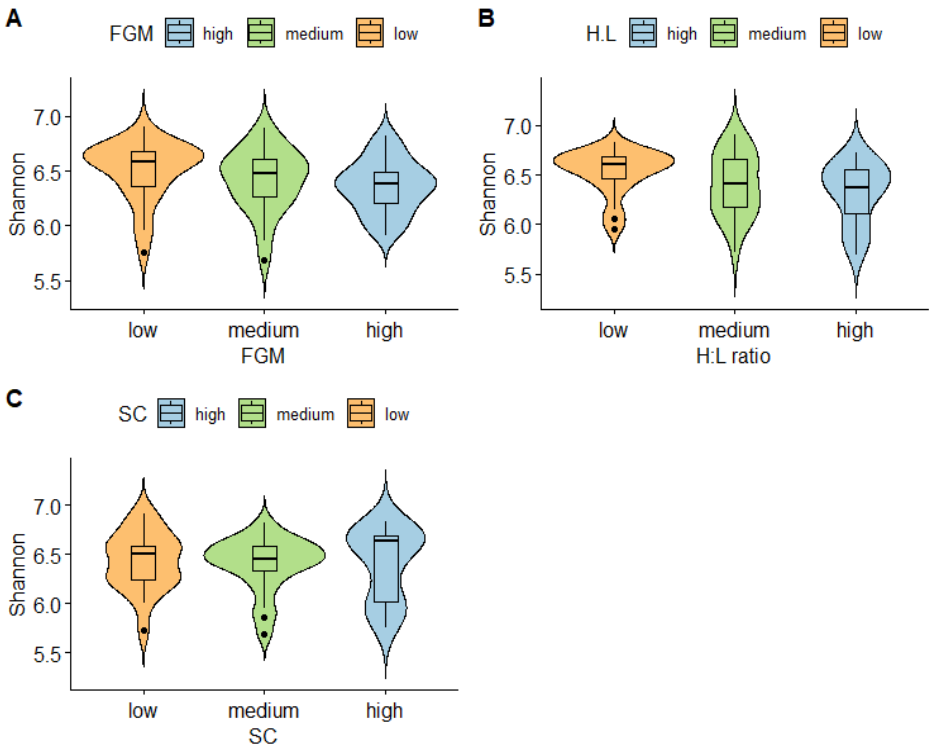


Figure 5. Shannon alpha diversity and the three different stress measures A: FGM, B: H:L ratio, C: SC. In the figure stress groups represent calculated Shannon alpha diversity of the gut microbiome on elephants belonging to that stress group. The diversity measure is not controlled for covariates. (FGM: low n=26, medium n=26 high n=25, H:L: low n=22, medium n=22 high n=21, SC: low n=21, medium n=22 high n=21)

Finally, the stress measures and microbiome composition were also analyzed using differential abundance analysis. The DA analysis showed statistically significant differences in the abundances of individual taxonomic groups between low and high stress groups in each stress measure, but there is no clear consistency in the results between the three stress measures. All measures (A: FGM, B: H:L ratio, C: SC) have an effect on the gut microbiome composition, but the affected taxa are different (Figure 6).

In the Figure 6, the y-axis represents log2 fold change of specific microbial taxa. Positive log2 fold change represents microbes that are more abundant in low stress groups and negative log2 fold change microbes that are more abundant in high stress groups. The threshold for statistical difference, used in the Desq2 algorithm, was p=0.01. Age and camp were used as covariates in the analysis.

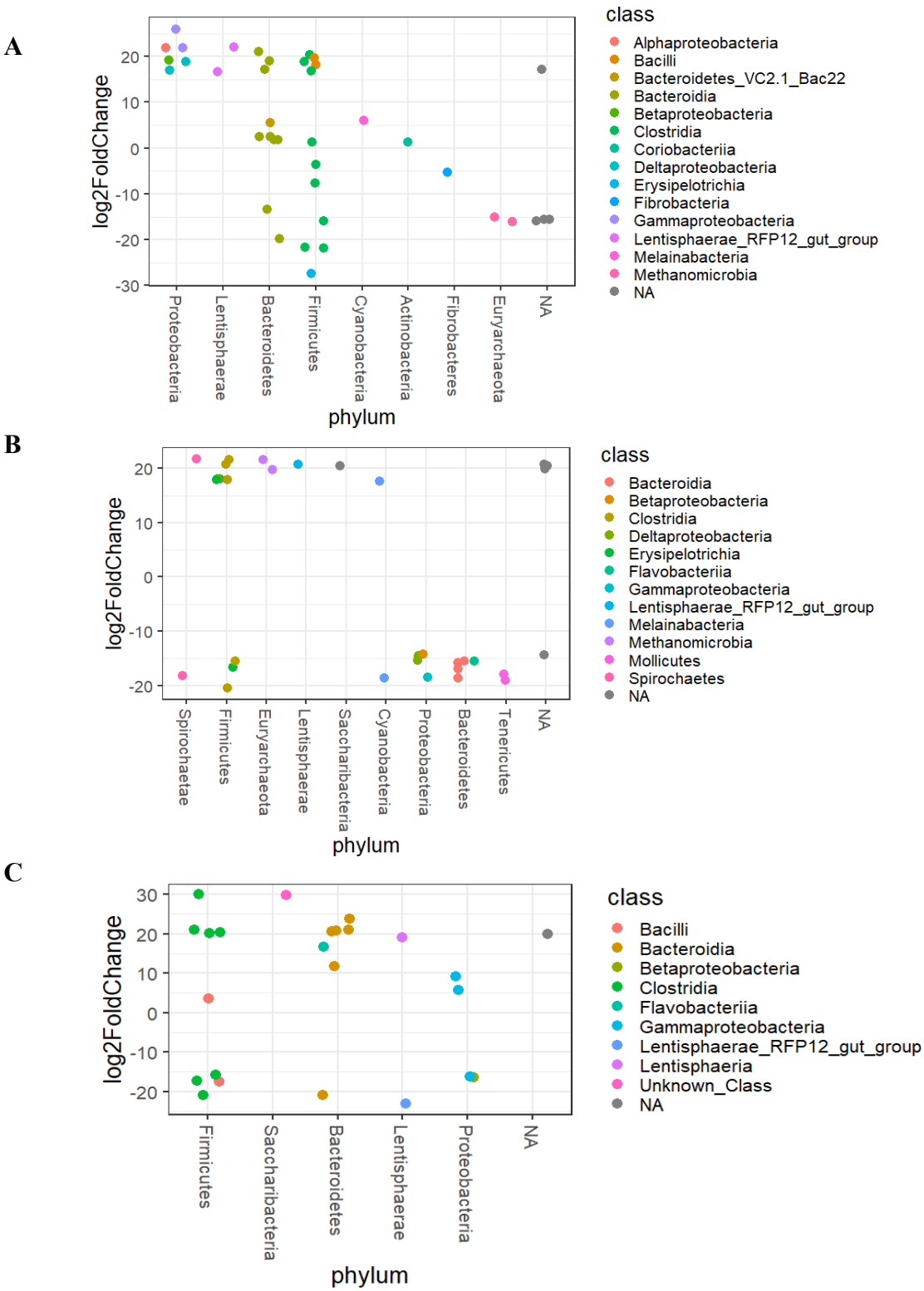


Figure 6. Differential abundance analysis between high and low stress groups in A: FGM, B: H:L ratio, and C: SC measures (FGM: low n=26, high n=25, H:L ratio: low n=22, high n=21 SC: low n=21, high n=21). In the figure each point represents a unique sequence that was found to be significantly higher in abundance in different stress groups. Positive log2 fold change represents high stress group and negative low2 stress group. The greater the log2 fold change is, the bigger the difference in abundance between compared groups.

Lentisphaerae is the only phylum that increases in abundance in all of the stress measures in high stress groups (Figure 6). Although no specific class were associated with all measures, class *Lentisphaerae_RFP12_gut_group* is increasing in both FGM and H:L ratio measures. In SC measure the increase is in class *Lentisphaeria* in genus *Victivallis*. Conversely, the class *Lentisphaerae_RFP12_gut_group* in SC measure is increasing in low stress groups. In phylum *Firmicutes* the class *Clostridia* is also increasing with stress in all of the measures. From the family *Ruminococcaceae* the genus *Ruminococcaceae_UCG-010* is increasing in H:L ratio and FGM measures and genus *Ruminococcaceae_UCG-002* in SC measure.

Another common feature in the Figure 6 is the increase in phylum *Bacteroidetes* in high stress groups in two of the stress measures (FGM and SC). Order *Bacteroidales* from phylum *Bacteroidetes* significantly increased in abundance in high stress groups with one common genus *Prevotellaceae_UCG-003*. In the Figure 6 the H:L ratio measure is acting differently to FGM and SC measures. The phyla *Proteobacteria* and *Bacteroidetes* is for example increasing in high stress groups in both FGM and SC measures, but not in H:L ratio measure. Also, the abundance of *Euryarchaeota* is behaving oppositely in FGM and H:L ratio measures. Notably, there are as many differences as there are similarities in the responses of the microbiome to different stress measures.

By viewing the Figure 6 it is possible to see that there are some genera of bacteria that have greater log2 fold change compared to others. The greatest log2 fold change in the Figure 6 can be seen in the phylum *Firmicutes.* In Figure 6A genus *Erysipelotrichaceae_UCG-004* from class *Erysipelotrichia* is significantly increased in the low stress group (Log2FoldChange= -27,4 p-value= 2.470499e-17). On the other hand, in Figure 6C genus *Ruminococcaceae_UCG-002* from class *Clostridia* is significantly more abundant in the high stress group (Log2FoldChange= 30,0 p-value= 2.83e-22). By generalizing the results, what it means in practice is that genus Ruminococcaceae*_UCG-002* is more abundant in elephants suffering from high levels of stress and genus *Erysipelotrichaceae_UCG-004* on the other hand is more common in elephants that are not showing physical signs of stress.

4. Discussion

This thesis provided a comprehensive picture of the composition of Asian elephant's gut microbiome by studying 94 semi-captive Asian elephants living in their natural environment, and with known identity, age and background information available, enabling the largest microbiome assessment of Asian elephants to-date. In addition to successfully characterizing the overall microbiome composition and its variation among individuals, this study also assessed the interaction between stress and gut microbes by utilizing a rare opportunity for a long-lived species to combine microbiome data with a range of stress measures obtained from blood sampling and glucocorticoid metabolite analysis. I found out that increases in the H:L ratio stress measures were associated with decline in microbiome alpha diversity, and that stress was associated with the changes of specific gut microbial taxa. The method established here to study gut microbes in elephant feces provides a convenient and reusable approach for future studies. The results of this thesis deepen our knowledge, not only on elephant gut microbiome, but also how stress and gut microbiome are linked to each other in general.

4.1 Gut microbiome composition

My first hypothesis was that the Asian elephant gut microbiome resembles other fermenters and that it is most dominated by phyla *Firmicutes* and *Bacteroidetes,* the two most common phyla on gut microbiome regardless of the species. My results confirmed that the same applied on Asian elephants and phyla *Firmicutes* (55%) and *Bacterioidetes* (25%) made up a total 80% of the gut microbiome. What comes after these two are more interesting and species specific. Previous studies on Asian elephants have provided contradictory information on the following phyla, but were all based on a small sample size ranging between 3 to 30 (Ilmberger et al., 2014; Moustafa et al., 2021; Zhang et al., 2019). The one common phylum in all the studies was *Spirochaetaes.* P*roteobacteria and Fibrobacteres* were abundant in two of the studies and phyla *Actinobacteria and Lentisphaerae* were reported to be abundant in one study. These phyla are similar to what I detected in this thesis, but their relative abundances differed. In this thesis the most abundant phyla after *Firmicutes* and *Bacterioidetes* were *Spirochaetae, Proteobacteria,*

Actinobacteria, and *Lentisphaerae.* These results confirm the previous findings of the composition of gut microbes at the phylum level. Furthermore, what these findings suggest is that the composition and abundance of different taxa is more related to the environment than elephant physiology.

On a genus level there was more variation between previous studies. A study conducted in China for three Asian elephants suggested that the most abundant genera were Fibrobacter, followed by *Treponema, Prevotella, Bacteroides, Bacillus, Butyrivibrio*, and *Ruminococcaceae* (Zhang et al., 2019). In another study conducted in a German zoo for two elephants, the most abundant genera were *Lachnospiraceae, Ruminococcaceae, Prevotellaceae* and the *RC9 gut group of Rickenellaceae* (Ilmberger et al., 2014). Only *Ruminococcaceae* were abundant in both studies. In this thesis I also found genus *Ruminococcaceae* to be abundant (Figure 2). Other previously reported genera that were also detected in this thesis were *Treponema, Butyrivibrio, and the RC9 gut group of Rickenellaceae.* Interestingly, what I found out to be the most abundant genus in the semi-captive elephants in Myanmar was not reported in other studies. Genus *Solibacillus* was the most abundant genus in this thesis, ranging from 20% to 50% in relative abundance, depending on the age of the elephant (Figure 2). In addition to limited sample size in previous studies based on mostly zoo animals, one reason for such clear differences in gut microbiome composition between studies could be the methods chosen for the analysis that can affect the detected composition. Nevertheless, in wild African buffalos, the two most abundant genera were *Solibacillus* and *Ruminococcaceae UCG-005* (Couch et al., 2021). These two genera, more specifically the enrichment of the genus, were also linked to putative enterotypes in wild African buffalo gut microbiome. In microbiome studies enterotypes are often linked to a stable gut microbiome which is dependent on long-term diet and host health. (Couch et al., 2021) In the African buffalo study the presence and enrichment of genus *Solibacillus* was linked to an enterotype prevalent under restricted dietary conditions whereas the enrichment of *Ruminococcaceae-UCG-005* was linked to resource-abundant dietary regimes. Many other genera were also similar between semi-captive Asian elephants and wild African buffalos. This highlights the importance of studying species in their natural environment. The previous studies on Asian elephants conducted in a zoo environment might not offer a comprehensive overall picture of the Asian elephant gut microbiome. In one recent study, the same effect was noticed where different zoo facilities had a strong influence on the Asian elephant gut microbiome community (Keady et al., 2021).

The abundance of *Solibacillus* in the Asian elephant gut microbiome, especially as great abundance as detected in this thesis, is possibly linked to restrictions in vegetation. During restricted resource periods, *Solibacillus* is known to increase in abundance. It can be because it is adapted to resource-restricted regimes and therefore can more likely survive in the gut better than other taxa. In some mammals, goats and cattle, genus *Solibacillus* has also been associated with reduction in forage intake, which can possibly be an adaptive response to increased dietary variability. (Couch et al., 2021) The abundance of *Solibacillus* in Asian elephant gut shown in this thesis calls for further research to determine the causes for this. For example, all samples analyzed in this thesis were collected at the turn of the hot and rainy season in Myanmar, when food abundance is at its lowest and elephant body weight declines. How seasonal variation in the Asian elephants' natural range area affects their gut microbiome composition remains to be studied.

4.1.1 Age and gut microbiome composition

I expected age to affect the composition of gut microbes in the study population that ranged from 1 to 67-year-old. In this thesis I observed variation in both alpha diversity and beta diversity between the study elephants. Alpha diversity among juveniles (11–20 years old) was higher than that of other age groups. The alpha diversity was lower in calves and seniors, and higher in juveniles and adults (Figure 3). The same trend has been detected also in human studies (Xu et al., 2019) as well as in other Asian elephant studies (Li et al., 2022). In the Li et al. (2022) study the diversity was reported to be highest in elephants at 25-year-old and then show a decreasing trend with age. The sample size in that study was however only two elephants per age group (calf: tow 3-year-old, juveniles: 13 and 18-year-old, adults: two 25-year-old, and old adults: two 40-year-old semi-captive elephants) but the results were in line with my observations.

Age was also linked to beta diversity of the gut microbiome (Table 5), with compositional changes across age in the core gut microbiome at the genus level (Figure 2). The biggest changes can be seen in the most abundant genus *Solibacillus*. In calves and seniors, the relative abundance is close to 50% and in juveniles the abundance is just 30% and in adults even lower. Abundance of genus *Pseudobutyrivibrio* behaves opposite to

Solibacillus meaning that it is more prevalent in juveniles and adults and less in calves and seniors. The effect of aging on the microbiome composition cannot be assessed from Figure 2 alone and it would require future testing. It is still the most comprehensive illustration of the relative abundances of gut microbes on semi-captive Asian elephants in different age groups.

4.1.2 Camp and gut microbiome composition

I expected the working camps of the elephants to have an effect on the composition of gut microbes within the study population. The location in which the elephants work, live and eat are referred to as the camps. In Figure 4, Shannon alpha diversity in camp Kawlin is significantly higher than in camp West Katha. The differences in alpha diversity can be due to the differences in vegetation of the areas, which are directly influenced by climate patterns. The climate in Kawlin is rainier than in west Katha. Average amount of rain at the time of sampling in Kawlin was 100 mm whereas in West Katha it was 37 mm. The rainfall should have a high impact on the plant diversity and quality, and this would be expected to explain some of the variation between the camps in alpha diversity.

The difference between the camps was also observed in the microbiome beta diversity (Table 5) and in community comparison analysis (Figure 3). In Kawlin phylum *Bacteroidetes* is clearly more abundant than in West Katha. The reason why this is, can also be linked to the vegetation of the area. *Bacteroidetes* phylum is the main cellulosic plant material degraders in herbivores, especially bacteria belonging to families *Bacteroidales*, *Bacteroidaceae* and *Bacteroides* (Li et al., 2022). The genera belonging to *Bacteroidales* in Kawlin is *Prevotellaceae UCG-001, Prevotellaceae UCG-03, Prevotella 1* and also two classes of *Bacteroidales: Bacteroidales_S24-7_group* and *Bacteroidales_BS11_gut_group.* The differences in the microbiome between Kawlin and West Katha do not only restrict to phylum *Bacteroidetes.* The genus *Succinivibrio* from phylum *Sphirochaetae* and genus *Treponema_2* from *Proteobacteria* is also more abundant in Kawlin. The abundance of *Succinivibrio* could also be explained by the vegetation, as it is known as Prevotella-type plant polysaccharide-fermenting commensal (Tan & Nie, 2020). The abundance of genus *Treponema_2* in camp Kawlin and its possible association to nutrient digestibility is something that should be studied more. There are previous studies on fermenters about nutrient digestibility and microbial

fermentation, where genus *Treponema_2* is mentioned (Greene et al., 2019; Mi et al., 2022). In one Asian elephant study the genus *Treponema* was also associated with elephants being semi-captive rather than captive elephants (Moustafa et al. 2021). Phylum *Spirochetes* plays an important role in degrading plant polymers materials such as xylan, pectin and arabinogalactan, *Treponema_2* particularly is associated with high pectin content. (Mi et al., 2022) Interestingly, many of the microbes associated with complex carbohydrate fermentation found in the caecum and colon of elephants are similar to those found in the rumen of ruminants. The same phenomenon has been noticed in the literature before (Greene et al., 2019). The main differences between the camps Kawlin and West Katha are most likely a result of different diets, which would explain the increase in different plant material degrading genus in Kawlin.

4.2 Stress and gut microbiome

In this thesis the main question was to assess the connection between stress and gut microbiome composition. My hypothesis was that I would find a decrease in the microbial richness in stressed elephants, in line with studies on stress and gut microbiome proposing such an association (Maltz et al., 2018; Mir et al., 2019; Wiley et al., 2017). I also expected to find variation in abundances in different bacterial taxa, for example reduction in genus *Lactobacillus* and increase in *Proteobacteria* in association with stress. In this thesis I found differences in alpha diversity in respect to different levels of stress. I also detected changes in the abundances of specific bacterial taxa in all of the used stress measures. However, I did not find reduction in genus *Lactobacillus* or a clear increase in *Proteobacteria* in association with stress.

In order to ask the question about the connection between gut microbiome and stress, one must first establish a way of measuring stress. I decided to use three physical stress measures, previously reported in the literature (Harper & Austad, 2000; Seltmann et al., 2020). The three measures were chosen to represent the overall stress at the time of the measurement. FGM and H:L ratio represented the long-term stress measures, and the SC was more an indication of acute stress reaction. Stress can reshape the gut bacterial composition through glucocorticoid secretion, inflammation, or autonomic alterations. The outcome of this is often gut bacterial imbalance, called dysbiosis, as well as low alpha diversity. In turn, the gut bacteria release metabolites, toxins, and neurohormones that can

affect the brain functions for example the eating behavior and in some cases also upregulate stress responsiveness. (Madison & Kiecolt-Glaser, 2019) The two-way interaction between the gut and the brain is important to keep in mind while interpreting the results of stress and gut microbes in this thesis.

The results on alpha diversity and stress, presented in the Figure 4 and Table 8, are partly in line with the hypothesis as well as previous studies (Maltz et al., 2018; Mir et al., 2019; Wiley et al., 2017). The decrease in alpha diversity with stress was detected for long-term stress measure that is descriptive for chronic stress. The statistical differences were detected between the lowest and the highest stress groups of H:L ratio measure, but no difference in the alpha diversity was detected in another long term stress measure FGM. I was not expecting to see changes in the alpha diversity with SC measure, for it works as an indication of acute stress, and in order to detect changes in the gut microbiome the physiological pathways of stress must have time to affect the gut microbe composition. Also, the causality of changes in the microbiome alpha diversity in respect to stress cannot be assessed. In more interventional studies the causality could be assessed, whereas in a natural setting such as the semi-captive elephants the drivers of stress differences between individuals are likely complex and diverse.

The compositional changes in the gut microbiome that associates with stress are visualized in the Figure 6. The absence of clear consistency in the results makes the interpretation of the results difficult. The results could be viewed by looking at each stress measurement separately, or by grouping it based on the biology behind the measures. One possible way of interpreting the results is looking the stress measures FGM and H:L ratio together. The FGM and the H:L ratio are both physiological markers of stress that reflect more integrated level of glucocorticoids, and therefore work as a measure of stress exposure over a longer time (Seltmann et al., 2020). The common feature in the fFigure 6 in these measures was the increase in class *Lentisphaerae_RFP12_gut_group* in high stress groups. There were no further taxonomic assignments for this class in the SILVA database, and in my knowledge, not many studies focusing on this class of bacteria either. It is therefore hard to find evidence supporting this finding and future research is needed to determine whether this class of bacteria is associated with chronic stress. In both FGM and H:L ratio measures class *Clostridia* from phylum *Firmicutes* were abundant in high stress group. Genus *Ruminococcaceae_UCG-010* from class *Clostridia* is referred as

pathogen in one study (Chen et al., 2019) and in other study it is associated with high sugar diets and decrease in *Firmicutes* to *Bacteroidetes* (F:B) ratio (Yue et al., 2019).

Another way of interpreting the results is to look the FGM and SC measures together. Both measures are more direct measures of glucocorticoid hormones, compared to H:L ratio. The abundance of phylum *Bacteroidetes* is increasing in high stress groups in these measures. One common genus *Prevotellaceae_UCG-003* was found to be significantly higher amongst stressed elephants. *Prevotellaceae_UCG-003* is a genus that is associated with dysbiosis (Z. L. Wu et al., 2022). The phylum *Proteobacteria* also increased in FGM an SC measures. Especially clear change was observed with the FGM measure, where all *Alpha-, Beta-, Delta-, and Gammaproteobacteria* increased in high stress groups. Reduction in F:B ratio, caused by the increase in *Bacteroidetes*, as well as increase in *Proteobacteria* is previously reported to be the indicator of dysbiosis (Mir et al., 2019; Shin et al., 2015). Genus *Acinetobacter* from class *Gammaproteobacteria* was increased in both SC and FGM measures in high stress groups. Gram-negative *Acinetobacter* genus is common in soil but also capable of occupying several ecological niches, including the mammalian intestine. The genus is often associated with infections. (Glover et al., 2022; Visca et al., 2011) Infections caused by *Acinetobacter* genus are also reposted to cause inflammatory responses in animal studies. For example, it induces inflammatory cytokines, such as IL-8 and tumor necrosis factor (TNF). (Wong et al., 2017) The low-grade inflammation is reported to be associated both with chronic diseases and with gut microbiome composition which are both associated with stress. Other *Proteobacteria* increased in high stress group in FGM measure were *Alphaproteobacteria Brevundimonas,* which is a non-fermenting, gram-negative, opportunistic pathogen (Ryan & Pembroke, 2018). Also, *Stenotrophomonas* from *Gammaproteobacteria* and *GR-WP33-58* from *Deltaproteobacteria* increased in high stress groups.

One study on Asian elephants shows that captivity induced stress increased the abundance of class *Clostridia* and *Bacteroidia* in the microbiome (Moustafa et al., 2021). In this thesis the number of bacteria in class *Clostridia* were also more abundant in high stress group compared to low stress groups in both FGM and SC measures. Also, class *Bacteroidia* increased in high stress group, but only for the FGM measure.

4.3 Strengths and limitations

This thesis has several strengths and limitations. One of the limitations is related to the overall limitation of microbiome analysis. By doing all of the steps from laboratory to computer analysis I gained understanding of the whole gut microbiome analysis pipeline, but this also introduces limitations. Because of the limited time I was not able to spend as much time on the statistical analysis of the data. This can affect the results as it is recommended in microbiome data analysis to use multiple different tests to confirm the results, because different methods use different approaches (parametric vs non-parametric, different normalization techniques, assumptions etc.) (Nearing et al., 2022). This is a common challenge in all microbiome analysis that needs to be considered.

The results can also be affected by methodological causes, that can explain the observed variation between different studies in Asian elephant gut microbiome. First, the fecal samples were collected in ethanol, and only after the samples were sent to the University of Turku, were they frozen. The procedure was the same for all stress groups, but could potentially affect the overall composition of gut microbes, especially to the relative abundances of different microbial taxa. Second, the DNA extraction method or the sequencing can cause bias to the gut microbe composition. Third, it is worth mentioning that in the laboratory analysis the final quality check of the sequencing library using Bioanalyzer produced data that was untypical (Supplement 3). (Leigh Greathouse et al., 2019)

One of the limitations of this thesis has to do with the Asian elephant study population. Myanmar has three seasons, and the effect of seasonality to the gut microbiome could be more closely studied. Season and other ecological conditions like vegetation or seasonal workload might have effects on the microbiome composition and stress levels. However, to minimize heterogeneity caused by such effects, samples used in this thesis were all collected at the turn of the hot and rainy season when none of the elephants had been working since mid-February. Also, there could have been more male elephants in this thesis in order to get the sex ratio more equal as well as more young elephants to get the age groups more equal. The statistical power was not limited in this thesis in regard to microbiome analysis, on the other hand there could have been more measures for the stress analysis.

Strengths of this thesis was the rare opportunity to study the gut microbiome composition of long-lived non-human species in their natural conditions, with known background information. The background information includes information about the exact birth dates, cause of death, maternal information, medications and other health parameters. All of this background information is especially important in gut microbiome studies because there are many factors affecting the complex ecosystems of gut microbes. One of the strengths of this thesis is also the used measures of stress. Three different measures, if working correctly to measure the stress levels in different aspects, gives a comprehensive picture of the stress level of each elephant.

4.4 Conclusions

In this thesis I was able to conduct a study answering the question about the composition of the Asian elephant gut microbiome and how stress associates with it. My results have diverse relevance. First, the characterization of the semi captive Asian elephant gut microbiome has value for previous studies that have mostly focused on zoo animals only. There is a limited number of previous studies on the composition of gut microbes in Asian elephants, considering that the importance of gut microbes is nowadays much appreciated. In this thesis I detected one especially abundant genus *Solibacillus* in Asian elephants that was not previously reported as an abundant genus in that species. This finding highlights the importance of studying animals in their natural environments as well as the importance of increasing studies focusing on the gut microbiome of various different species. By understanding the structure and dynamics of microbial communities in wild populations it is possible to find ecological patterns in host–microbiome relationships that can help to understand these complicated relationships and add to our knowledge on gut microbes also in humans.

Second, the differences in the gut microbiome composition between elephants in different working camps is of relevance. This is something that was also reported to Myanma Timber Enterprise (MTE) who owns and takes care of the elephants' health. The difference in the microbiome alpha diversity between the camps, with alpha diversity being lower in West Katha, is important to acknowledge. Lower levels of diversity are

associated with several acute and chronic diseases, and are also linked to chronic stress (Manor et al., 2020). In the case of alpha diversity differences observed in this thesis the probable cause is the different diets of the elephants in different camps. I didn't test in this thesis whether the stress levels of elephants in different camps would be different, but this would be interesting to assess in future studies.

Third, this thesis gives a snapshot of the current levels of stress of Asian elephants working in the logging industry in Myanmar. In addition, this thesis tries to assess the complicated relationship between gut microbiome and brain function, using these stress levels as a response to brain function. The results about stress and gut microbiome are partly in line with previous literature. The strongest results were observed with stress measure H:L ratio and microbiome alpha diversity. With increasing stress, the diversity of microbes decreases. If the same would have been detected also in the FGM stress measure the results would have been stronger and indicating the link between chronic stress and lower microbial alpha diversity.

Lastly, the methods used in this study provide tools for the future analysis of elephant gut microbiome in challenging environments. One of the key findings in the methods was to see that fecal samples could be preserved in ethanol before DNA extraction. In the literature the fecal samples are often immediately frozen after, but in many distant locations, like forest in Myanmar, that is not possible. The ethanol preserved fecal samples worked in the analysis, but it must be kept in mind while interoperating the results. This thesis aims for full reproducibility of the results, with the methods described in detail, and the bioinformatics pipeline for the data analysis in the supplements.

5. Acknowledgements

Firstly, I want to thank my supervisors Virpi Lummaa and Eero Vesterinen for their endless patience and help during this process. I would also like to thank Manu Tamminen for introducing me to the workflow of microbial analysis and for valuable help in bioinformatics. A special thanks to Meri Lindqvist for all the help in the laboratory. Finally, I want to thank MTE for the provided samples, Diogo dos Santos and Jennifer Crawley and others not mentioned here, but who contributed to the process.

6. References

Amato, K. R. (2013). Co-evolution in context: The importance of studying gut microbiomes in wild animals. *Microbiome Science and Medicine*, *1*(1). https://doi.org/10.2478/micsm-2013-0002

Barko, P. C., McMichael, M. A., Swanson, K. S., & Williams, D. A. (2018). The Gastrointestinal Microbiome: A Review. *Journal of Veterinary Internal Medicine*, *32*(1), 9. https://doi.org/10.1111/JVIM.14875

Boers, S. A., Jansen, R., & Hays, J. P. (2019). Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. *European Journal of Clinical Microbiology & Infectious Diseases*, *38*(6), 1059. https://doi.org/10.1007/S10096-019-03520-3

Chen, R., Wang, J., Zhan, R., Zhang, L., & Wang, X. (2019). Fecal metabonomics combined with 16S rRNA gene sequencing to analyze the changes of gut microbiota in rats with kidney-yang deficiency syndrome and the intervention effect of You-gui pill. *Journal of Ethnopharmacology*, *244*, 112139. https://doi.org/10.1016/J.JEP.2019.112139

Couch, C. E., Stagaman, K., Spaan, R. S., Combrink, H. J., Sharpton, T. J., Beechler, B. R., & Jolles, A. E. (2021). Diet and gut microbiome enterotype are associated at the population level in African buffalo. *Nature Communications 2021 12:1*, *12*(1), 1–11. https://doi.org/10.1038/s41467-021-22510-8

Crawley, J. A. H., Lierhmann, O., Santos, D. J. F., Brown, J., Nyein, U. K., Aung, H. H., Htut, W., Oo, Z. M., Seltmann, M. W., Webb, J. L., Lahdenperä, M., & Lummaa, V. (2021). Influence of handler relationships and experience on health parameters, glucocorticoid responses and behaviour of semi-captive Asian elephants. *Conservation Physiology*. https://doi.org/10.1093/conphys/coaa116

Crawley, J.A.H., Lahdenperä, M., Min Oo, Z., Htut, W., Nandar, H., & Lummaa, V. (2020). Taming age mortality in semi-captive Asian elephants. *Scientific Reports*, *10*(1). https://doi.org/10.1038/s41598-020-58590-7

Cryan, J. F., O'riordan, K. J., Cowan, C. S. M., Sandhu, K. V., Bastiaanssen, T. F. S., Boehme, M., Codagnone, M. G., Cussotto, S., Fulling, C., Golubeva, A. V., Guzzetta, K. E., Jaggar, M., Long-Smith, C. M., Lyte, J. M., Martin, J. A., Molinero-Perez, A., Moloney, G., Morelli, E., Morillas, E., … Dinan, T. G. (2019). The microbiota-gut-brain axis. *Physiological Reviews*, *99*(4), 1877–2013. https://doi.org/10.1152/PHYSREV.00018.2018/ASSET/IMAGES/LARGE/Z9J004 1929160006.JPEG

Davis, A. K., Maney, D. L., & Maerz, J. C. (2008). The use of leukocyte profiles to measure stress in vertebrates: a review for ecologists. *Functional Ecology*, *22*(5), 760–772. http://doi.wiley.com/10.1111/j.1365-2435.2008.01467.x

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J. B., Massart, S., Collini, S., Pieraccini, G., & Lionetti, P. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(33), 14691–14696. https://doi.org/10.1073/PNAS.1005963107/-/DCSUPPLEMENTAL

Dickens, M. J., & Romero, L. M. (2013). A consensus endocrine profile for chronically stressed wild animals does not exist. *General and Comparative Endocrinology*, *191*, 177–189. https://doi.org/10.1016/J.YGCEN.2013.06.014

Edgar, R. C., & Bateman, A. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460–2461. https://doi.org/10.1093/BIOINFORMATICS/BTQ461

Farzi, A., Fröhlich, E. E., & Holzer, P. (2018). Gut Microbiota and the Neuroendocrine System. *Neurotherapeutics*, *15*(1), 5. https://doi.org/10.1007/S13311-017-0600-5

Francino, M. P. (2018). Birth Mode-Related Differences in Gut Microbiota Colonization and Immune System Development. *Annals of Nutrition and Metabolism*, *73*(3), 12–16. https://doi.org/10.1159/000490842

Fox J, Weisberg S (2019). *An R Companion to Applied Regression*, Third edition. Sage, Thousand Oaks CA. https://socialsciences.mcmaster.ca/jfox/Books/Companion/. (read 01/2023)

Glover, J. S., Browning, B. D., Ticer, T. D., Engevik, A. C., & Engevik, M. A. (2022). Acinetobacter calcoaceticus is Well Adapted to Withstand Intestinal Stressors and Modulate the Gut Epithelium. *Frontiers in Physiology*, *13*, 1. https://doi.org/10.3389/FPHYS.2022.880024/FULL

Golofast, B., & Vales, K. (2020). The connection between microbiome and schizophrenia. *Neuroscience & Biobehavioral Reviews*, *108*, 712–731. https://doi.org/10.1016/J.NEUBIOREV.2019.12.011

Goodrich, J. K., Waters, J. L., Poole, A. C., Sutter, J. L., Koren, O., Blekhman, R., Beaumont, M., Van Treuren, W., Knight, R., Bell, J. T., Spector, T. D., Clark, A. G., & Ley, R. E. (2014). Human Genetics Shape the Gut Microbiome. *Cell*, *159*(4), 789–799. https://doi.org/10.1016/J.CELL.2014.09.053

Greene, W., Dierenfeld, E. S., & Mikota, S. (2019). A Review of Asian and African Elephant Gastrointestinal Anatomy, Physiology, and Pharmacology. *Journal of Zoo and Aquarium Research*, *7*(1), 1–14. https://doi.org/10.19227/JZAR.V7I1.329

Harper, J. M., & Austad, S. N. (2000). Fecal Glucocorticoids: A Noninvasive Method of Measuring Adrenal Activity in Wild and Captive Rodents. *Https://Doi.Org/10.1086/316721*, *73*(1), 12–22. https://doi.org/10.1086/316721

Herlemann, D. P., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J. J., & Andersson, A. F. (2011). Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *The ISME journal*, *5*(10), 1571-1579.

Hill, J. H., & Round, J. L. (2021). SnapShot: Microbiota effects on host physiology. *Cell*, *184*, 2796-2796.e1. https://doi.org/10.1016/j.cell.2021.04.026

Ilmberger, N., Güllert, S., Dannenberg, J., Rabausch, U., Torres, J., Wemheuer, B., Alawi, M., Poehlein, A., Chow, J., Turaev, D., Rattei, T., Schmeisser, C., Salomon, J., Olsen, P. B., Daniel, R., Grundhoff, A., Borchert, M. S., & Streit, W. R. (2014). A Comparative Metagenome Survey of the Fecal Microbiota of a Breast- and a Plant-Fed Asian Elephant Reveals an Unexpectedly High Diversity of Glycoside Hydrolase Family Enzymes. *PLOS ONE*, *9*(9), e106707. https://doi.org/10.1371/JOURNAL.PONE.0106707

Keady, M. M., Prado, N., Lim, H. C., Brown, J., Paris, S., & Muletz-Wolz, C. R. (2021). Clinical health issues, reproductive hormones, and metabolic hormones associated with gut microbiome structure in African and Asian elephants. *Animal Microbiome*, *3*(1). https://doi.org/10.1186/S42523-021-00146-9

Lahdenperä, M., Mar, K. U., Courtiol, A., & Lummaa, V. (2018). Differences in age-specific mortality between wild-caught and captive-born Asian elephants. *Nature communications*, *9*(1), 3023.

Lahti, L., Shetty, S., et al. (2017). Tools for microbiome analysis in R. Version. URL: http://microbiome.github.com/microbiome. (read 07/2022)

Lahti, L., Shetty, S., Broman, T., Ernst, F.M. (2021). Orchestrating Microbiome Analysis with Bioconductor [Beta Version]. microbiome.github.io/oma/. (read 12/2022)

Leigh Greathouse, K., Sinha, R., & Vogtmann, E. (2019). DNA extraction for human microbiome studies: the issue of standardization. *Genome Biology*, *20*(1). https://doi.org/10.1186/S13059-019-1843-8

Lenth R (2023). _emmeans: Estimated Marginal Means, aka Least-Squares Means_. R package version 1.8.4-1, <https://CRAN.R-project.org/package=emmeans>. (read 01/2023)

Li, G., Jiang, Y., Li, Q., An, D., Bao, M., Lang, L., Han, L., Huang, X., & Jiang, C. (2022). Comparative and functional analyses of fecal microbiome in Asian elephants. *Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology*, *115*(9), 1187–1202. https://doi.org/10.1007/S10482-022-01757-1/FIGURES/9

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 1–21. https://doi.org/10.1186/S13059-014-0550-8/FIGURES/9. (read 12/2022)

Million, M, Lagier, J. C., Yahav, D., & Paul, M. (2013). Gut bacterial microbiota and obesity. *Clinical Microbiology and Infection*, *19*(4), 305–313. https://doi.org/10.1111/1469-0691.12172

Madison, A., & Kiecolt-Glaser, J. K. (2019). Stress, depression, diet, and the gut microbiota: human–bacteria interactions at the core of psychoneuroimmunology and nutrition. *Current Opinion in Behavioral Sciences*, *28*, 105–110. https://doi.org/10.1016/j.cobeha.2019.01.011

Maltz, R. M., Keirsey, J., Kim, S. C., Mackos, A. R., Gharaibeh, R. Z., Moore, C. C., Xu, J., Bakthavatchalu, V., Somogyi, A., & Bailey, M. T. (2018). Prolonged restraint stressor exposure in outbred CD-1 mice impacts microbiota, colonic inflammation, and short chain fatty acids. *PLoS ONE*, *13*(5). https://doi.org/10.1371/JOURNAL.PONE.0196961

Mangiola, F., Nicoletti, A., Gasbarrini, A., & Ponziani, F. R. (2018). Gut microbiota and aging. *European Review for Medical and Pharmacological Sciences*, *22*(21), 7404–7413. https://doi.org/10.26355/EURREV_201811_16280

Manor, O., Dai, C. L., Kornilov, S. A., Smith, B., Price, N. D., Lovejoy, J. C., Gibbons, S. M., & Magis, A. T. (2020). Health and disease markers correlate with gut microbiome composition across thousands of people. *Nature Communications*, *11*(1). https://doi.org/10.1038/S41467-020-18871-1

Marcobal, A., Barboza, M., Sonnenburg, E. D., Pudlo, N., Martens, E. C., Desai, P., Lebrilla, C. B., Weimer, B. C., Mills, D. A., German, J. B., & Sonnenburg, J. L. (2011). Bacteroides in the Infant Gut Consume Milk Oligosaccharides via Mucus-Utilization Pathways. *Cell Host & Microbe*, *10*(5), 507. https://doi.org/10.1016/J.CHOM.2011.10.007

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10–12. https://doi.org/10.14806/EJ.17.1.200

McMurdie, P. J., & Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PloS one*, *8*(4), e61217.

Messaoudi, M., Lalonde, R., Violle, N., Javelot, H., Desor, D., Nejdi, A., Bisson, J. F., Rougeot, C., Pichelin, M., Cazaubiel, M., & Cazaubiel, J. M. (2011). Assessment of psychotropic-like properties of a probiotic formulation (Lactobacillus helveticus R0052 and Bifidobacterium longum R0175) in rats and human subjects. *British Journal of Nutrition*, *105*(5), 755–764. https://doi.org/10.1017/S0007114510004319

Mi, H., Ren, A., Zhu, J., Ran, T., Shen, W., Zhou, C., Zhang, B., & Tan, Z. (2022). Effects of different protein sources on nutrient disappearance, rumen fermentation

parameters and microbiota in dual-flow continuous culture system. *AMB Express*, *12*(1), 1–10. https://doi.org/10.1186/S13568-022-01358-1/TABLES/6

Mir, R. A., Kleinhenz, M. D., Coetzee, J. F., Allen, H. K., & Kudva, I. T. (2019). Fecal microbiota changes associated with dehorning and castration stress primarily affects light-weight dairy calves. *PLoS ONE*, *14*(1). https://doi.org/10.1371/journal.pone.0210203

Misiak, B., Łoniewski, I., Marlicz, W., Frydecka, D., Szulc, A., Rudzki, L., & Samochowiec, J. (2020). The HPA axis dysregulation in severe mental illness: Can we shift the blame to gut microbiota? *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *102*, 109951. https://doi.org/10.1016/J.PNPBP.2020.109951

Molina-Torres, G., Rodriguez-Arrastia, M., Roman, P., Sanchez-Labraca, N., & Cardona, D. (2019). Stress and the gut microbiota-brain axis. *Behavioural Pharmacology*, *30*(2and3-SpecialIssue), 187–200. https://doi.org/10.1097/FBP.0000000000000478

Moustafa, M. A. M., Chel, H. M., Thu, M. J., Bawm, S., Htun, L. L., Win, M. M., Oo, Z. M., Ohsawa, N., Lahdenperä, M., Mohamed, W. M. A., Ito, K., Nonaka, N., Nakao, R., & Katakura, K. (2021). Anthropogenic interferences lead to gut microbiome dysbiosis in Asian elephants and may alter adaptation processes to surrounding environments. *Scientific Reports*, *11*(1), 741. https://doi.org/10.1038/s41598-020-80537-1

Musazzi, L., & Marrocco, J. (2016). The Many Faces of Stress: Implications for Neuropsychiatric Disorders. *Neural Plasticity*, *2016*. https://doi.org/10.1155/2016/8389737

Nearing, J. T., Douglas, G. M., Hayes, M. G., MacDonald, J., Desai, D. K., Allward, N., Jones, C. M. A., Wright, R. J., Dhanani, A. S., Comeau, A. M., & Langille, M. G. I. (2022). Microbiome differential abundance methods produce different results across 38 datasets. *Nature Communications*, *13*(1), 342. https://doi.org/10.1038/S41467-022-28034-Z

Odamaki, T., Kato, K., Sugahara, H., Hashikura, N., Takahashi, S., Xiao, J. Z., Abe, F., & Osawa, R. (2016). Age-related changes in gut microbiota composition from newborn to centenarian: A cross-sectional study. *BMC Microbiology*, *16*(1), 90. https://doi.org/10.1186/s12866-016-0708-5

Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J (2022). _vegan: Community Ecology Package_. R package version 2.6-4, <https://CRAN.R-project.org/package=vegan>. (read 12/2022)

Palma, G. De, Collins, S. M., Bercik, P., & Verdu, E. F. (2014). The microbiota–gut–brain axis in gastrointestinal disorders: stressed bugs, stressed brain or both? *The Journal of Physiology*, *592*(Pt 14), 2989. https://doi.org/10.1113/JPHYSIOL.2014.273995

Peirce, J. M., & Alviña, K. (2019). The role of inflammation and the gut microbiome in depression and anxiety. *Journal of Neuroscience Research*, *97*(10), 1223–1241. https://doi.org/10.1002/jnr.24476

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, *41*(Database issue), D590. https://doi.org/10.1093/NAR/GKS1219

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, *4*(10).

https://doi.org/10.7717/PEERJ.2584
Rojas, C. A., Ramírez-Barahona, S., Holekamp, K. E., & Theis, K. R. (2021). Host phylogeny and host ecology structure the mammalian gut microbiota at different taxonomic scales. *Animal Microbiome*, *3*(1). https://doi.org/10.1186/S42523-021-00094-4
Ryan, M. P., & Pembroke, J. T. (2018). Brevundimonas spp: Emerging global opportunistic pathogens. *Virulence*, *9*(1), 480. https://doi.org/10.1080/21505594.2017.1419116
Sapolsky, R. M., Romero, L. M., & Munck, A. U. (2000). How Do Glucocorticoids Influence Stress Responses? IntegratingPermissive, Suppressive, Stimulatory, and Preparative Actions. *Endocrine Reviews*, *21*(1), 55–89. https://doi.org/10.1210/EDRV.21.1.0389
Savage, D. C. (1977). Microbial Ecology of the Gastrointestinal Tract. *Annual Review of Microbiology*, *31*(1), 107–133. https://doi.org/10.1146/annurev.mi.31.100177.000543
Seltmann, M. W., Ukonaho, S., Reichert, S., Dos Santos, D., Nyein, U. K., Htut, W., & Lummaa, V. (2020). Faecal Glucocorticoid Metabolites and H/L Ratio Are Related Markers of Stress in Semi-Captive Asian Timber Elephants. *Animals: An Open Access Journal from MDPI*, *10*(1). https://doi.org/10.3390/ANI10010094
Sharpton, T. J., Stapleton, A. E., Ben-Hur, A., & Blanchard, J. (2014). *An introduction to the analysis of shotgun metagenomic data*. https://doi.org/10.3389/fpls.2014.00209
Shin, N. R., Whon, T. W., & Bae, J. W. (2015). Proteobacteria: microbial signature of dysbiosis in gut microbiota. *Trends in Biotechnology*, *33*(9), 496–503. https://doi.org/10.1016/J.TIBTECH.2015.06.011
Strandwitz, P. (2018). Neurotransmitter modulation by the gut microbiota. *Brain Research*, *1693*(Pt B), 128. https://doi.org/10.1016/J.BRAINRES.2018.03.015
Sukumar, R. (2006). A brief review of the status, distribution and biology of wild Asian elephants Elephas maximus. *International Zoo Yearbook,* 40(1), 1-8.
Tan, H., & Nie, S. (2020). Deciphering diet-gut microbiota-host interplay: Investigations of pectin. *Trends in Food Science & Technology*, *106*, 171–181. https://doi.org/10.1016/J.TIFS.2020.10.010
Tanaka, M., & Nakayama, J. (2017). Development of the gut microbiota in infancy and its impact on health in later life. *Allergology International*, *66*(4), 515–522. https://doi.org/10.1016/J.ALIT.2017.07.010
Tannock, G. W., & Savage, D. C. (1974). Influences of Dietary and Environmental Stress on Microbial Populations in the Murine Gastrointestinal Tract. *Infection and Immunity*, *9*(3), 591–598. https://doi.org/10.1128/IAI.9.3.591-598.1974
Thursby, E., & Juge, N. (2017). Introduction to the human gut microbiota. *Biochemical Journal*, *474*(11), 1823. https://doi.org/10.1042/BCJ20160510
Ursell, L. K., Metcalf, J. L., Parfrey, L. W., & Knight, R. (2012). *Defining the human microbiome*. https://doi.org/10.1111/j.1753-4887.2012.00493.x
Vesterinen, E. J., Ruokolainen, L., Wahlberg, N., Peña, C., Roslin, T., Laine, V. N., Vasko, V., Sääksjärvi, I. E., Norrdahl, K., & Lilley, T. M. (2016). What you need is what you eat? Prey selection by the bat Myotis daubentonii. *Molecular Ecology*, *25*(7), 1581–1594. https://doi.org/10.1111/mec.13564
Visca, P., Seifert, H., & Towner, K. J. (2011). Acinetobacter infection – an emerging threat to human health. *IUBMB Life*, *63*(12), 1048–1054. https://doi.org/10.1002/IUB.534
Watson, R., Munro, C., Edwards, K. L., Norton, V., Brown, J. L., & Walker, S. L. (2013). Development of a versatile enzyme immunoassay for non-invasive assessment of glucocorticoid metabolites in a diversity of taxonomic

species. *General and Comparative Endocrinology*, *186*, 16-24.
Wiley, N. C., Dinan, T. G., Ross, R. P., Stanton, C., Clarke, G., & Cryan, J. F. (2017). The microbiota-gut-brain axis as a key regulator of neural function and the stress response: Implications for human and animal health. *Journal of Animal Science*, *95*(7), 3225–3246. https://doi.org/10.2527/JAS.2016.1256
Willing, B. P., Russell, S. L., & Finlay, B. B. (2011). Shifting the balance: antibiotic effects on host–microbiota mutualism. *Nature Reviews Microbiology 2011 9:4*, *9*(4), 233–243. https://doi.org/10.1038/NRMICRO2536
Wong, D., Nielsen, T. B., Bonomo, R. A., Pantapalangkoor, P., Luna, B., & Spellberg, B. (2017). Clinical and Pathophysiological Overview of Acinetobacter Infections: A Century of Challenges. *Clinical Microbiology Reviews*, *30*(1), 409. https://doi.org/10.1128/CMR.00058-16
Wu, G. D., Bushmanc, F. D., & Lewis, J. D. (2013). Diet, the human gut microbiota, and IBD. *Anaerobe*, *24*, 117–120. https://doi.org/10.1016/J.ANAEROBE.2013.03.011
Wu, G. D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y.-Y., Keilbaugh, S. A., Bewtra, M., Knights, D., Walters, W. A., Knight, R., Sinha, R., Gilroy, E., Gupta, K., Baldassano, R., Nessel, L., Li, H., Bushman, F. D., & Lewis, J. D. (2011). Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes. *Science (New York, N.Y.)*, *334*(6052), 105. https://doi.org/10.1126/science.1208344
Wu, Z. L., Wei, R., Tan, X., Yang, D., Liu, D., Zhang, J., & Wang, W. (2022). Characterization of gut microbiota dysbiosis of diarrheic adult yaks through 16S rRNA gene sequences. *Frontiers in Veterinary Science*, *9*. https://doi.org/10.3389/FVETS.2022.946906/FULL
Xia, Y., Sun, J., & Chen, D.G., (2018). In: Statistical Analysis of Microbiome Data with R, pp. 5, Springer Nature Singapore Pte Ltd. 2018
Xu, C., Zhu, H., & Qiu, P. (2019). Aging progression of human gut microbiota. *BMC Microbiology*, *19*(1), 1–10. https://doi.org/10.1186/S12866-019-1616-2/FIGURES/4
Yue, S., Zhao, D., Peng, C., Tan, C., Wang, Q., & Gong, J. (2019). Effects of theabrownin on serum metabolites and gut microbiome in rats with a high-sugar diet. *Food & Function*, *10*(11), 7063–7080. https://doi.org/10.1039/C9FO01334B
Zhang, C., Xu, B., Lu, T., & Huang, Z. (2019). Metagenomic Analysis of the Fecal Microbiomes of Wild Asian Elephants Reveals Microflora and Enzymes that Mainly Digest Hemicellulose. *Journal of Microbiology and Biotechnology*, *29*(8), 1255–1265. https://doi.org/10.4014/jmb.1904.04033
Zheng, D., Liwinski, T., & Elinav, E. (2020). Interaction between microbiota and immunity in health and disease. *Cell Research*, *30*(6), 492. https://doi.org/10.1038/S41422-020-0332-7
Zoetendal, E. G., Akkermans, A. D. L., Akkermans-van Vliet, W. M., De Visser, J. A. G. M., & De Vos, W. M. (2001). The host genotype affects the bacterial community in the human gastrointestinal tract. *Microbial Ecology in Health and Disease*, *13*(3), 129–134. https://doi.org/10.1080/089106001750462669

Supplement materials

Supplement 1:

i5_sequence_in_adapter

1	TGCATGAT	33	CACTTCGA	101	GGTATACT
2	TGACATCG	34	CAGCGTTA	102	GTAGTACG
3	TTGCCTAG	35	CATACCAA	103	GTACACGT
4	TGTGGTCT	36	CCAGTTCA	69	AACTCACC
5	TCCACTGT	37	CCGAAGTA	70	AAGAGATC
6	GCATTGGT	38	CCGTGAGA	71	AAGGACAC
7	TGGATCTG	39	CCTCCTGA	72	AATCCGTC
8	GCTCAAGT	40	CGAACTTA	73	AATGTTGC
9	CGCTGATC	41	CGACTGGA	74	ACACGACC
10	ACAAGCTA	42	CGCATACA	75	ACAGATTC
11	CTGTAGCC	43	CTCAATGA	76	AGATGTAC
12	AGTACAAG	44	CTGAGCCA	77	AGCACCTC
13	GTCAACCT	45	CTGGCATA	78	AGCCATGC
14	TTCCGAGT	46	GAATCTGA	79	AGGCTAAC
15	TTCGCTTG	47	CAAGACTA	80	ATAGCGAC
16	GTGACGGT	48	GAGCTGAA	81	ATCATTCC
17	TAGGTACT	49	GATAGACA	82	ATTGGCTC
18	GCACAGAG	50	GCCACATA	83	CAAGGAGC
19	TCAGCAGT	51	GCGAGTAA	84	CACCTTAC
20	GCCTCCAG	52	GCTAACGA	85	CCATCCTC
21	ACGCTCGT	53	GCTCGGTA	86	CCGACAAC
22	ACGTATCG	54	GGAGAACA	87	CCTAATCC
23	ACTATGCG	55	GGTGCGAA	88	CCTCTATC

24	AGAGTCAT	56	GTACGCAA	89	CGACACAC
25	AGATCGCT	57	GTCGTAGA	90	CGGATTGC
26	TGCAGGTG	58	GTCTGTCA	91	CTAAGGTC
27	AGTCACTG	59	GTGTTCTA	92	GAACAGGC
28	ATCCTGTG	60	TAGGATGA	93	GACAGTGC
29	ATTGAGGT	61	TATCAGCA	94	GAGTTAGC
30	CAACCACA	62	TCCGTCTA	95	GATGAATC
31	GACTAGTA	63	TCTTCACA	96	GCCAAGAC
32	CAATGGAA	64	TGAAGAGA		
33	CACTTCGA	100	GTCACGAT		

Supplement 2:

SPRI bead solution recipe

See Vesterinen *et al.* 2016 for further specifications and references.

Materials

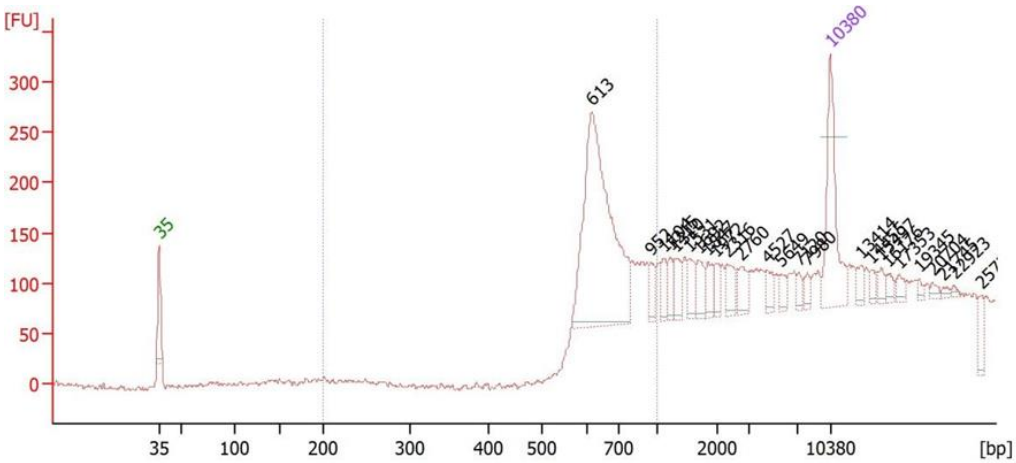
- Sera-Mag magnetic carboxylate modified particles (Sigma-Aldrich GE44152105050250)
- PEG-8000 (Sigma-Aldrich 89510-250G-F)
- 0.5 M EDTA, pH 8.0
- 1.0 M Tris, pH 8.0
- Tween 20 (Sigma-Aldrich P1379-25ML)
- 5 M NaCL
- 50 bp ladder (Thermo Scientific (Life Technologies) SM0371)
- Rare-earth magnet stand (Ambion AM10055 or NEB S1506S)

Sera-Mag SPRI bead substitute recipe

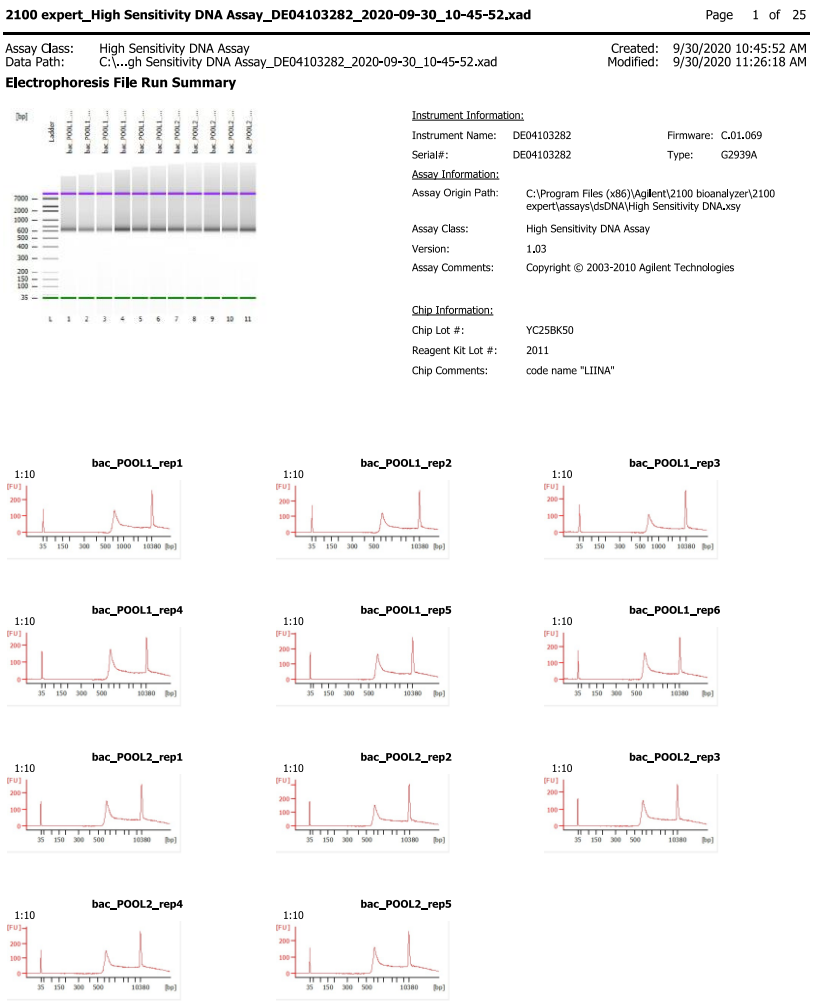
2 Mix Sera-mag SpeedBeads and transfer 1 mL to a 1.5 mL microtube.
3 Place SpeedBeads on magnet stand until beads are drawn to magnet.
4 Remove supernatant with P200 or P1000 pipetter.
5 Add 1 mL TE to beads, remove from magnet, mix, and return to magnet.
6 Remove supernatant with P200 or P1000 pipetter.
7 Add 1 mL TE to beads, remove from magnet, mix, and return to magnet.
8 Remove supernatant with P200 or P1000 pipetter.
9 Add 1 mL TE to beads and remove from magnet. Fully re-suspend and set tube in rack (i.e. not on magnet stand).
10 Add 9 g PEG-8000 to a new 50 mL, sterile conical.
11 Add 10 mL 5 M NaCL (or 2.92 g) to conical.
12 Add 500 µL 1 M Tris-HCL to conical.
13 Add 100 µL 0.5 M EDTA to conical.
14 Fill conical to ~ 49 mL using sterile dH20. You can do this by eye, just go slowly.
15 Mix conical for about 3-5 minutes until PEG goes into solution (solution, upon sitting, should be clear).
16 Add 27.5 uL Tween 20 to conical and mix gently.
17 Mix 1 mL SpeedBead + TE solution and transfer to 50 mL conical.
18 Fill conical to 50 mL mark with dH20 (if not already there) and gently mix 50 mL conical until brown.
19 Test against AMPure XP using aliquots of ladder (Fermentas GeneRuler). I recommend the 50 bp ladder in place of the ultra-low range ladder.
20 Wrap in tinfoil (or place in dark container) and store at 4°C.
21 Test monthly – see Testing, next page.

Supplement 3:

BioAnalyzer (FFGC) final sequencing library



BioAnalyzer (University of Turku)



2100 expert_High Sensitivity DNA Assay_DE04103282_2020-09-30_10-45-52.xad Page 1 of 25

Assay Class: High Sensitivity DNA Assay
Data Path: C:\...gh Sensitivity DNA Assay_DE04103282_2020-09-30_10-45-52.xad

Created: 9/30/2020 10:45:52 AM
Modified: 9/30/2020 11:26:18 AM

Electrophoresis File Run Summary

Instrument Information:

Instrument Name:	DE04103282	Firmware:	C.01.069
Serial#:	DE04103282	Type:	G2939A

Assay Information:

Assay Origin Path:	C:\Program Files (x86)\Agilent\2100 bioanalyzer\2100 expert\assays\dsDNA\High Sensitivity DNA.xsy
Assay Class:	High Sensitivity DNA Assay
Version:	1.03
Assay Comments:	Copyright © 2003-2010 Agilent Technologies

Chip Information:

Chip Lot #:	YC25BK50
Reagent Kit Lot #:	2011
Chip Comments:	code name "LIINA"

bac_POOL1_rep1 bac_POOL1_rep2 bac_POOL1_rep3
bac_POOL1_rep4 bac_POOL1_rep5 bac_POOL1_rep6
bac_POOL2_rep1 bac_POOL2_rep2 bac_POOL2_rep3
bac_POOL2_rep4 bac_POOL2_rep5

2100 Expert (B.02.10.SI764) © Copyright 2003 - 2018 Agilent, Inc. Printed: 9/30/2020 11:56:58 AM

Supplement 4:

Script (first bash script and second R script)

#----------CSC----------

File primers.fst contains these two lines:
>341_805
CCTACGGGNGGCWGCAG...GGATTAGATACCCBDGTAGTC

Script file contains these lines (server-specific lines removed):

```
VSEARCH=vsearch
USEARCH=usearch11.0.667

trunclen=250
QC=fastq_maxee
maxee=2
maxdiffs=50
mergeminlen=400
minlen=350
primererr=0.20
uniqsize=2

module load biokit

rm notmerged/ -rf
mkdir notmerged

for f in $LOCAL_SCRATCH/*_R1*; do

        r=$(sed -e "s/_R1_/_R2_/" <<< "$f")
        i=$(echo $f | awk 'BEGIN { FS = "/" } ; {print $(NF)}' | cut -
d"_" -f1)

$VSEARCH \
--threads $SLURM_NTASKS \
--fastq_filter $f \
--reverse $r \
--fastq_trunclen $trunclen \
--fastqout $LOCAL_SCRATCH/$i"_trimmed_R1.fastq" \
--fastqout_rev $LOCAL_SCRATCH/$i"_trimmed_R2.fastq"

$VSEARCH \
--threads $SLURM_NTASKS \
--fastq_mergepairs $LOCAL_SCRATCH/$i"_trimmed_R1.fastq" \
--reverse $LOCAL_SCRATCH/$i"_trimmed_R2.fastq" \
--fastq_allowmergestagger \
--fastq_maxdiffs $maxdiffs \
--fastq_minmergelen $mergeminlen \
--$QC $maxee \
--relabel barcodelabel=$i\; \
--fastaout_notmerged_fwd notmerged/$i"_notmerged_fwd.fasta" \
--fastaout_notmerged_rev notmerged/$i"_notmerged_rev.fasta" \
--fastaout $LOCAL_SCRATCH/$i".fasta"

cat $LOCAL_SCRATCH/$i".fasta" >> $LOCAL_SCRATCH/superfasta.fasta

rm $LOCAL_SCRATCH/$i".fq" -rf
```

```
rm $LOCAL_SCRATCH/$i".fasta" -rf

done

cutadapt \
--cores=$SLURM_NTASKS \
-a file:primers.fst \
-e $primererr \
--minimum-length $minlen \
--too-short-output tooShort.fa \
--untrimmed-output noPrimers.fa \
-o primertrimmed.fast $LOCAL_SCRATCH/superfasta.fasta

for x in *.fast; do

        g=$(echo $x | awk 'BEGIN { FS = "/" } ; {print $(NF)}' | cut -
d"." -f1)

$VSEARCH --derep_fulllength $x \
    --threads $SLURM_NTASKS \
    --minuniquesize $uniqsize \
    --sizein \
    --sizeout \
    --fasta_width 0 \
    --output $LOCAL_SCRATCH/$g"_uniq.fa"

$USEARCH \
      -unoise3 $LOCAL_SCRATCH/$g"_uniq.fa" \
    -threads $SLURM_NTASKS \
    -minsize 8 \
    -unoise_alpha 2 \
    -zotus zotus.fa

$VSEARCH \
    --usearch_global primertrimmed.fast \
    --id 0.97 \
    --threads $SLURM_NTASKS \
    --db zotus.fa \
    --strand plus \
    --otutabout zotutab_global.txt \
    --biomout biom_zotutab_global.txt

done

# Assign taxonomy
# Bacteria using RDP database
vsearch \
--sintax zotus.fa \
--db
/projappl/project_2000450/refDB/16S/rdp_16s_v16_sp_plus_ZymoMock.udb \
--tabbedout rdp_16s_v16_sp_plus_ZymoMock_out.txt

# Bacteria using SILVA database
vsearch \
--sintax zotus.fa \
--db
/projappl/project_2000450/refDB/16S/silva_16s_v123_plus_ZymoMock.udb \
--tabbedout silva_16s_v123_plus_ZymoMock_out.txt

################################################
```

```
#------------SUMMARY-----------------
Job starts
31Mar2022_0903
Original reads: 18274368
Reads that were successfully merged and quality-filtered: 13166863
Minimum length after primer trimming: 100
Accepted error rate for primers: 20.00%
Primer-trimming and filtering summary:
=== Summary ===

Total reads processed:              13,252,416
Reads with adapters:                13,060,204 (98.5%)
Reads that were too short:             553,895 (4.2%)
Reads written (passing filters):    12,694,914 (95.8%)

Total basepairs processed: 5,626,074,837 bp
Total written (filtered):  5,110,709,874 bp (90.8%)
=== Adapter 341_805 ===

Sequence: CCTACGGGNGGCWGCAG...GGATTAGATACCCBDGTAGTC; Type: linked;
Length: 17+21; 5' trimmed: 13032718 times; 3' trimmed: 13043350 times
Unique non-singleton sequences: 1165438
Number of ZOTUs: 8809
Vsearch global summary Matching unique query sequences: 7897507 of
12694914 (62.21%)
Job efficiency
Job ID: 11170513
Cluster: puhti
User/Group: ekrates/pepr_ekrates
State: RUNNING
Nodes: 1
Cores per node: 4
CPU Utilized: 00:00:00
CPU Efficiency: 0.00% of 09:54:44 core-walltime
Job Wall-clock time: 02:28:41
Memory Utilized: 0.00 MB (estimated maximum)
Memory Efficiency: 0.00% of 62.50 GB (15.62 GB/core)
Job consumed 26.89 CSC billing units based on following used resources
Billed project: project_2000450
CPU BU: 9.91
Mem BU: 15.49
NVME BU: 1.49
WARNING: Efficiency statistics may be misleading for RUNNING jobs.
Time and memory usage:
    ReqMem     MaxRSS     AveRSS    Elapsed  AllocCPUS
---------- ---------- ---------- ---------- ----------
   16000Mc                          02:28:41          4
   16000Mc                          02:28:41          4
   16000Mc                          02:28:41          4
31Mar2022_1203
Job finished

#----------------SUMMARY END---------------

#-------------data to my own computer---------

#-------------- RStudio--------------
```

These are the packages I need for the analysis:

```
library(BiocManager)
```

```
library(microbiome)

library(ggplot2)# graphics

library(readxl)# necessary to import the data from Excel file

library(dplyr)# filter and reformat data frames

library(tibble)# Needed for converting column to row names

library(phyloseq)

library(Biostrings)

library(tidyverse)

library(seqinr)

library(readr)

library(cli)

library(vegan)

library(DESeq2)

library(plyr)

library(knitr)

library(hrbrthemes)

library(gcookbook)

library(ggpubr)

library(knitr)

library(dplyr)

library(DHARMa)
```

In the statistical analysis different online tutorials have been used.

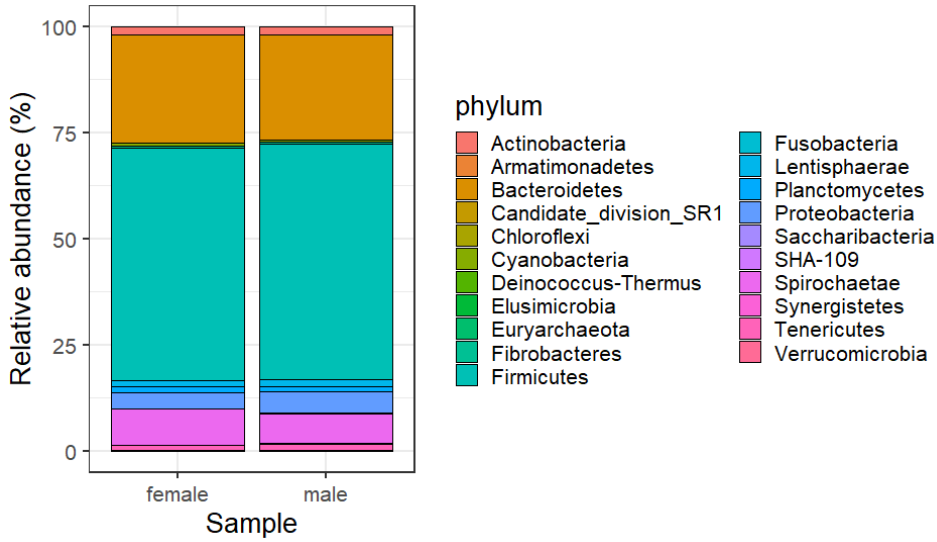
Set working diary. Place where you have all the data files.

```
path<-("Miseq")

#sample data
samdf <- read_excel("Miseq/gradu_aineisto_2022A.xlsx")
samdf <- column_to_rownames(samdf, var = "Sample")
sampledata = sample_data(data.frame(samdf))
physeq <- readRDS(file = "Miseq_final/physeq.rds")
physeq

## phyloseq-class experiment-level object
## otu_table()   OTU Table:         [ 8809 taxa and 94 samples ]
## tax_table()   Taxonomy Table:    [ 8809 taxa by 7 taxonomic ranks ]

library("ape")
random_tree = rtree(ntaxa(physeq), rooted=TRUE, tip.label=taxa_names(p
```

```
hyseq))

physeq1 = merge_phyloseq(physeq, sampledata, random_tree)
physeq1

## phyloseq-class experiment-level object
## otu_table()   OTU Table:         [ 8809 taxa and 94 samples ]
## sample_data() Sample Data:       [ 94 samples by 17 sample variable
s ]
## tax_table()   Taxonomy Table:    [ 8809 taxa by 7 taxonomic ranks ]
## phy_tree()    Phylogenetic Tree: [ 8809 tips and 8808 internal node
s ]

saveRDS(physeq1, file.path(path, "physeq1.rds"))
```

The composition of gut microbiome

```
# Data normalization
pn = transform_sample_counts(physeq1, function(x) 100 * x/sum(x))

# `tax_glom()` function from phyloseq collapses  OTU table at the Phyl
um level. This takes all OTUs under a given phylum and combines them,
adding up the total sequence counts.
phylum = tax_glom(pn, taxrank="phylum")

# Next the average community by sex and not each individual sample:`me
rge_samples()` function to do this.

pm  = merge_samples(phylum, "Sex")

# this next line of code makes sure that everything will be labelled c
orrectly for the  figures.
sample_data(pm)$DeRep <- levels(sample_data(pm)$Sample)

# make a sample_data dataframe for future reference
pm.m = pm %>% sample_data

# Since each 'sample' in this new object is an agglomerate sample, I r
e-normalize the sequence counts for each OTU.
pm = transform_sample_counts(pm, function(x) 100 * x/sum(x))

# `plot_bar()` function
a<-plot_bar(pm, "Sample", fill = "phylum")
a + ylab("Relative abundance (%)") + (theme_bw(base_size = 20))
```

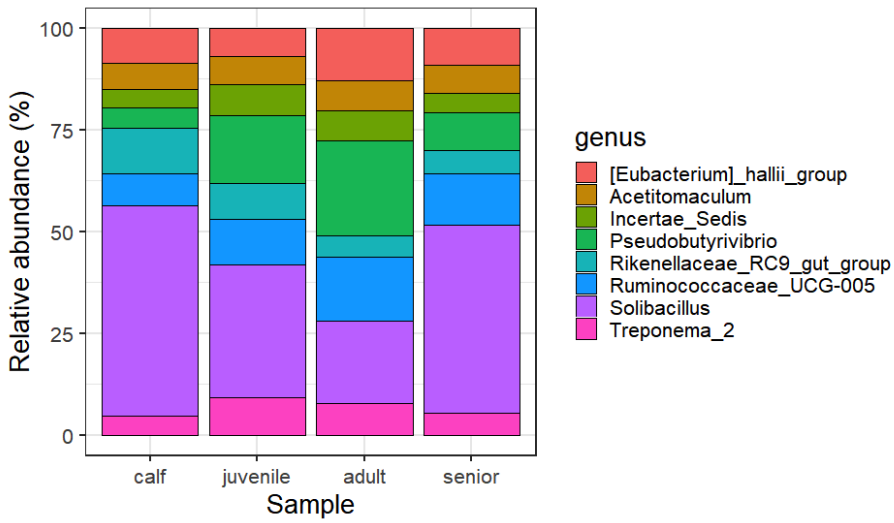



```
#calculate top 20 phylas/class and plot
top20 <- names(sort(taxa_sums(physeq1), decreasing=TRUE))[1:20]
ps.top20 <- transform_sample_counts(physeq1, function(OTU) OTU/sum(OTU
))
ps.top20 <- prune_taxa(top20, ps.top20)
```

Age and top 20 phylas on microbiome

Age and microbiome

```
pn = transform_sample_counts(ps.top20, function(x) 100 * x/sum(x))
phylum = tax_glom(pn, taxrank="genus")
pm  = merge_samples(phylum, "Age")
sample_data(pm)$DeRep <- levels(sample_data(pm)$Sample)
pm.m = pm %>% sample_data
pm = transform_sample_counts(pm, function(x) 100 * x/sum(x))
b<-plot_bar(pm, "Sample", fill = "genus")
b + scale_x_discrete(limits= c("calf", "juvenile", "adult", "senior"))
+ ylab("Relative abundance (%)")+ (theme_bw(base_size = 20))
```



Camp and microbiome

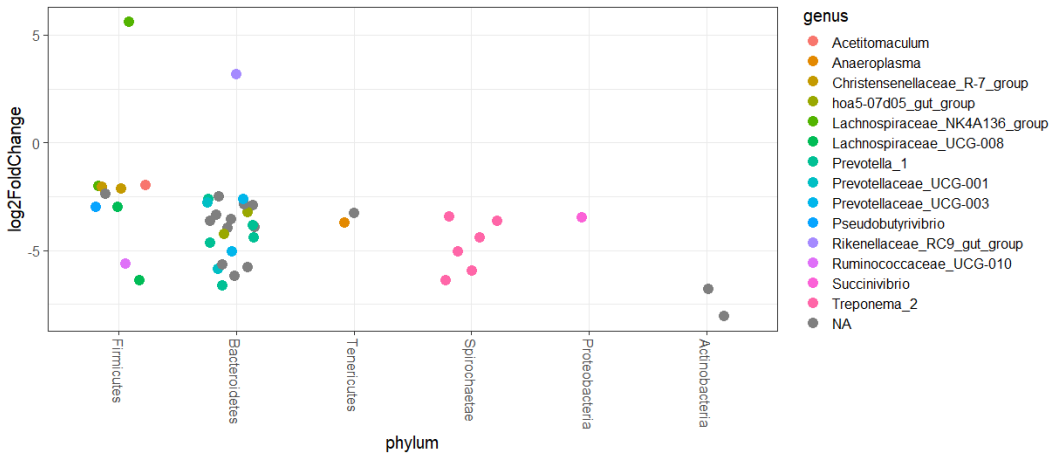
Statistical differences in the microbiome composition between camps

```
physeq_Camp = subset_samples(physeq1, Camp != "Na")

# Create a new `phyloseq` object with only the high and low groups
mb = subset_samples(physeq_Camp, Camp == "West Katha" | Camp == "Kawli
n") %>%
  filter_taxa(function(x) sum(x) > 0, TRUE)

# Next, check the order of our levels. DESeq2 takes the first level as
the 'Control' and the second level as the 'Treatment', and this is nee
ded for downstream interpretation of results.
head(sample_data(mb)$Camp)
```

```
## [1] "Kawlin"     "Kawlin"     "Kawlin"     "West Katha" "Kawlin"
## [6] "Kawlin"
```

```
#"control":Kawlin, "treatment": West Katha
#any sequence that is more abundant in the ‘Control’(Kawlin) variable
will have a negative log2 fold change

# First, convert the sequence count data from phyloseq object into the
proper format for DESeq2
mb.dds <- phyloseq_to_deseq2(mb, ~ Camp)

# estimate the size factors for our sequences
gm_mean = function(x, na.rm=TRUE){
  exp(sum(log(x[x > 0]), na.rm=na.rm) / length(x))
}
geoMeans = apply(counts(mb.dds), 1, gm_mean)
mb.dds = estimateSizeFactors(mb.dds, geoMeans = geoMeans)

# rRun the core DESeq2 algorithm
mb.dds = DESeq(mb.dds, fitType="local")

# Create a new dataframe from the results of our DESeq2 run
mb.res = results(mb.dds)

# Reorder the sequences by their adjusted p-values
mb.res = mb.res[order(mb.res$padj, na.last=NA), ]

# Set alpha for testing significance, and filter out non-significant r
esults
alpha = 0.00000005 #the number needed to be small in order to fit the
genus in the picture
#alpha = 0.01
mb.sigtab = mb.res[(mb.res$padj < alpha), ]

# Add taxonomy information to each sequence
mb.sigtab = cbind(as(mb.sigtab, "data.frame"), as(tax_table(mb)[rownam
es(mb.sigtab), ], "matrix"))
```

```
#OTUs that showed a significant difference in abundance between Camps
#PLOT

# any sequence that is more abundant in the 'Kawlin' variable will hav
e a negative log2 fold change value,
#and any sequence more abundant in the 'West Katha' variable will have
a positive value.

# Set ggplot2 options
theme_set(theme_bw(base_size = 15))
scale_fill_discrete <- function(palname = "Set1", ...) {
  scale_fill_brewer(palette = palname, ...)
}
# Rearrange the order of our OTUs by Phylum.
x = tapply(mb.sigtab$log2FoldChange, mb.sigtab$phylum, function(x) max
(x))
x = sort(x, TRUE)
mb.sigtab$phylum = factor(as.character(mb.sigtab$phylum), levels=names
(x))

# Create and display the plot
p = ggplot(mb.sigtab, aes(x=phylum, y=log2FoldChange, color=genus)) +
geom_jitter(size=4, width=0.25) +
  theme(axis.text.x = element_text(angle = -90, hjust = 0, vjust=0.5),
)

p
```

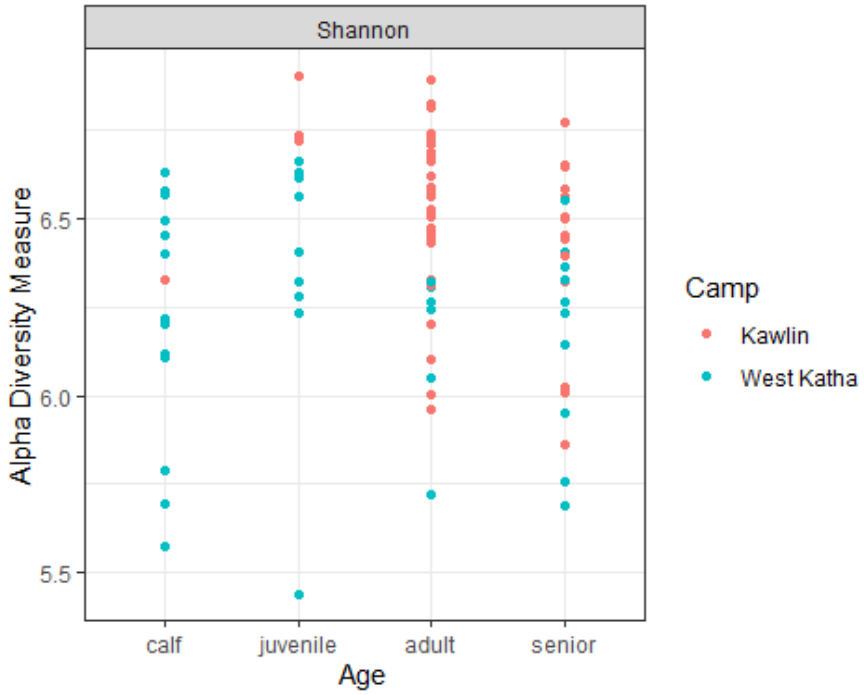


Alpha diversity

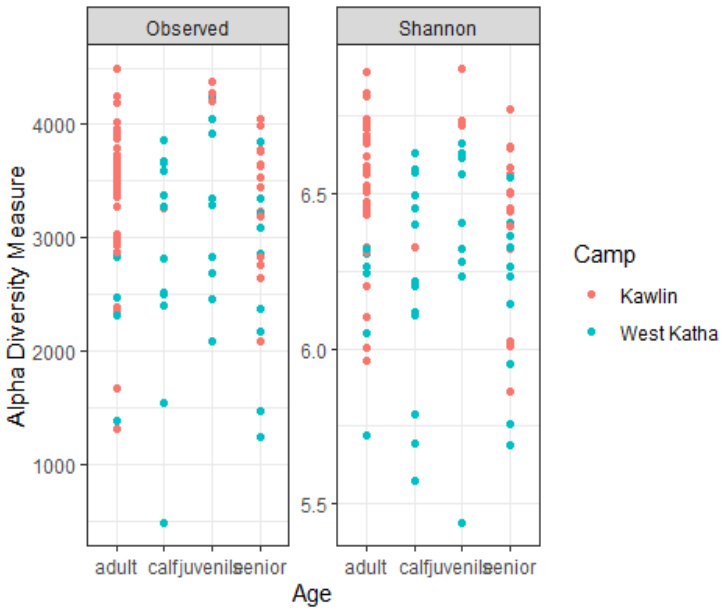
Alpha diversity

```
physeq_camp = subset_samples(physeq1, Camp != "Na")

Shannon<- plot_richness(physeq_camp, x="Age", measures=c("Shannon"), c
olor="Camp") +   theme_bw()
age_ar <- c("calf","juvenile","adult","senior")
```

```
Shannon$data$Age <- as.character(Shannon$data$Age)
Shannon$data$Age <- factor(Shannon$data$Age, levels = age_ar)
print(Shannon)
```




```
#Phyloseq contains the `plot_richness()` function to display multiple alpha diversity measures at once.

plot_richness(physeq_camp, x="Age", measures=c("Observed", "Shannon"), color="Camp") +   theme_bw()
```



Alpha diversity and stress groups

These plots include statistical analysis to support the results.

analysis was made using microbiome R package

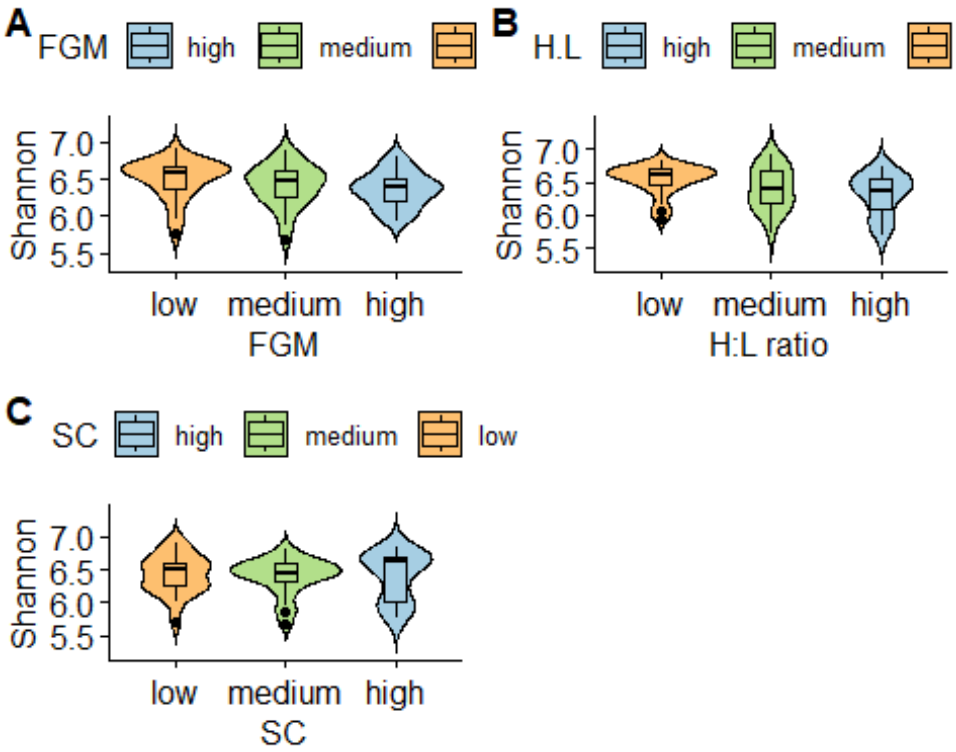
```r
#H:L

physeq_HL = subset_samples(physeq1, H.L != "Na")

ps1_HL <- prune_taxa(taxa_sums(physeq_HL) > 0, physeq_HL)
tab <- microbiome::alpha(ps1_HL, index = "all")
ps1.meta_HL <- meta(ps1_HL)
kable(head(ps1.meta_HL))

ps1.meta_HL$Shannon <- tab$diversity_shannon
ps1.meta_HL$InverseSimpson <- tab$diversity_inverse_simpson

a <- ggviolin(ps1.meta_HL, x = "H.L", y = "Shannon",
              add = "boxplot", fill = "H.L", palette = c("#a6cee3", "#
b2df8a", "#fdbf6f"), xlab = "H:L ratio")
alpha_HL = a+ scale_x_discrete(limits= c("low", "medium", "high"))
alpha_HL
```

```r
#FGM

physeq_FGM = subset_samples(physeq1, FGM_values != "Na")
ps1_FGM <- prune_taxa(taxa_sums(physeq_FGM) > 0, physeq_FGM)
tab <- microbiome::alpha(ps1_FGM, index = "all")
kable(head(tab))
```

```r
ps1.meta_FGM$Shannon <- tab$diversity_shannon
ps1.meta_FGM$InverseSimpson <- tab$diversity_inverse_simpson

a <- ggviolin(ps1.meta_FGM, x = "FGM", y = "Shannon",
              add = "boxplot", fill = "FGM", palette = c("#a6cee3", "#
b2df8a", "#fdbf6f"), xlab = "FGM")
alpha_fgm = a+ scale_x_discrete(limits= c("low", "medium", "high"))
alpha_fgm
```

```r
#SC
physeq_SC = subset_samples(physeq1, SC != "Na")
ps1_SC <- prune_taxa(taxa_sums(physeq_SC) > 0, physeq_SC)
tab <- microbiome::alpha(ps1_SC, index = "all")
kable(head(tab))

ps1.meta_SC <- meta(ps1_SC)
kable(head(ps1.meta_SC))

ps1.meta_SC$Shannon <- tab$diversity_shannon
ps1.meta_SC$InverseSimpson <- tab$diversity_inverse_simpson

a <- ggviolin(ps1.meta_SC, x = "SC", y = "Shannon",
              add = "boxplot", fill = "SC", palette = c("#a6cee3", "#b
2df8a", "#fdbf6f"), xlab = "SC")
alpha_SC = a+ scale_x_discrete(limits= c("low", "medium", "high"))
alpha_SC
```

```
par(mfrow=c(3,1))
library("ggpubr")
figure_alpha <- ggarrange(alpha_fgm, alpha_HL, alpha_SC,
                          labels = c("A", "B", "C"),
                          ncol = 2, nrow = 2)
figure_alpha
```



Statistical analysis to support alpha diversity plots

GLM

Alpha diversity, especially the differences in alpha diversity, was tested using general linear model (GLM). In the model the dependent variable were observed unique sequences representing the alpha diversity. The data were normally distributed, and the residuals were tested for over/underdispersion, outliers and for normal distribution (Kolmogorov–Smirnov test) using DHARMa package.

mod1 and mod2 was compared

ANOVA was used to analyze the GLM results

```
physeq.m_camp = sample_data(physeq_camp)

sample_data(physeq_camp)$Age <- as.factor(sample_data(physeq_camp)$Age
)
adiv <- estimate_richness(physeq_camp, measures = c("Observed", "Shann
on", "Simpson", "Chao1", "InvSimpson"))
```

```r
mod1 <- glm(adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Camp + ph
yseq.m_camp$Sex + physeq.m_camp$cw)
summary(mod1)
```

```
##
## Call:
## glm(formula = adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Camp
+
##     physeq.m_camp$Sex + physeq.m_camp$cw)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.94310  -0.09771   0.04545   0.19727   0.42820
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    6.517828   0.047236 137.986  < 2e-16
***
## physeq.m_camp$Agecalf          0.002411   0.107760   0.022    0.982
## physeq.m_camp$Agejuvenile      0.182231   0.099665   1.828    0.071
.
## physeq.m_camp$Agesenior       -0.069683   0.096391  -0.723    0.472
## physeq.m_camp$CampWest Katha -0.319509   0.074829  -4.270 5.12e-05
***
## physeq.m_camp$Sexmale          0.002523   0.073348   0.034    0.973
## physeq.m_camp$cwwild           0.002148   0.100270   0.021    0.983
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.07716361)
##
##     Null deviance: 8.8258  on 90  degrees of freedom
## Residual deviance: 6.4817  on 84  degrees of freedom
## AIC: 33.837
##
## Number of Fisher Scoring iterations: 2
```

```r
# for Anova
library(car)

# for categorical variables
#install.packages("emmeans")
library(emmeans)


Anova(mod1, type="III", test.statistic="F")
```

```
## Analysis of Deviance Table (Type III tests)
##
## Response: adiv$Shannon
## Error estimate based on Pearson residuals
##
##                     Sum Sq Df F values    Pr(>F)
## physeq.m_camp$Age   0.3855  3   1.6652    0.1807
## physeq.m_camp$Camp 1.4068  1  18.2317 5.119e-05 ***
```

```
## physeq.m_camp$Sex  0.0001  1   0.0012    0.9726
## physeq.m_camp$cw   0.0000  1   0.0005    0.9830
## Residuals          6.4817 84
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
age <- emmeans(mod1, "Age", type="response")
```

```
contrast(age, method = "pairwise", type = "response",adjust="none")
```

```
##  contrast          estimate     SE df t.ratio p.value
##  adult - calf      -0.00241 0.1078 84  -0.022  0.9822
##  adult - juvenile  -0.18223 0.0997 84  -1.828  0.0710
##  adult - senior     0.06968 0.0964 84   0.723  0.4717
##  calf - juvenile   -0.17982 0.1090 84  -1.650  0.1027
##  calf - senior      0.07209 0.1349 84   0.535  0.5944
##  juvenile - senior  0.25191 0.1286 84   1.958  0.0535
##
## Results are averaged over the levels of: Camp, Sex, cw
```

```
camp <- emmeans(mod1, "Camp", type="response")
```

```
contrast(camp, method = "pairwise", type = "response",adjust="none")
```

```
##  contrast            estimate     SE df t.ratio p.value
##  Kawlin - West Katha     0.32 0.0748 84   4.270  0.0001
##
## Results are averaged over the levels of: Age, Sex, cw
```

```
mod2 <- glm(adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex)
summary(mod2)
```

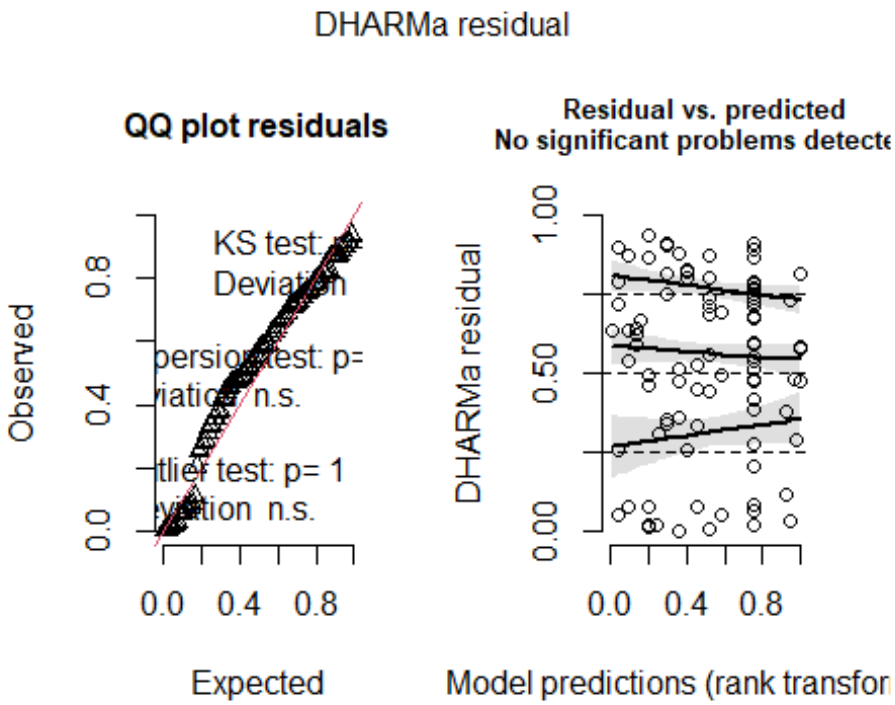
```
##
## Call:
## glm(formula = adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -1.06113  -0.16774   0.08513   0.22284   0.43246
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 6.47717    0.04959 130.624   <2e-16 ***
## physeq.m_camp$Agecalf      -0.22157    0.09877  -2.243   0.0274 *
## physeq.m_camp$Agejuvenile   0.02142    0.09817   0.218   0.8278
## physeq.m_camp$Agesenior    -0.13716    0.07863  -1.744   0.0847 .
## physeq.m_camp$Sexmale      -0.06157    0.07668  -0.803   0.4242
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.09212379)
##
##     Null deviance: 8.8258  on 90  degrees of freedom
## Residual deviance: 7.9226  on 86  degrees of freedom
## AIC: 48.104
```

```
##
## Number of Fisher Scoring iterations: 2
```

```
model<-anova(mod1, mod2)
summary(model)
```

```
##    Resid. Df      Resid. Dev          Df        Deviance
##  Min.   :84.0   Min.   :6.482   Min.   :-2   Min.   :-1.441
##  1st Qu.:84.5   1st Qu.:6.842   1st Qu.:-2   1st Qu.:-1.441
##  Median :85.0   Median :7.202   Median :-2   Median :-1.441
##  Mean   :85.0   Mean   :7.202   Mean   :-2   Mean   :-1.441
##  3rd Qu.:85.5   3rd Qu.:7.562   3rd Qu.:-2   3rd Qu.:-1.441
##  Max.   :86.0   Max.   :7.923   Max.   :-2   Max.   :-1.441
##                                 NA's   :1    NA's   :1
```

```
simulationOutput <- simulateResiduals(fittedModel = mod1)
plot(simulationOutput)
```



Alpha diversity and stress measures (GLM)

Alpha diversity and categorical variables

```
mod6 <- glm(adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex + phy
seq.m_camp$Camp + physeq.m_camp$cw + physeq.m_camp$H.L)
summary(mod6)
```

```
##
## Call:
## glm(formula = adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex
+
##     physeq.m_camp$Camp + physeq.m_camp$cw + physeq.m_camp$H.L)
##
```

```
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.52346  -0.14689   0.04474   0.15056   0.44961
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    6.43683    0.06011 107.075  < 2e-16 *
**
## physeq.m_camp$Agecalf          0.21862    0.15669   1.395 0.168444
## physeq.m_camp$Agejuvenile      0.21188    0.09488   2.233 0.029563 *
## physeq.m_camp$Agesenior       -0.12232    0.10855  -1.127 0.264571
## physeq.m_camp$Sexmale         -0.14486    0.08011  -1.808 0.075929 .
## physeq.m_camp$CampWest Katha -0.28155    0.07630  -3.690 0.000509 *
**
## physeq.m_camp$cwwild           0.07133    0.11802   0.604 0.548056
## physeq.m_camp$H.Llow           0.20398    0.07525   2.711 0.008896 *
*
## physeq.m_camp$H.Lmedium        0.10584    0.07506   1.410 0.164046
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.05641737)
##
##     Null deviance: 5.5044  on 64  degrees of freedom
## Residual deviance: 3.1594  on 56  degrees of freedom
##   (26 observations deleted due to missingness)
## AIC: 7.9011
##
## Number of Fisher Scoring iterations: 2
```

```
mod7 <- glm(adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex + phy
seq.m_camp$Camp + physeq.m_camp$cw + physeq.m_camp$FGM)
summary(mod7)
```

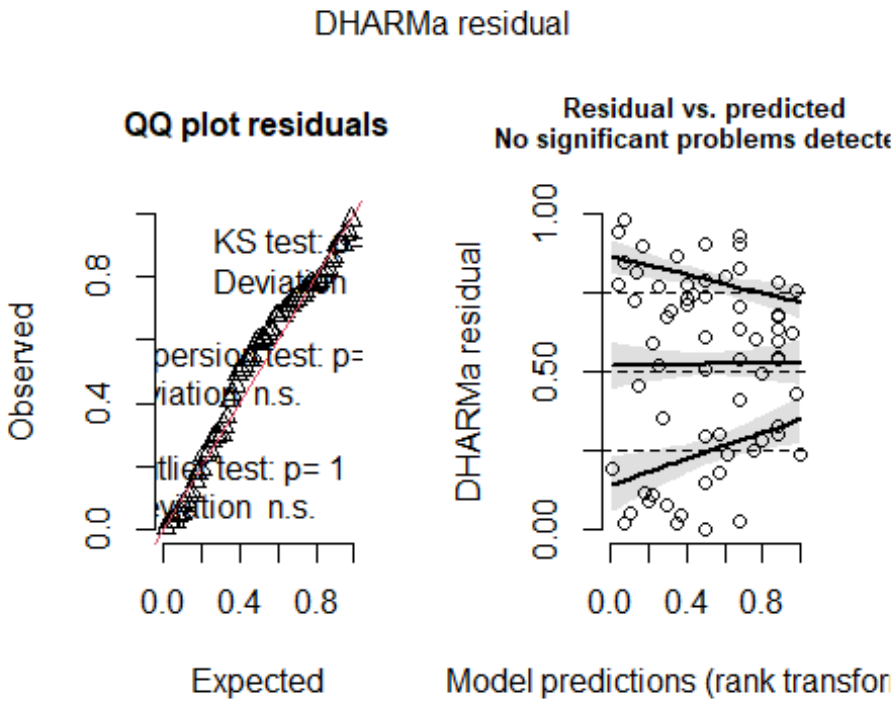
```
##
## Call:
## glm(formula = adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex
+
##     physeq.m_camp$Camp + physeq.m_camp$cw + physeq.m_camp$FGM)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.59286  -0.13430   0.04727   0.16719   0.40182
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    6.55434    0.06677  98.168  < 2e-16 *
**
## physeq.m_camp$Agecalf          0.10945    0.12719   0.861  0.39261
## physeq.m_camp$Agejuvenile      0.21829    0.09516   2.294  0.02499 *
## physeq.m_camp$Agesenior       -0.08595    0.08573  -1.003  0.31973
## physeq.m_camp$Sexmale         -0.05061    0.07901  -0.640  0.52408
## physeq.m_camp$CampWest Katha -0.24078    0.06858  -3.511  0.00081 *
**
## physeq.m_camp$cwwild          -0.03645    0.09465  -0.385  0.70138
```

```
## physeq.m_camp$FGMlow          -0.01420    0.07388  -0.192  0.84818
## physeq.m_camp$FGMmedium       -0.03909    0.07258  -0.538  0.59205
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.05410916)
##
##     Null deviance: 5.1094  on 74  degrees of freedom
## Residual deviance: 3.5712  on 66  degrees of freedom
##   (16 observations deleted due to missingness)
## AIC: 4.4969
##
## Number of Fisher Scoring iterations: 2
```

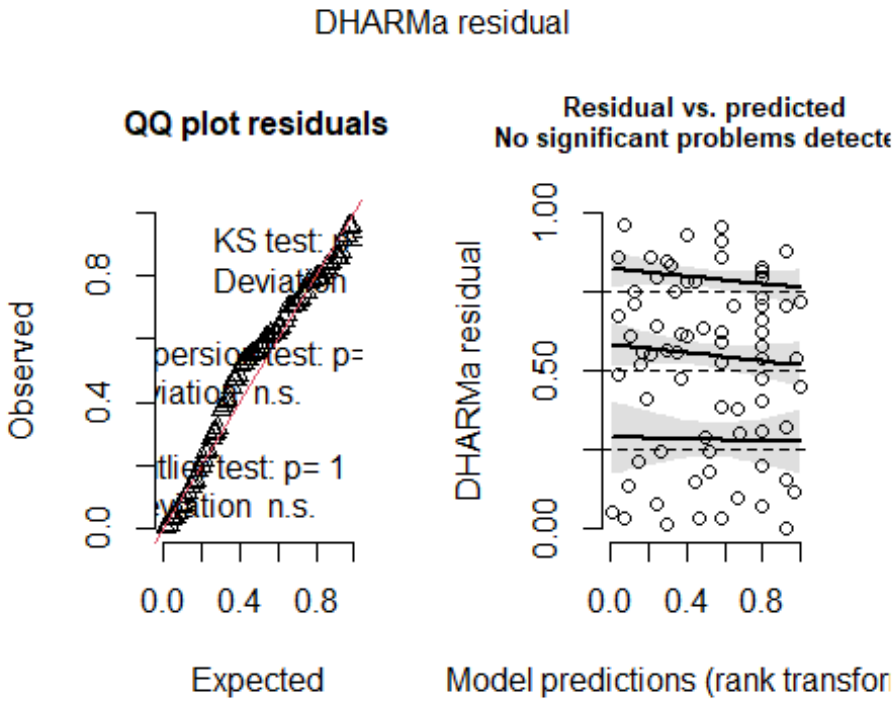
```
mod8 <- glm(adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex + phy
seq.m_camp$Camp + physeq.m_camp$cw + physeq.m_camp$SC)
summary(mod8)
```

```
##
## Call:
## glm(formula = adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex
+
##     physeq.m_camp$Camp + physeq.m_camp$cw + physeq.m_camp$SC)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.64210  -0.13409   0.03667   0.16783   0.43083
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   6.524146   0.069087  94.433  < 2e-16
***
## physeq.m_camp$Agecalf        -0.034742   0.170383  -0.204  0.83922
## physeq.m_camp$Agejuvenile     0.282573   0.110396   2.560  0.01343
*
## physeq.m_camp$Agesenior      -0.215935   0.142985  -1.510  0.13705
## physeq.m_camp$Sexmale        -0.149535   0.087021  -1.718  0.09168
.
## physeq.m_camp$CampWest Katha -0.291298   0.086167  -3.381  0.00138
**
## physeq.m_camp$cwwild          0.169663   0.149946   1.131  0.26304
## physeq.m_camp$SClow           0.008974   0.088208   0.102  0.91936
## physeq.m_camp$SCmedium        0.026599   0.089480   0.297  0.76745
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.06813868)
##
##     Null deviance: 5.5538  on 60  degrees of freedom
## Residual deviance: 3.5432  on 52  degrees of freedom
##   (30 observations deleted due to missingness)
## AIC: 19.514
##
## Number of Fisher Scoring iterations: 2
```

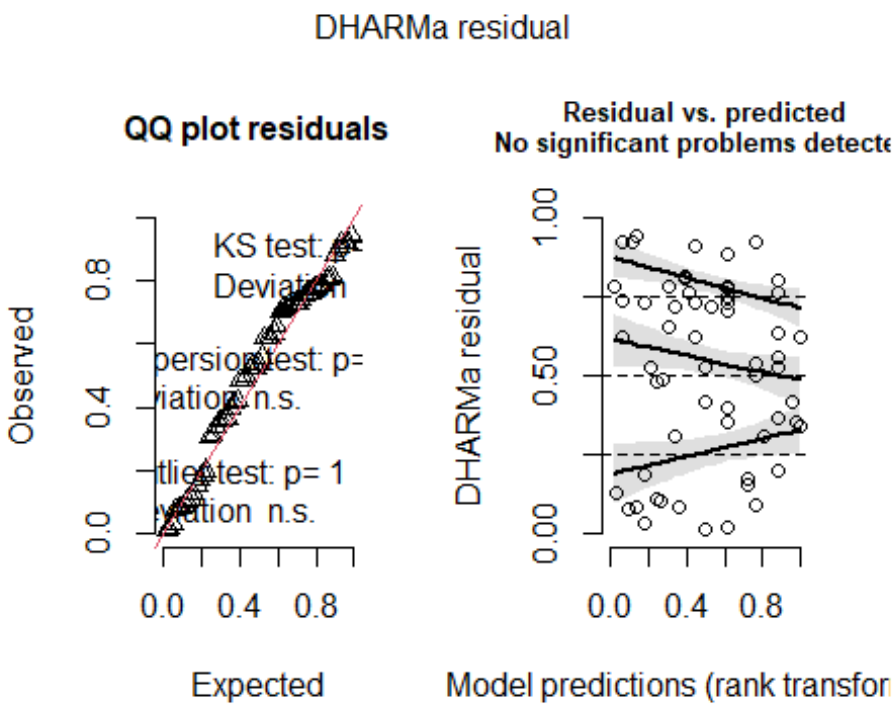
```
simulationOutput <- simulateResiduals(fittedModel = mod6)
plot(simulationOutput)
```




```
simulationOutput <- simulateResiduals(fittedModel = mod7)
plot(simulationOutput)
```




```
simulationOutput <- simulateResiduals(fittedModel = mod8)
plot(simulationOutput)
```



ANOVA for the GLM about stress and alpha diversity

```
#H:L
Anova(mod6, type="III", test.statistic="F")

## Analysis of Deviance Table (Type III tests)
##
## Response: adiv$Shannon
## Error estimate based on Pearson residuals
##
##                      Sum Sq Df F values    Pr(>F)
## physeq.m_camp$Age   0.43675  3   2.5805 0.0625189 .
## physeq.m_camp$Sex   0.18448  1   3.2700 0.0759292 .
## physeq.m_camp$Camp  0.76816  1  13.6156 0.0005093 ***
## physeq.m_camp$cw    0.02061  1   0.3652 0.5480556
## physeq.m_camp$H.L   0.41508  2   3.6787 0.0315457 *
## Residuals           3.15937 56
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

HL_anova <- emmeans(mod6, "H.L", type="response")

contrast(HL_anova, method = "pairwise", type = "response",adjust="none
")

##  contrast      estimate     SE df t.ratio p.value
##  high - low     -0.2040 0.0752 56  -2.711  0.0089
##  high - medium  -0.1058 0.0751 56  -1.410  0.1640
##  low - medium    0.0981 0.0760 56   1.292  0.2018
##
## Results are averaged over the levels of: Age, Sex, Camp, cw
```

```r
#FGM
Anova(mod7, type="III", test.statistic="F")
```

```
## Analysis of Deviance Table (Type III tests)
##
## Response: adiv$Shannon
## Error estimate based on Pearson residuals
##
##                     Sum Sq Df F values    Pr(>F)
## physeq.m_camp$Age   0.3959  3   2.4389 0.0721989 .
## physeq.m_camp$Sex   0.0222  1   0.4102 0.5240818
## physeq.m_camp$Camp  0.6671  1  12.3279 0.0008099 ***
## physeq.m_camp$cw    0.0080  1   0.1483 0.7013807
## physeq.m_camp$FGM   0.0169  2   0.1561 0.8558183
## Residuals           3.5712 66
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
FGM_anova <- emmeans(mod7, "FGM", type="response")

contrast(FGM_anova, method = "pairwise", type = "response",adjust="non
e")
```

```
##  contrast      estimate     SE df t.ratio p.value
##  high - low      0.0142 0.0739 66   0.192  0.8482
##  high - medium   0.0391 0.0726 66   0.538  0.5920
##  low - medium    0.0249 0.0670 66   0.371  0.7115
##
## Results are averaged over the levels of: Age, Sex, Camp, cw
```

```r
#SC
Anova(mod8, type="III", test.statistic="F")
```

```
## Analysis of Deviance Table (Type III tests)
##
## Response: adiv$Shannon
## Error estimate based on Pearson residuals
##
##                     Sum Sq Df F values   Pr(>F)
## physeq.m_camp$Age   0.7672  3   3.7533 0.016270 *
## physeq.m_camp$Sex   0.2012  1   2.9528 0.091676 .
## physeq.m_camp$Camp  0.7787  1  11.4286 0.001379 **
## physeq.m_camp$cw    0.0872  1   1.2803 0.263039
## physeq.m_camp$SC    0.0064  2   0.0471 0.954073
## Residuals           3.5432 52
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
SC_anova <- emmeans(mod8, "SC", type="response")

contrast(SC_anova, method = "pairwise", type = "response",adjust="none
")
```

```
##  contrast      estimate     SE df t.ratio p.value
##  high - low    -0.00897 0.0882 52  -0.102  0.9194
##  high - medium -0.02660 0.0895 52  -0.297  0.7675
##  low - medium  -0.01763 0.0839 52  -0.210  0.8344
```

```
##
## Results are averaged over the levels of: Age, Sex, Camp, cw
```

Alpha diversity and continuous stress levels

```
mod9 <- glm(adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex + phy
seq.m_camp$Camp + physeq.m_camp$cw + physeq.m_camp$H.L_values)
summary(mod9)

##
## Call:
## glm(formula = adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex
+
##     physeq.m_camp$Camp + physeq.m_camp$cw + physeq.m_camp$H.L_value
s)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.58635  -0.16322   0.06956   0.14405   0.42231
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   6.66923    0.07676  86.879  < 2e-16 *
**
## physeq.m_camp$Agecalf         0.19879    0.15755   1.262  0.21218
## physeq.m_camp$Agejuvenile     0.22319    0.09554   2.336  0.02302 *
## physeq.m_camp$Agesenior      -0.15195    0.10744  -1.414  0.16274
## physeq.m_camp$Sexmale        -0.16522    0.08035  -2.056  0.04435 *
## physeq.m_camp$CampWest Katha -0.25977    0.07633  -3.403  0.00122 *
*
## physeq.m_camp$cwwild          0.10279    0.11724   0.877  0.38433
## physeq.m_camp$H.L_values     -0.11284    0.05150  -2.191  0.03253 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.05783755)
##
##     Null deviance: 5.5044  on 64  degrees of freedom
## Residual deviance: 3.2967  on 57  degrees of freedom
##   (26 observations deleted due to missingness)
## AIC: 8.6676
##
## Number of Fisher Scoring iterations: 2

mod10 <- glm(adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex + ph
yseq.m_camp$Camp + physeq.m_camp$cw + physeq.m_camp$FGM_values)
summary(mod10)

##
## Call:
## glm(formula = adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex
+
##     physeq.m_camp$Camp + physeq.m_camp$cw + physeq.m_camp$FGM_value
s)
##
## Deviance Residuals:
```

```
##      Min        1Q    Median        3Q       Max
## -0.59745  -0.13081   0.03796   0.17183   0.38847
##
## Coefficients:
##                                Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    6.464499   0.114908  56.258  < 2e-16
***
## physeq.m_camp$Agecalf          0.110158   0.126068   0.874 0.385348
## physeq.m_camp$Agejuvenile      0.213825   0.093236   2.293 0.024971
*
## physeq.m_camp$Agesenior       -0.092184   0.084511  -1.091 0.279269
## physeq.m_camp$Sexmale         -0.053618   0.078472  -0.683 0.496792
## physeq.m_camp$CampWest Katha -0.247365   0.068099  -3.632 0.000545
***
## physeq.m_camp$cwwild          -0.038971   0.092493  -0.421 0.674855
## physeq.m_camp$FGM_values       0.001181   0.001790   0.660 0.511550
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.05320776)
##
##     Null deviance: 5.1094  on 74  degrees of freedom
## Residual deviance: 3.5649  on 67  degrees of freedom
##   (16 observations deleted due to missingness)
## AIC: 2.3648
##
## Number of Fisher Scoring iterations: 2
```

```
mod11 <- glm(adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex + ph
yseq.m_camp$Camp + physeq.m_camp$cw + physeq.m_camp$cortisol)
summary(mod11)
```

```
##
## Call:
## glm(formula = adiv$Shannon ~ physeq.m_camp$Age + physeq.m_camp$Sex
+
##     physeq.m_camp$Camp + physeq.m_camp$cw + physeq.m_camp$cortisol)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.63069  -0.12015   0.04529   0.17242   0.42703
##
## Coefficients:
##                                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)                    6.534e+00  8.426e-02  77.547  < 2e-16
***
## physeq.m_camp$Agecalf         -4.521e-02  1.736e-01  -0.260  0.79556
## physeq.m_camp$Agejuvenile      2.814e-01  1.086e-01   2.592  0.01231
*
## physeq.m_camp$Agesenior       -2.203e-01  1.411e-01  -1.561  0.12444
## physeq.m_camp$Sexmale         -1.478e-01  8.618e-02  -1.716  0.09208
.
## physeq.m_camp$CampWest Katha -2.876e-01  8.478e-02  -3.392  0.00132
**
## physeq.m_camp$cwwild           1.792e-01  1.458e-01   1.229  0.22432
```

```
## physeq.m_camp$cortisol        1.446e-05  2.053e-03   0.007  0.99441
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.06697398)
##
##     Null deviance: 5.5538  on 60  degrees of freedom
## Residual deviance: 3.5496  on 53  degrees of freedom
##   (30 observations deleted due to missingness)
## AIC: 17.624
##
## Number of Fisher Scoring iterations: 2
```

Phyla that differ by stress groups

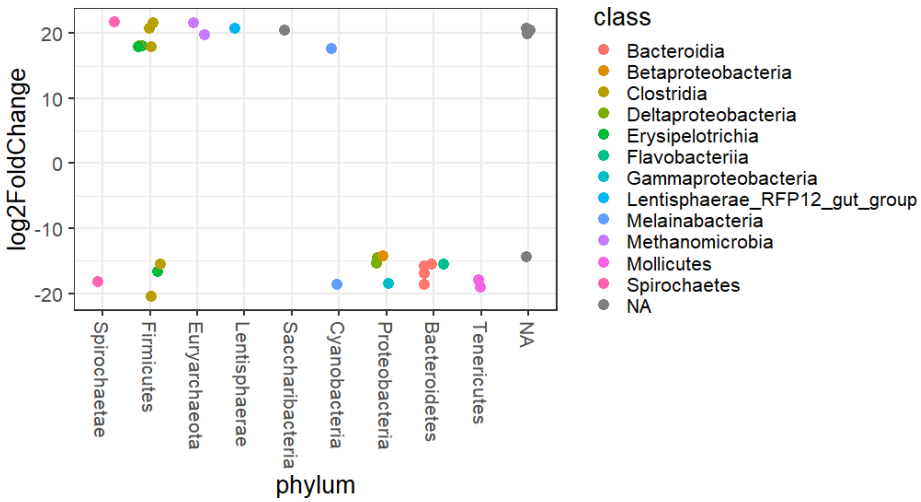
Differential Abundance Testing

H:L ratio

covariates in the analysis Camp and Age

```
physeq_HL = subset_samples(physeq1, H.L_values != "Na")
physeq_HL_Camp = subset_samples(physeq_HL, Camp != "Na")

# Create a new `phyloseq` object with only the high and low groups
mb = subset_samples(physeq_HL_Camp, H.L == "high" | H.L == "low") %>%
  filter_taxa(function(x) sum(x) > 0, TRUE)


# First, convert the sequence count data from our phyloseq object into
the proper format for DESeq2
mb.dds <- phyloseq_to_deseq2(mb, ~ H.L + Camp + Age)

# Next estimate the size factors for the sequences
gm_mean = function(x, na.rm=TRUE){
  exp(sum(log(x[x > 0]), na.rm=na.rm) / length(x))
}
geoMeans = apply(counts(mb.dds), 1, gm_mean)
mb.dds = estimateSizeFactors(mb.dds, geoMeans = geoMeans)

# Finally, run the core DESeq2 algorithm
mb.dds = DESeq(mb.dds, fitType="local")

# Create a new dataframe from the results of our DESeq2 run
mb.res <- results(mb.dds, contrast=c("H.L","high","low")) # by adding
contrast function I can name which variable  I'm interested in

# Reorder the sequences by their adjusted p-values
mb.res = mb.res[order(mb.res$padj, na.last=NA), ]

# Set alpha for testing significance, and filter out non-significant r
esults
alpha = 0.01
mb.sigtab = mb.res[(mb.res$padj < alpha), ]
```

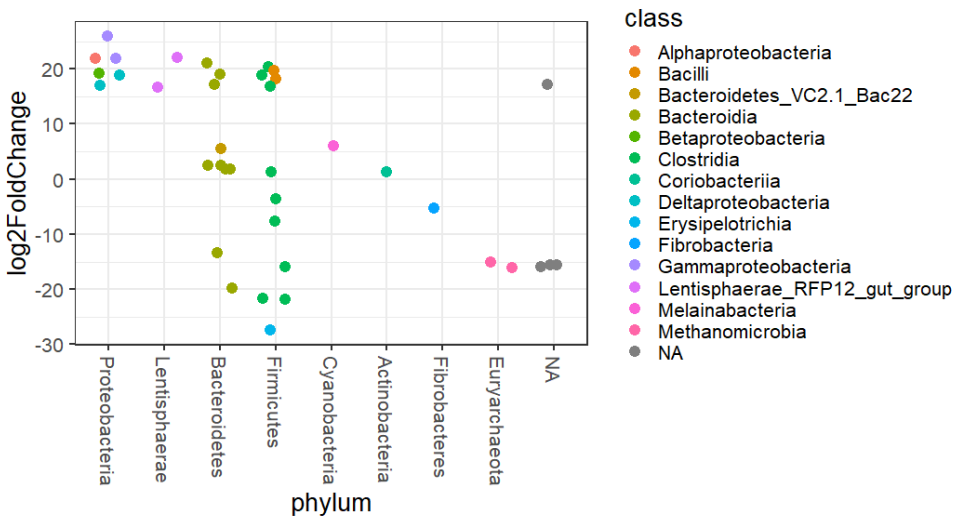
```
# Add taxonomy information to each sequence
mb.sigtab = cbind(as(mb.sigtab, "data.frame"), as(tax_table(mb)[rownam
es(mb.sigtab), ], "matrix"))


# Set ggplot2 options
theme_set(theme_bw(base_size = 20))
scale_fill_discrete <- function(palname = "Set1", ...) {
  scale_fill_brewer(palette = palname, ...)
}
# Rearrange the order of  OTUs by Phylum.
x = tapply(mb.sigtab$log2FoldChange, mb.sigtab$phylum, function(x) max
(x))
x = sort(x, TRUE)
mb.sigtab$phylum = factor(as.character(mb.sigtab$phylum), levels=names
(x))

# Create and display the plot
p_hl = ggplot(mb.sigtab, aes(x=phylum, y=log2FoldChange, color=class))
+ geom_jitter(size=4, width=0.25) +
  theme(axis.text.x = element_text(angle = -90, hjust = 0, vjust=0.5),
)

p_hl+theme(plot.title = element_text(size=18))
```



FGM

```
physeq_FGM = subset_samples(physeq1, FGM_values != "Na")
physeq_FGM_Camp = subset_samples(physeq_FGM, Camp != "Na")
physeq_FGM_Camp

## phyloseq-class experiment-level object
## otu_table()   OTU Table:         [ 8809 taxa and 75 samples ]
## sample_data() Sample Data:       [ 75 samples by 17 sample variable
s ]
## tax_table()   Taxonomy Table:    [ 8809 taxa by 7 taxonomic ranks ]
## phy_tree()    Phylogenetic Tree: [ 8809 tips and 8808 internal node
s ]
```

```
mb = subset_samples(physeq_FGM_Camp, FGM == "high" | FGM == "low") %>%
  filter_taxa(function(x) sum(x) > 0, TRUE)
mb.dds <- phyloseq_to_deseq2(mb, ~ FGM + Camp + Age)
gm_mean = function(x, na.rm=TRUE){
  exp(sum(log(x[x > 0]), na.rm=na.rm) / length(x))
}
geoMeans = apply(counts(mb.dds), 1, gm_mean)
mb.dds = estimateSizeFactors(mb.dds, geoMeans = geoMeans)
mb.dds = DESeq(mb.dds, fitType="local")
mb.res <- results(mb.dds, contrast=c("FGM","high","low"))
mb.res = mb.res[order(mb.res$padj, na.last=NA), ]
alpha = 0.01
mb.sigtab = mb.res[(mb.res$padj < alpha), ]
mb.sigtab = cbind(as(mb.sigtab, "data.frame"), as(tax_table(mb)[rownam
es(mb.sigtab), ], "matrix"))

theme_set(theme_bw(base_size = 20))
scale_fill_discrete <- function(palname = "Set1", ...) {
  scale_fill_brewer(palette = palname, ...)
}

x = tapply(mb.sigtab$log2FoldChange, mb.sigtab$phylum, function(x) max
(x))
x = sort(x, TRUE)
mb.sigtab$phylum = factor(as.character(mb.sigtab$phylum), levels=names
(x))
p_fgm = ggplot(mb.sigtab, aes(x=phylum, y=log2FoldChange, color=class)
) + geom_jitter(size=4, width=0.25) +
  theme(axis.text.x = element_text(angle = -90, hjust = 0, vjust=0.5),
)
p_fgm+theme(plot.title = element_text(size=18))
```



SC

```
physeq_SC = subset_samples(physeq1, cortisol != "Na")
physeq_SC_Camp = subset_samples(physeq_SC, Camp != "Na")
physeq_SC_Camp
```

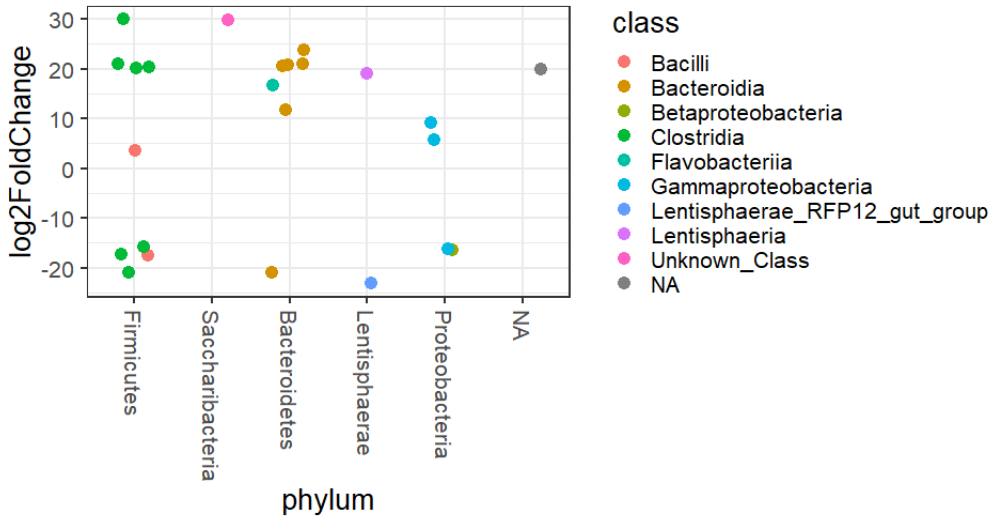
```
## phyloseq-class experiment-level object
## otu_table()   OTU Table:         [ 8809 taxa and 61 samples ]
## sample_data() Sample Data:       [ 61 samples by 17 sample variable
s ]
## tax_table()   Taxonomy Table:    [ 8809 taxa by 7 taxonomic ranks ]
## phy_tree()    Phylogenetic Tree: [ 8809 tips and 8808 internal node
s ]
```

```
mb = subset_samples(physeq_SC_Camp, SC == "high" | SC == "low") %>%
  filter_taxa(function(x) sum(x) > 0, TRUE)
mb.dds <- phyloseq_to_deseq2(mb, ~ SC + Camp + Age)
gm_mean = function(x, na.rm=TRUE){
  exp(sum(log(x[x > 0]), na.rm=na.rm) / length(x))
}
geoMeans = apply(counts(mb.dds), 1, gm_mean)
mb.dds = estimateSizeFactors(mb.dds, geoMeans = geoMeans)
mb.dds = DESeq(mb.dds, fitType="local")
mb.res <- results(mb.dds, contrast=c("SC","high","low"))
mb.res = mb.res[order(mb.res$padj, na.last=NA), ]
alpha = 0.01
mb.sigtab = mb.res[(mb.res$padj < alpha), ]
mb.sigtab = cbind(as(mb.sigtab, "data.frame"), as(tax_table(mb)[rownam
es(mb.sigtab), ], "matrix"))
```

```
theme_set(theme_bw(base_size = 20))
scale_fill_discrete <- function(palname = "Set1", ...) {
  scale_fill_brewer(palette = palname, ...)
}
```

```
x = tapply(mb.sigtab$log2FoldChange, mb.sigtab$phylum, function(x) max
(x))
x = sort(x, TRUE)
mb.sigtab$phylum = factor(as.character(mb.sigtab$phylum), levels=names
(x))
```

```
p_sc = ggplot(mb.sigtab, aes(x=phylum, y=log2FoldChange, color=class))
+ geom_jitter(size=4, width=0.25) +
  theme(axis.text.x = element_text(angle = -90, hjust = 0, vjust=0.5),
)
p_sc+theme(plot.title = element_text(size=18))
```



Beta diversity

Permanova

Community similarity (beta diversity) was tested using PERMANOVA Community similarity was assessed using normalized sequence counts

```
#PERMANOVA:
library(vegan)
# First, normalize the sequence counts by converting from raw abundanc
e to relative abundance. This removes any bias due to total sequence c
ounts per sample.
pn = transform_sample_counts(physeq_camp, function(x) 100 * x/sum(x))
p.df = as(sample_data(pn), "data.frame")
p.d = phyloseq::distance(pn, method = "bray")

p.adonis = adonis2(p.d ~ Camp + Age  + Sex + cw, p.df)
p.adonis

## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = p.d ~ Camp + Age + Sex + cw, data = p.df)
##          Df SumOfSqs      R2      F Pr(>F)
## Camp      1   1.4971 0.08266 8.2147  0.001 ***
## Age       3   0.8908 0.04918 1.6292  0.003 **
## Sex       1   0.2422 0.01337 1.3287  0.092 .
## cw        1   0.1733 0.00957 0.9509  0.496
## Residual 84  15.3087 0.84522
## Total    90  18.1120 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Stress and beta diversity

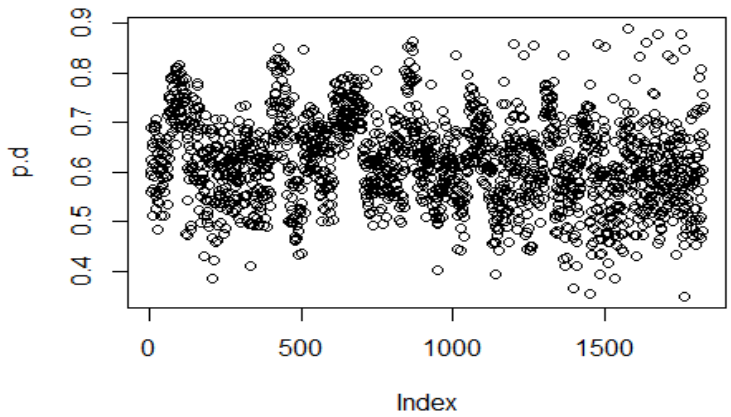
First categorical then continuous stress values

```
#FGM

physeq_FGM_camp = subset_samples(physeq_camp, FGM != "Na")

pn = transform_sample_counts(physeq_FGM_camp, function(x) 100 * x/sum(
x))
p.df = as(sample_data(pn), "data.frame")
p.d = phyloseq::distance(pn, method = "bray")
p.adonis2 = adonis2(p.d ~ Camp + Age  + Sex + cw + FGM, p.df)
p.adonis2

## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = p.d ~ Camp + Age + Sex + cw + FGM, data = p.df)
##          Df SumOfSqs      R2      F Pr(>F)
## Camp      1   1.1856 0.08365 6.8845  0.001 ***
## Age       3   0.8633 0.06091 1.6709  0.002 **
## Sex       1   0.2163 0.01526 1.2558  0.137
## cw        1   0.1977 0.01395 1.1478  0.231
## FGM       2   0.3441 0.02428 0.9991  0.451
## Residual 66  11.3661 0.80195
## Total    74  14.1730 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

summary(p.adonis2)

##        Df           SumOfSqs              R2                F
##  Min.   : 1.00   Min.   : 0.1977   Min.   :0.01395   Min.   :0.9991
##  1st Qu.: 1.00   1st Qu.: 0.2802   1st Qu.:0.01977   1st Qu.:1.1478
##  Median : 2.00   Median : 0.8633   Median :0.06091   Median :1.2558
##  Mean   :21.14   Mean   : 4.0494   Mean   :0.28571   Mean   :2.3916
##  3rd Qu.:34.50   3rd Qu.: 6.2759   3rd Qu.:0.44280   3rd Qu.:1.6709
##  Max.   :74.00   Max.   :14.1730   Max.   :1.00000   Max.   :6.8845
##                                                      NA's   :2
##      Pr(>F)
##  Min.   :0.0010
##  1st Qu.:0.0020
##  Median :0.1370
##  Mean   :0.1644
##  3rd Qu.:0.2310
##  Max.   :0.4510
##  NA's   :2

physeq_FGM_values_camp = subset_samples(physeq_camp, FGM_values != "Na
")

pn = transform_sample_counts(physeq_FGM_values_camp, function(x) 100 *
```

```
x/sum(x))
p.df = as(sample_data(pn), "data.frame")
p.d = phyloseq::distance(pn, method = "bray")
p.adonis3 = adonis2(p.d ~ Camp + Age  + Sex + cw + FGM_values, p.df)
p.adonis3

## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = p.d ~ Camp + Age + Sex + cw + FGM_values, data =
p.df)
##            Df SumOfSqs      R2      F Pr(>F)
## Camp        1   1.1856 0.08365 6.8787  0.001 ***
## Age         3   0.8633 0.06091 1.6695  0.002 **
## Sex         1   0.2163 0.01526 1.2548  0.131
## cw          1   0.1977 0.01395 1.1468  0.226
## FGM_values  1   0.1622 0.01145 0.9412  0.505
## Residual   67  11.5480 0.81479
## Total      74  14.1730 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#H:L

physeq_HL_camp = subset_samples(physeq_camp, H.L != "Na")

pn = transform_sample_counts(physeq_HL_camp, function(x) 100 * x/sum(x
))
p.df = as(sample_data(pn), "data.frame")
p.d = phyloseq::distance(pn, method = "bray")
p.adonis4 = adonis2(p.d ~ Camp + Age  + Sex + cw + H.L, p.df)
p.adonis4

## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = p.d ~ Camp + Age + Sex + cw + H.L, data = p.df)
##          Df SumOfSqs      R2      F Pr(>F)
## Camp      1   0.9683 0.07797 5.5053  0.001 ***
## Age       3   0.7902 0.06363 1.4977  0.002 **
## Sex       1   0.2622 0.02111 1.4907  0.053 .
## cw        1   0.1719 0.01384 0.9771  0.464
## H.L       2   0.3769 0.03035 1.0714  0.281
## Residual 56   9.8492 0.79310
## Total    64  12.4186 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

physeq_HL_values_camp = subset_samples(physeq_camp, H.L_values != "Na"
)
```

```r
pn = transform_sample_counts(physeq_HL_values_camp, function(x) 100 *
x/sum(x))
p.df = as(sample_data(pn), "data.frame")
p.d = phyloseq::distance(pn, method = "bray")
p.adonis5 = adonis2(p.d ~ Camp + Age  + Sex + cw + H.L_values, p.df)
p.adonis5

## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = p.d ~ Camp + Age + Sex + cw + H.L_values, data =
p.df)
##            Df SumOfSqs      R2      F Pr(>F)
## Camp        1   0.9683 0.07797 5.5048  0.001 ***
## Age         3   0.7902 0.06363 1.4976  0.006 **
## Sex         1   0.2622 0.02111 1.4906  0.046 *
## cw          1   0.1719 0.01384 0.9770  0.469
## H.L_values  1   0.2000 0.01611 1.1371  0.232
## Residual   57  10.0260 0.80734
## Total      64  12.4186 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#SC

physeq_SC_camp = subset_samples(physeq_camp, SC != "Na")

pn = transform_sample_counts(physeq_SC_camp, function(x) 100 * x/sum(x
))
p.df = as(sample_data(pn), "data.frame")
p.d = phyloseq::distance(pn, method = "bray")
p.adonis6 = adonis2(p.d ~ Camp + Age  + Sex + cw + SC, p.df)
p.adonis6

## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = p.d ~ Camp + Age + Sex + cw + SC, data = p.df)
##          Df SumOfSqs      R2      F Pr(>F)
## Camp      1   0.9219 0.07806 5.2548  0.001 ***
## Age       3   0.9907 0.08389 1.8824  0.001 ***
## Sex       1   0.2094 0.01773 1.1938  0.165
## cw        1   0.1796 0.01520 1.0235  0.390
## SC        2   0.3859 0.03268 1.0998  0.236
## Residual 52   9.1228 0.77244
## Total    60  11.8103 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

physeq_SC_values_camp = subset_samples(physeq_camp, cortisol != "Na")
```

```
pn = transform_sample_counts(physeq_SC_values_camp, function(x) 100 *
x/sum(x))
p.df = as(sample_data(pn), "data.frame")
p.d = phyloseq::distance(pn, method = "bray")
p.adonis7 = adonis2(p.d ~ Camp + Age  + Sex + cw + cortisol, p.df)
p.adonis7

## Permutation test for adonis under reduced model
## Terms added sequentially (first to last)
## Permutation: free
## Number of permutations: 999
##
## adonis2(formula = p.d ~ Camp + Age + Sex + cw + cortisol, data = p.
df)
##          Df SumOfSqs      R2      F Pr(>F)
## Camp      1   0.9219 0.07806 5.2318  0.001 ***
## Age       3   0.9907 0.08389 1.8742  0.001 ***
## Sex       1   0.2094 0.01773 1.1886  0.179
## cw        1   0.1796 0.01520 1.0190  0.412
## cortisol  1   0.1696 0.01436 0.9628  0.508
## Residual 53   9.3390 0.79075
## Total    60  11.8103 1.00000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#cheking of the residuals
plot(p.d)
```



Stress measures (H:L, FGM, SC) didn't have effect on the beta diversity.

Supplement 5:

Sample	baby	sex	age	birth	cw	age_capture	camp	serum_cortisol	FGM_values	H:L_values	FGM	H:L	SC
2568	N	F	senior	3.9.1959	captive		Kawlin	66,56	76,30464078	1,4375	high	high	high
2630	N	F	senior	6.12.1961	captive		Kawlin		68,58625547	1,541666667	medium	high	
3215	N	F	senior	1.2.1954	wild	17	West Katha	22,49	75,85098039	1,826086957	high	high	medium
3222	N	F	senior	30.11.1964	wild	6	Kawlin	26,57	78,01621359		high		medium
3258	N	F	senior	30.11.1968	wild	2	Kawlin	37,95	98,94046712	0,96969697	high	medium	high
3372	N	F	senior	1.1.1956	wild	16	West Katha	32,86	60,23137255	0,75	low	low	medium
3378	N	F	senior	1.1.1963	wild	9	Kawlin	19,21	66,40156863	1,583333333	medium	high	low
3589	N	F	senior	30.11.1958	wild	14	West Katha		76,52156863	0,634146341	high	low	
3591	N	F	senior	30.11.1965	wild	7	Kawlin	22,96	69,88846154	1,518518519	medium	high	medium
3818	N	F	senior	1.1.1953	wild	20	West Katha	23,73	70,96078431	3,923076923	medium	high	medium
3855	N	F	senior	30.11.1968	wild	5	Kawlin	16,13	58,88654902	0,735294118	low	low	low
3954	N	F	senior	11.6.1969	captive		West Katha	14,84	60,664	0,923076923	medium	medium	low
4024	N	F	senior	21.3.1969	captive		Kawlin	34,68	45,25148515	1,151515152	low	medium	high
4129	N	F	senior	26.11.1969	captive		Kawlin		73,57623762		medium		
4193	N	F	adult	30.11.1970	wild	5	Kawlin		85,42100302		high		
4196	N	F	senior	1.1.1966	wild	10	Kawlin	26,66	60,62698462	0,512820513	medium	low	medium
4365	N	F	adult	14.10.1973	captive		Kawlin	38,55	79,592	0,709677419	high	low	high
4463	N	F	senior	30.11.1967	wild	12	West Katha	19,72	67,25098039	1,321428571	medium	high	low
4616	N	F	adult	29.12.1976	captive		West Katha		58,66	2,117647059	low	high	
4655	N	F	adult	11.2.1976	captive		Kawlin	23,82	65,95535848	1,36	medium	high	medium
4717	N	F	adult	12.5.1977	captive		Kawlin	52,61	67,04145088	0,666666667	medium	low	high
4767	N	F	adult	7.6.1978	captive		Kawlin	29,07	82,5063202		high		medium
4811	N	F	adult	2.1.1978	captive		Kawlin	42,87	45,07658824	1,142857143	low	medium	high
5048	N	F	senior	30.11.1958	wild	25	Kawlin						
5095	N	F	adult	17.6.1996	captive		Kawlin		58,15203307	0,756756757	low	low	
5098	N	F	senior	30.11.1958	wild	25	Kawlin	10,05	73,89932673		high		low
5102	N	F	adult	30.11.1978	wild	7	Kawlin		79,95298		high		
5733	N	F	adult	17.8.1987	captive		Kawlin	35,62	53,81356939	0,8125	low	low	high
5844	N	F	adult	5.3.1988	captive		Kawlin	29,75	90,86778218	1,178571429	high	medium	medium
5955	N	F	adult	3.10.1989	captive		Kawlin	39,16	45,71763952	0,820512821	low	low	high
5962	N	F	adult	24.8.1989	captive		Kawlin	34,48	79,93333333	1,363636364	high	high	high
6020	N	F	adult	15.7.1989	captive		West Katha		63,68		medium		
6077	N	F	adult	23.5.1989	captive		Kawlin	25,09	47,68	0,666666667	low	low	medium
6080	N	F	adult	3.7.1990	captive		Kawlin	22,16	42,076	1,03125	low	medium	medium
6081	N	F	adult	10.9.1990	captive		Kawlin		34,53861386	1,56	low	high	
6084	N	F	adult	24.7.1990	captive		Kawlin		64,54174757	0,727272727	medium	low	
6085	N	F	adult	16.8.1990	captive		Kawlin		73,28543689		medium		
6092	N	F	adult	9.3.1991	captive		Kawlin	31,09	77,72771493	1,185185185	high	medium	medium
6100	N	F	adult	10.10.1991	captive		Kawlin	33,99	41,96470588	1,259259259	low	medium	medium
6101	N	F	adult	24.9.1990	captive		Kawlin	23,17	65,05934936		medium		medium
6196	N	F	adult	24.7.1992	captive			57,01	60,91141026	1,258064516	medium	medium	high
6260	N	F	adult	14.9.1992	captive		West Katha	16,8	83,42135922		high		low
6263	N	F	adult	1.4.1993	captive		Kawlin	21,37	68,53976848	1,133333333	medium	medium	low
6264	N	F	adult	4.6.1993	captive		Kawlin		70,36625397	1,028571429	medium	medium	
6383	N	F	adult	8.9.1993	captive		Kawlin	19,28	56,67529703	0,75	low	low	low
6386	N	F	adult	15.12.1994	captive		Kawlin	80,36	39,17607843	2,181818182	low	high	high
6388	N	F	adult	1.1.1995	captive		Kawlin	47,78	42,21923077	1,533333333	low	high	high
6464	N	F	adult	11.3.1997	captive		Kawlin	40,56	52,56698039	0,75	low	low	high
6465	N	F	adult	2.4.1997	captive		Kawlin	21,8	56,18446602	1,428571429	low	high	low
6520	N	F	adult	13.3.1998	captive		Kawlin		69,52196117	0,892857143	medium	medium	
6521	N	F	adult	1.4.1998	captive		Kawlin	13,21	41,76923077	1,947368421	low	high	low
6525	N	F	adult	26.6.1998	captive		Kawlin	41,66	42,52952381	1,777777778	low	high	high
6617	N	F	juvenile	1.6.2000	captive		Kawlin	72,2	49,62480769	0,681818182	low	low	high
6620	N	F	juvenile	11.11.2000	captive		Kawlin	20,73	52,35839643	0,951219512	low	medium	low
6800	N	F	juvenile	21.2.2003	captive		West Katha	41,29	43,5049505		low		high
6818	N	F	juvenile	10.10.2002	captive		Kawlin			2,647058824		high	
6852	N	F	juvenile	6.5.2004	captive		West Katha	32,85	84,52038835	0,316666667	high	low	medium
6853	N	F	juvenile	24.10.2004	captive		West Katha	11,01	62,96470588	0,925925926	medium	medium	low
6893	N	F	juvenile	24.10.2004	captive		West Katha						
6968	N	F	juvenile	31.12.2006	captive		West Katha	4,424	62,42	0,84375	medium	medium	low
6989	N	F	adult	30.11.1992	wild		West Katha						
7028	N	F	juvenile	10.7.2006	captive		Kawlin	15,67	65,50637624	1,172413793	medium	medium	low
7271	N	F	calf	27.7.2010	captive		West Katha	63,9	46,912	0,47826087	low	low	high
B4705	Y	F	calf	22.2.2015	captive		West Katha						
B4932	Y	F	calf	1.1.2015	captive		West Katha	86,24					high
B5133	Y	F	calf	26.2.2015	captive		West Katha						
B5852	Y	F	calf	3.1.2015	captive		West Katha						
B5948	Y	F	calf	18.1.2015	captive		West Katha						
B6800	Y	F	calf	1.9.2018	captive		West Katha		68,42718447		medium		
2888	N	M	senior	1.1.1963	wild	6	Kawlin	67,57	77,28565327	0,918918919	high	medium	high
3297	N	M	senior	30.11.1960	wild	10	West Katha	30,93	78,28420463	0,970588235	high	medium	medium
3357	N	M	senior	30.11.1964	wild	7	Kawlin	32,82	86,08067588	1,04	high	medium	medium
3486	N	M	senior	30.11.1956	wild	15	West Katha	26,02	82,63366337	0,935483871	high	medium	medium
3805	N	M	senior	30.11.1961	wild	11	West Katha	27,91	99,24660194	0,702702703	high	low	medium
3884	N	M	senior	30.11.1965	wild	8	Kawlin	24,55	101,645098	2,789473684	high	high	medium
4017	N	M	senior	30.11.1967	wild	7		42,13	98,57425743	2,08	high	high	high
4023	N	M	senior	30.11.1968	wild	6	West Katha	50,69	46,18431373	1,35483871	low	high	high
4035	N	M	adult	30.11.1971	captive			10,92	83,02509804	1,066666667	high	medium	low
4254	N	M	adult	30.11.1970	wild	6	Kawlin	20,98	61,89654369	1,5	medium	high	low
4459	N	M	adult	22.4.1974	captive		Kawlin	8,636		1,375		high	low
4921	N	M	adult	1.1.1973	wild	10	West Katha		86,508	0,75	high	low	
5393	N	M	adult	16.2.1984	captive		West Katha	22,02		1,275862069		medium	low
6566	N	M	adult	6.5.1999	captive		Kawlin						
6804	N	M	juvenile	17.2.2004	captive		West Katha	7,4	71,11538462	0,970588235	medium	medium	low
7070	N	M	juvenile	12.10.2007	captive		West Katha	29,48	64,164	0,842105263	medium	low	medium
7104	N	M	juvenile	10.7.2008	captive		West Katha	23,64	69,724	0,538461538	medium	low	medium
7192	N	M	juvenile	21.4.2009	captive		West Katha	13	60,48627451	0,769230769	low	low	low
7314	N	M	calf	28.2.2012	captive		Kawlin	13,41	29,73137255	1,310344828	low	high	low
7394	N	M	calf	22.9.2014	captive		West Katha						
7593	N	M	calf	1.7.2014	captive		West Katha			0,272727273		low	
B4463	Y	M	calf	3.10.2016	captive		West Katha		103,5686275		high		
B4517	Y	M	calf	28.6.2015	captive		West Katha						
B6020	Y	M	calf	1.1.2017	captive		West Katha		101,48		high		
B6627	Y	M	calf	1.3.2016	captive		West Katha						