Elucidating the gut microbiomes of healthy South African individuals from rural and urban localities

by

Khodani Mulaudzi

04901178

Thesis presented in partial fulfilment of the requirements for the degree of

Magister Scientiae

In the Faculty of Natural and Agricultural Sciences

Department of Biochemistry, Genetics and Microbiology

University of Pretoria

Pretoria

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

Supervisor: Dr. Thulani P. Makhalanyane

Co-supervisor: Dr Jarishma K. Gokul

November 2019

Declaration

I, Khodani Mulaudzi, hereby declare that this dissertation submitted for the degree Magister Scientiae (M.Sc.) at the University of Pretoria, is my own work and has not previously been submitted for a degree at this, or other university.

_____

Khodani Mulaudzi

MSc Candidate

November 2019

Acknowledgments

I would like to thank the National Research Foundation and South African Medical Research Council for funding various aspects of the project.

Dr Thulani Peter Makhalanyane, thank you for the opportunity to study under your supervision. I would like to express my sincere gratitude for everything you taught me and all the guidance, insightful comments, support and patience that you gave me throughout past three years.

Dr Jarishma K. Gokul Thank you for your guidance, motivation, and immense knowledge on how to write.

Dr Surendra Vikram thank you for your support, guidance and insight into bioinformatics. I was able to learn and understand R studio and Linux language because of you.

Ms Mubanga H. Kabwe and Mr Percy L. Mutseka thank you both for the endless support, encouragement and motivation.

To my family, you all were my pillar of strength.

Table of Contents

Abstract

The diversity and functional potential of the gut bacterial community of healthy South African individuals are not well understood. This study investigated the gut microbiomes of two populations: individuals living in rural (n = 45) and urban (n = 50) areas using high-throughput 16S rRNA gene sequencing. The bacterial diversity (alpha diversity) was not significantly different between rural and urban populations (Welch test: $p = 0.3$; Shannon) and that the composition differed significantly between urban and rural cohorts (PERMANOVA: $R2 = 0.03$, $p < 0.001$, higher β-diversity). At the phylum level, both populations were dominated by Firmicutes, while Spirochete was uniquely associated with the rural population. Although 24 core bacterial genera were detected across the two distinct populations, the gut bacterial composition revealed signatures that was specific to the geographical location and dietary intake of the individuals. Analysis of the predicted metabolic pathways showed that urban population had carbohydrates, amino acids, lipids, and xenobiotics metabolism. The rural cohort was characterized by species richness and more interindividual homogeneity. In conclusion, the variability of gut bacterial communities within and between populations differed dietary habits based on lifestyle and geographic location.

List of tables

List of figures

List of abbreviations

ANOVA	analysis of variance
BLAST	basic local alignment search tool
BSC	Bristol stool chart
β	beta
DGGE	denaturing gradient gel electrophoresis
DNA	deoxyribonucleic acid
KEEG	Kyoto encyclopaedia of genes and genomes
mM	millimolar
nMDS	non-metric multidimensional scaling
OTUs	operation taxonomic units
PERMANOVA	permutational multivariate analysis of variance
PICRUSt	phylogenetic investigation of communities by reconstruction of unobserved states
rRNA	ribosomal ribonucleic acid
STAMP	statistical analysis of taxonomic and functional profiles
TGGE	temperature gradient gel electrophoresis;
T-RFLP	terminal restriction fragment length polymorphism
µ	micro

Annexures

Chapter 1

Literature review

1.1 Introduction

In the environment, microorganisms are pivotal to ecosystem functionality and help facilitate nutrient cycling, organic matter composition and the sequestration of carbon (Bodelier, 2011). Microorganisms have shaped the transition of the earth from anoxic to oxic, which has led to the transition towards modern earth (Falkowski et al., 2008). Although the roles of microorganisms (i.e bacteria, archaea, fungi and microeukaryotes) in the environment are increasingly clear, key mechanistic insights regarding microbial communities associated with animals, including humans is lacking (Ley et al., 2008).

Recently, several studies suggest that the human microbiome (i.e microorganisms and their genes) is crucial for human health and pivotal in efforts towards defining the one health concept (Clemente et al., 2012; Human Microbiome Project Consortium, 2012). It has been suggested that trillions of microorganisms found on human skin and inside the intestinal tract (Hugon et al., 2015), is now the central focus in efforts to understand the role of the microbiome (Thursby & Juge, 2017). These microbial communities are primarily dominated by bacteria, with about 90% of the species (Tremaroli & Backhed, 2012), and include viruses (Reyes et al., 2012), fungi (Limon et al., 2017), archaea (Hoffmann et al., 2013) in low abundances. These microorganisms perform various functions for the host, such as breaking down and fermenting indigestible food components in the colon (Keim & Martin, 2014), production of vitamins and antibiotics (Nicholson et al., 2012), protection against pathogen invasion (Baumler & Sperandio, 2016), modulating host immunity and maintaining intestinal homeostasis (Kinross et al., 2011). However, there is potential for these functions to be disrupted due to changes in the microbial composition, known as dysbiosis (Carding et al., 2015). The human gastrointestinal tract (GIT) represents one of the primary boundaries between the host and external environmental factors (Thursby & Juge, 2017). Throughout

human life, ingested foods pass through the GIT, along with an abundance of microorganisms from the environment leading to shifts in the resident microbial population and possibly effecting changes to the gut (Thursby & Juge, 2017). The microbial composition and diversity are influenced by many factors including diet (De Filippis & Ercolini, 2018), the use of antibiotics (Blaser, 2016), host genetics (Rodríguez et al., 2015), hygiene practices (Martinez et al., 2015) and other lifestyle factors (Conlon & Bird, 2014). Resultantly, changes in the gut microbiome may lead to the development of diseases such as inflammatory bowel disease, colon cancer, obesity, and diabetes (Nagao-Kitamoto et al., 2016).

In the last few decades, the human gut microbiome has been extensively investigated for its role in physiology and diseases (Ji & Nielsen, 2015; Turnbaugh et al., 2007). With increasingly sophisticated methods (16S ribosomal-RNA (rRNA) gene, or shotgun sequencing) used to profile and characterise the fecal microbiome (Song et al., 2018), these studies have revealed significant inter-individual variability across populations from the same country and even greater variability between those from different countries (European, US and Asia) (Turnbaugh et al., 2007; Consortium, 2010; Huttenhower et al., 2012). In addition, the human gut microbiome is incredibly diverse and complex with the local environmental or external influences playing a significant role in its structure (Qin et al., 2010; Human Microbiome Project, 2012). Previous studies have shown even greater variability in the gut microbiomes of the patients with inflammatory bowel disease (IBD), cardiovascular disease, diabetes, colon cancer, and obesity compared to healthy controls (Lynch & Pedersen, 2016; Schippa & Conte, 2014). In addition to age- and gender-related gut microbial inconsistencies, it is difficult to establish the precise relationship between human health and the presence of specific microbial communities (Carding et al., 2015).

On a global scale, the diversity of the human gut microbiome remains inadequately characterized, with serious knowledge gaps regarding the gut microbiome of non-western populations. Ongoing human gut microbiome projects aimed at profiling, characterizing and creating a reference genome catalog of the gut microbiome have only focused on European,

North American, and Asian populations (Consortium, 2010; The Human Microbiome Project et al., 2012). Larger longitudinal studies that include populations from Africa, the Middle-East, and South America are crucial to ascertain the composition and diversity of the intestinal microbiota of distinct human populations.

There is increasing evidence of significant declines in the number of gut bacterial species and diversity in urban populations compared to those of rural communities (Gomez et al., 2016; Ayeni et al., 2018). For instance, the gut microbial taxa involved in breaking down fiber-rich or resistant-starch food components appear to be among the least abundant taxa in urban populations (De Filippo et al., 2010). The common hypothesis is that observed specific gut microbiome patterns are likely due to differences in lifestyle and dietary habits across the studied populations (Gupta et al., 2017).

In South Africa, the emergence of chronic diseases associated with urbanization and westernised lifestyles, has increased over the last few decades (Steyna, 2006; van der Merwe & Pepper, 2006). However, there is no available data regarding the composition and functional diversity of the gut microbiome of healthy South African individuals.

1.2 The human gut microbiome

The gut microbiome coexists with its host throughout life and plays a significant role in human health (Baumler & Sperandio, 2016). There is a growing realization that the gut microbiome develops with the host and that the composition is altered in response to both internal and external influences (Walsh et al., 2014). Although there are inter-individual and inter-population variations in the gut microbiome, a conserved set of encoded functions and microbes are shared between populations or individuals (often referred to as the ‘core gut microbiome’) (Fujio-Vejar et al., 2017; Qin et al., 2010). In addition to the biochemical and metabolic functions mentioned earlier, the gut microbiome is crucial for retrieving nutrients such as amino acids and short-chain fatty acids (SCFAs) from indigestible food components (i.e resistant-starch), protect the host from pathogenic invasion, and immune system

modulation (Kinross et al., 2011). As a species, we collectively depend on the intestinal microbiota for various essential functions and thus in a mutualistic relationship (Nagao-Kitamoto et al., 2016). However, these functions and mechanisms may potentially be disrupted due to dysbiosis (Barbara et al., 2016).

The dysbiotic state of the gut microbiome has been attributed to various physiological conditions including obesity, IBD, colon cancer, diabetes, and cardiovascular disease (Belizario & Napolitano, 2015). Despite several studies suggesting that alterations in the composition of the gut microbiome may be linked to various diseases, the term “healthy gut microbiome” remains somewhat contentious and poorly defined (Schippa & Conte, 2014). Although the composition of the gut microbiota and its role in health and diseases has recently been subject to intensive studies, our understanding of factors leading to differences in the gut microbiome of healthy individuals remains unclear (Lloyd-Price et al., 2016). Some reports suggest a large degree of inter-individual variations in the human gut microbiomes even without any diseases (Belizario & Napolitano, 2015; Walsh et al., 2014). To determine the physiological significance of different gut microbial composition associated with either a normal or dysbiotic state, it is necessary to define a healthy gut microbiome (Konturek et al., 2015).

Considering the substantial variation between individuals and the fact that several environmental factors may drive its composition, the microbiome of adult humans is relatively stable at the phylum level (Hongfei Cui, 2016). Currently, approximately 1 000 species collectively comprise the gut microbiota of healthy adults (Bik, 2009; Walsh et al., 2014). Among these, the most abundant species belong to members of the phyla *Bacteroidetes* and *Firmicutes*. In almost all individuals, *Bacteroidetes* and *Firmicutes* are present albeit at different abundances (Bull & Plummer, 2014). For example, urban populations generally have higher abundances of *Firmicutes* and less *Bacteroidetes,* while is directly opposite for rural populations where *Bacteroidetes* appear to dominate (De Filippo et al., 2017; Schnorr et al., 2014). Other prevalent but less abundant members are *Proteobacteria, Actinobacteria,*

Verrocumicrobia, Cyanobacteria, and *Fusobacteria* (O'Hara & Shanahan, 2006). Inter-individual variation, at the species level, is higher than at the phylum level (Eckburg et al., 2005). For instance, African children from rural Burkina Faso had the exclusive presence of *Prevotella, Xylanibacter, Butyrivibrio,* and *Treponema* species which, were absent in children from an urban society in Italy (De Filippo et al., 2010). A recent study of individuals from Denmark, Spain, China and the US identified country- or population-specific microbial signatures compositions and concluded that local environmental factors such as diet, and host genetics may play an important role in the composition of the gut microbiome (Li et al., 2014). Due to the differences in the structure of the microbiome across urban and rural gradients, it is unclear whether these patterns may hold for individuals from other parts of the world, including South Africa.

1.3 The role of the human gut microbiome

The gut bacteria colonise the intestinal surfaces, creating stable environments which prevent invasion by external microorganisms (Guarner & Malagelada, 2003). The functions include protective, metabolic, trophic (Purchiaroni et al., 2013) and immunological (Guarner & Malagelada, 2003). Metabolic functions of metabolome include the breakdown of undigested food (often resistant starch and plant-derived foods) components by anaerobic fermentation, which produces short-chain fatty acids (SCFAs) (Tan et al., 2014).

SCFAs includes acetate, propionate, and butyrate which are important by-products. These metabolites play significant roles in maintaining intestinal homeostasis and contribute to host immune system maturation (Pryde et al., 2002; Thorburn et al., 2014). SCFAs have been shown to maintain epithelial integrity, alter phagocytosis and chemotaxis, change cells proliferation and function, and have anti-inflammatory and anti-tumorigenic activities (Barcenilla et al., 2000; Tan et al., 2014). High levels of SCFAs are associated with regular consumption of a fiber-rich diet and higher bacterial diversity in the gut (De Filippo et al., 2010). The substantial health benefits associated with SCFAs linked with the low consumption of

fiber-rich foods, known to produce them, possibly explains the high prevalence of chronic diseases such as inflammatory bowel syndrome and irritable bowel disorder in the urban societies.

1.4 The colonization and establishment of the gut microbiome

Humans are born sterile and acquire microorganisms during birth, which colonize the skin and the gastrointestinal tracts during infancy (O'Hara & Shanahan, 2006; Koenig et al., 2011; Rodríguez et al., 2015; Backhed et al., 2015). Microbial colonization of the gut by newborns is continuous and may originate from several sources, although the mother is the primary source (Penders et al., 2006; Palmer et al., 2007). Newborns are first colonized by facultative anaerobic bacteria, which include *Enterobacteriaceae*, *Streptococcus spp., Staphylococcus spp* and *Enterococcus* (Guaraldi & Salvatori, 2012). The majority of these first colonizers belong to species, which may be pathogenic (Marques et al., 2010). However, these microbes have co-evolved for years with the human host and play important roles such as reducing the levels of oxygen in the neonate's GIT (Marques et al., 2010). Thus, oxygen-free environments will promote the proliferation of obligate anaerobes dominated by *Bifidobacteria*, *Bacteroides,* and *Clostridium* (Guaraldi & Salvatori, 2012). The mode of birth (caesarean or vaginal) (Dominguez-Bello et al., 2010), type of nourishment (breast-milk, and formula-milk) (Voreades et al., 2014), administration of antibiotics (either to the mother during pregnancy or infant after birth), and illness (Koenig et al., 2011; Stearns et al., 2017) influence the structure of the infant’s gut bacterial community. Consequently, these results in increased variation in the gut microbiota of individuals and may have a long-term influence on the associated health outcomes of the host (Marques et al., 2010).

The pattern in which intestinal bacteria colonise newborns delivered vaginally differs from that born through cesarean method (Dominguez-Bello et al., 2010; Backhed et al., 2015). During a vaginal delivery, the vaginal and fecal microbiota become major sources of inoculum (Koenig

et al., 2011). Several studies show that the fecal microbiota (termed meconium) of infants delivered vaginally is dominated by *Prevotella*, *Parabacteroides*, *Bacteroides,* and *Lactobacillus* and, after a few months, there are high occurrences of *Bifidobacterium* and *Bacteroides* (Dominguez-Bello et al., 2011; Koenig et al., 2011; Yang et al., 2016). In contrast, infants delivered through cesareans have fecal microbiota profiles like the mother's skin microbiota (Palmer et al., 2007). Their gut microbiota is dominated by relatively high abundance of *Staphylococcus* and *Propionibacterium* spp. and *Corynebacteria*, with low proportions of *Bacteroides spp.* and *Bifidobacteria* (Backhed et al., 2015; Dominguez-Bello et al., 2010; Ravel et al., 2011). The differences between the two delivery modes steadily decrease over a 4 to 12-month period, with the composition of infants delivered through caesareans remaining more heterogeneous (Backhed et al., 2015). Although the presence of other bacterial communities such as *Bifidobacterium* and *Lactobacillus* are known to promote host health (Rastall, 2004), the physiological impact (short- and long-term effects) by other initial bacterial colonizers is less clear.

The growth and development of the gut microbiota continues during the breast and/or formula feeding period (Penders et al., 2006). Breastfeeding is a continuous source of potential probiotic bacteria, which are known to promote initial gut colonization and growth possibly influencing long-term health (Guaraldi & Salvatori, 2012; Jost et al., 2013). This is due to the presence of oligosaccharides in breast milk, which promote the growth of *Bifidobacterium* and *Lactobacillus* and further strengthen the development of the immune system (Conlon & Bird, 2014). Several studies have demonstrated that the intestinal colonization of formula-fed infants follows a different trajectory to that of breastfed infants. Fan et al., (2013) investigated 12 healthy Chinese infants (aged 3 – 6 months) and showed that although the fecal microbiota of both breastfed and formula-fed may be dominated by similar phyla (*Actinobacteria, Firmicutes,* and *Proteobacteria*) the relative abundance of *Proteobacteria* in breastfed subjects and *Firmicutes* in formula-fed subjects differed significantly (Fan et al., 2013). These findings

provide information on the first diet-related intestinal colonization and the importance of diet in shaping the structure of the human gut microbiome (Fan et al., 2013).

The introduction of solid food in the infant's diet, generates a shift towards bacterial consortia like that of adults (Guaraldi & Salvatori, 2012). A significant decline in *Lactobacillus* and *Bifidobacterium* abundance is observed when milk consumption is reduced, and replaced by solid foods (Koenig et al., 2011; Voreades et al., 2014; Backhed et al., 2015). At approximately 2 – 3 years old, the bacterial community structure begins to stabilize and is dominated by *Bacteroidetes* and *Firmicutes* (Koenig et al., 2011; Voreades et al., 2014).

1.5 Factors influencing the gut microbiome composition

1.5.1 Diet

Of all the factors which likely influence the composition and structure of microbes in the human gut, diet remains the most studied due to its significant role in modulating microbiota either beneficially or detrimentally (Conlon & Bird, 2014). Both long- and short-term dietary factors have been shown to induce changes in the composition and functional potential of the intestinal microbiota (David et al., 2014; Scott et al., 2013; Wu et al., 2011). Other studies have also shown that the gut microbiome may be resilient to change. For instance, the gut microbiome of mice was shown to revert to the original composition following changes in diet (Zhang et al., 2012). However, the effects of both long- and short-term dietary patterns on the gut microbiome are relatively unexplored and may provide clearer cues regarding the effects of the microbiome on human health.

Carbohydrates, proteins, and fats are major components in the diet of humans, albeit with differing consumption proportions depending on geography (Benítez-páez et al., 2016, Gupta et al., 2017). The type and amount of these components present in diets influences bacterial species and host metabolism (Rajoka et al., 2017). A high-fat and high-sugar diet increases the ratio of *Firmicutes* to *Bacteroidetes*, effects changes on the metabolic pathways in the

microbiome, and alters microbiome-associated gene expression (Turnbaugh et al., 2009). In a separate study, a high-fat diet was correlated to increased lipopolysaccharides (LPS) circulation and intestinal permeability (Cândido et al., 2018). Intestinal permeability mediates obesity and systemic inflammation (Moreira et al., 2012). Fat stimulates or increases the production of bile acids by gut microbiota (O'Keefe, 2008), and has been linked to increased colon cancer risk (Ou et al., 2013). Some of the bile acids from the gallbladder escape into the small intestine and enter the colon through the hepatic system, and this process generates a feedback mechanism increasing dietary fat intake (O'Keefe, 2008). Within the large intestine, colonic microbiota produce secondary bile acids, using 7 α-dehydroxylation from primary bile acids (bile acid from the gallbladder) (Ou et al., 2012). However, the type and amount of bile acids which escape to the colon remain unclear (Conlon & Bird, 2014). Furthermore, how these mechanisms may be influenced by the gut microbiota in different individuals remains unclear.

Degradation of proteins normally occurs in the colon where proteolytic bacteria are present (Macfarlane et al., 1986). A high protein diet increased proportions of *Bifidobacteria* and *Lactobacilli* in mice (Rajoka et al., 2017). Most ingested carbohydrates are fermented by colonic microbiota (Conlon & Bird, 2014). Fermentation of carbohydrates, such as fiber-rich components, increases the bacterial biomass in the colon (Simpson & Campbell, 2015). This leads to the production of metabolites such as SCFAs, major end-product of biodegraded carbohydrates (West et al., 2013). Clarke and colleagues revealed that *Parabacteroides distasonis* counts increases and was responsible for approximately 90% of SCFAs produced in individuals with high carbohydrates intake, particularly high-amylose maize starch (Clarke et al., 2011).

Current knowledge about the influence of diet on the diversity and functional potential of intestinal bacteria is based mainly on Caucasian populations, living in western societies and consuming western diet (i.e. high in sugar and saturated fats). Very few studies have examined the gut microbiome of traditional farmers/agriculturists (Morton et al., 2015; Gomez

et al., 2016), hunter-gatherers (Schnorr et al., 2014; Obregon-Tito et al., 2015), vegetarians and omnivores (Kabeerdoss et al., 2012; Kim et al., 2013; Matijašić et al., 2014). Comparative studies consistently found major differences among the intestinal microbiomes of urban populations consuming a western diet, and rural cohort with a low-fat diet (De Filippo et al., 2010; Schnorr et al., 2014; Martinez et al., 2015; Gomez et al., 2016). Studies by Obregon-Tito et al. (2015) and Schnorr et al. (2014) reported a significant decline in the gut microbial diversity of urban or western populations. These studies attributed this decline in diversity to dietary habits and showed the absence and reduction of *Treponema* and *Prevotella* in urban populations. Developing countries, such as South Africa, are subject to increased rates of urbanization with increasing populations adopting western lifestyles (Mutavhadsindi & Meiring, 2014). This may affect the gut microbiota and functional processes in both rural and urban populations. However, there are currently no data available regarding the gut microbiome of South African individuals and the impact of dietary differences remains undetermined.

1.5.2 Lifestyle

Although a few studies have provided insights regarding the effects of lifestyle on the microbiome, the influence of non-dietary factors including cigarette smoking, stress, lack of physical exercise on the gut microbiota is less clear. Cigarette smoking, and the lack of exercise may alter the intestinal microbiota, which may increase non-communicable diseases such as colorectal cancer (CRC) (Conlon & Bird, 2014). One study reported a significant increase of *Bacteroides* in cigarettes-smoking individuals compared to non-smoking participants (Benjamin et al., 2012). The influence of other non-dietary factors such as stress (Mackos et al., 2016), and a lack of exercise (Booth et al., 2012; Campbell & Wisniewski, 2017) on the human gut microbiome have been reported. Stress-induced reductions in *Lactobacilli* have been reported in the gut of stressor-exposed animals, such as mice (Tannock & Savage, 1974; Bailey & Coe, 1999; Sakuma et al., 2013).

1.5.3 Host genetics

Host genetics may influence the composition of the gut microbiota and metabolic pathways (Turpin et al., 2016; Goodrich et al., 2016; Xie et al., 2016). To reduce other confounding variables, several studies have assessed the fecal microbiota of monozygotic twin pairs and unrelated individuals (Zoetendal et al., 2001; Turnbaugh et al., 2009). Results revealed that the gut microbiomes of monozygotic twins were structurally more similar compared to unrelated individuals. Turnbaugh and colleagues further showed that the fecal microbiota was more similar between a mother and a daughter compared to unrelated individuals (Turnbaugh et al., 2009). Taken together, these studies suggest that the gut microbiota may be determined by host genetics to a significant degree.

1.5.4 Antibiotic

Evidence of both short and long-term use of antibiotics (for infection treatment) result in distinct temporal perturbation in the human gut microbiota (Jakobsson et al., 2010; Löfmark et al., 2006). It has been demonstrated that the composition of the gut microbiota may be disrupted for prolonged periods depending on the type of antibiotics (Jernberg et al., 2005; Jernberg et al., 2007). Jernberg et al. showed that within two years, *Bacteroides* species did not revert to the original count, after 7 days intake of clindamycin (Jernberg et al., 2007). A study by Jakobsson and colleagues assessed the short- and long-term effects of metronidazole and clarithromycin, a standard treatment against *Helicobacter pylori* in patients with gastric ulcer (Jakobsson et al., 2010). The authors observed a significant decline in members of *Actinobacteria* immediately after treatment, in both throat and fecal microbiota of patients. However, a study using animal models (murine) suggested that, although the gut microbiome can be disrupted by antimicrobial agents, it could steadily revert to a relatively new and stable composition (Cho et al., 2012). More studies are needed to determine antibiotic-based impact on gut bacterial population over extended periods.

1.5.5 Geography

A study by Suzuki and Worobey reported that there are variations in the gut microbiome of individuals from colder regions compared to those living in warmer regions. This is ascribed to the fact that individuals from colder regions potentially consumed, extracted and stored more energy from dietary fats (Suzuki & Worobey, 2014). In a separate study, the functional profiles and fecal microbiota differed significantly between individuals living in the US and those living in rural areas of Malawi and Venezuela (Yatsunenko et al., 2012). This study demonstrates that geography has a significant influence on the composition of the gut microbiome (Gupta et al., 2017). However, knowledge about variations in the gut microbiome of individuals from the same country and ethnic background, but different local geography (rural and urban) is limited.

1.5.6 Age and gender

Due to altered physiology in elderly individuals, decreased intestinal motility and secretion of gastric acids, and changes in dietary and lifestyle habits, the gut microbiome of the elderly differs markedly from that of younger adults (Biagi et al., 2010). For instance, most elderly individuals have relatively low proportions of *Bacteroides* and *Bifidobacteria* species (Biagi et al., 2010), reduced amylolysis processes, and low levels of SCFAs (Rampelli, et al., 2013). Whilst, *Fusobacteria*, *Clostridia*, *Eubacteria,* and facultative anaerobes are higher in abundance (Woodmansey, 2007). Studies on intestinal microbiota with large cohorts of Europeans revealed distinct age- and gender-related variations. These studies showed that the ration of *Bacteroides-Prevotella* was higher in males than in females (Mueller et al., 2006).

1.6 The gut microbiome in diseases

The significance of the intestinal microbiota in human health is clear (D'Argenio & Salvatore, 2015). A substantial body of research suggests that the intestinal microbiota may play a key role in the manifestation of various metabolic and GIT-related diseases (Cho & Blaser, 2012; Belizario & Napolitano, 2015; Nagao-Kitamoto et al., 2016). The evidence suggests that

dysbiosis of the gut bacterial community may cause various metabolic diseases (Carding et al., 2015). These include inflammatory bowel disease (Wu et al., 2013), obesity (Walters et al., 2014), colon cancer (Hagland & Søreide, 2015), type 1 and type 2 diabetes mellitus (Bertoni et al., 2010; Barlow, Yu, & Mathur, 2015), atherosclerosis (Koeth et al., 2013), and asthma (West et al., 2015), prevalent in western or urban societies.

It has been proposed that the development of obesity in children (Karlsson et al., 2012) and adults (Turnbaugh & Gordon, 2009) is influenced by the relative proportions of *Bacteroidetes* and *Firmicutes*. However, the role of intestinal bacteria on the disease appears to be inconsistent from one study to the next. Schwiertz and colleagues observed very low *Firmicutes* to *Bacteroidetes* ratios in obese individuals compared to lean controls (Schwiertz et al., 2010). In contrast, Duncan et al., (2008) reported no differences in *Bacteroidetes* counts between obese and non-obese controls nor was there evidence that *Bacteroidetes* and *Firmicutes* contributed to obesity (Duncan et al., 2008).

1.7 Current methods for studying the human gut microbiome

For years, human gastrointestinal tract microorganisms were studied using traditional cultivation-based methods or techniques (McDonald et al. 2015). The advantage of using cultivation techniques is that the isolates can be recovered and used for other downstream analysis (Dicksved, 2008; Browne et al. 2016). However, these approaches require a large amount of work in relation to the output, and only about 10-25% of the GIT microorganisms have been successfully cultured (Eckburg et al., 2005). Culture-independent methods, which analyse DNA extracted directly from a sample, have expedited studies on the gut microbiome. These include taxonomic diversity, how many of which microbes are present in a community (Franzosa et al., 2015) and functional metagenomics which describes biological functions of an entire community or members of a community (Morgan and Huttenhower 2012). To investigate bacterial communities efficiently, studies are combining high-throughput DNA

sequencing with other genome-scale platforms such as proteomics and metabolomics (Vernocchi et al., 2016; Maier et al., 2017).

Briefly, metagenomic studies on the human gut microbiota comprise several experimental steps, namely collection of fecal samples, metagenomic DNA extraction, massive DNA sequencing, and data analyses with bioinformatics (Song et al., 2018). Several molecular techniques including culture-independent methods have been developed to elucidating the gut microbiota (**Table 1.1**). Each technique has different applications to the study of the gut microbiota. Many researchers have employed sequence analysis of the 16S rRNA marker gene (Hold et al., 2002; Human Microbiome Project Consortium, 2012; Martínez et al., 2015; Carbonetto et al., 2016; Fujio-Vejar et al., 2017; Hansen et al., 2019). Using 16S rRNA gene is relatively cheap and simple to sequence only the 16S sequences from a microbiome (Caporaso et al., 2010). This 1.5 kbp gene is sufficiently conserved to allow amplification and the ubiquitous sequences and regions vary with greater or lesser frequency over evolutionary time (Jovel et al., 2016). Bacterial 16S rRNA genes has 9 hypervariable regions (V1–V9) that show sequence diversity and uses a barcode-like method to differentiate bacterial taxa, and sometimes but not always, species level resolution is possible (Allaband et al., 2018). There are many considerations in choosing which primers to use (Soergel et al., 2012). However, the best option is to use the same PCR primers as other studies which would be easy to compare the results, such as V1 - V3 or V3 - V5 primers previously used for the Human Microbiome Project (Human Microbiome Project, 2012). The 16S amplicons are matched against several databases including GreenGenes (DeSantis et al., 2006), the Ribosomal Database Project (Cole et al., 2009), and Silva (Pruesse et al., 2007). Sequences assayed in this manner have been characterized for close association of the gut microbiota with human health and diseases (Vesterbacka et al., 2017, Mullish et al., 2018). Although more recently this approach has been overshadowed by metagenomics (i.e shotgun sequencing), 16S rRNA gene sequencing is still robust and well characterized approach that yields sufficient

information about bacterial communities, even with relatively small number of sequences per samples (below 200 thousand) (Laudadio et al., 2018).

More recent researchers employ whole genome shotgun sequencing (WGS) for human gut microbiome (Ranjan et al., 2018; Rampelli et al., 2019). Whole genome shotgun sequencing is rapidly displacing 16S rRNA amplicon analysis because it can provide genome content, functional potential and readily resolve species-level and strain-level classification (Allaband et al., 2018). This method bypasses gene-specific amplification and randomly sequences all fragmented DNA in a community including other microorganisms such as eukaryotes and viruses (Clooney et al., 2016). An important consideration is that WGS and 16S methods use different databases for classification of taxa (Laudadio et al., 2018). However, the amount of obtained WGS sequence data leads to few challenges in terms of data processing, analysis and storage. Sequencing platforms for WGS such as Illumina HiSeq 2500 can yield over 1 Tbp of raw sequence data which increases several-fold during downstream processing and analysis (Clooney et al., 2016). It also requires higher coverage (10–30 million of reads), to identify and understand the bacterial genes (Ranjan et al., 2016). Overall, WGS remains more expensive than 16S sequencing, less tolerant of low microbial biomass or contaminated samples, and analytic approaches are complex and computationally expensive (Allaband et al., 2018).

Other prevalent non-sequencing culture-independent methoods for molecular profiling of the intestinal microbiome include metabolomic and metaproteomic techniques that use mass spectrometry (MS) (Melnik et al., 2017) and nuclear magnetic resonance (NMR) spectrometry (Jacobs et al., 2008; Cox et al., 2019). The MS and NMR profiling techniques has been used to identify bacterial products and metabolites, such as fatty acids, vitamins, bile salts, and polyphenols (Vernocchi et al., 2016; Cox et al., 2019).

Table 1.1: Techniques used to characterize the gut microbiota.

Technique	Description	Advantages	Limitations	References
Culture-based	Isolation of bacteria on selective media	Cheap, semi-quantitative	Labour intensive, culture <30% of gut microbiota	Guarner, F. & Malagelada, 2003; Ingham et al., 2007
DGGE/TGGE	Gel separation of 16S rRNA amplicons using denaturant/temperature	Fast, semi-quantitative, bands can be excised for further analysis	No phylogenetic identification, PCR bias	Muyzer, 1999
T-RFLP Fluorescently	labelled primers are amplified and then restriction enzymes are used to digest the 16S rRNA amplicon. Digested fragments separated by gel electrophoresis	Fast, semi-quantitative, cheap	No phylogenetic identification, PCR bias, low resolution	Marsh, 1999; Hayashi et al., 2003
DNA microarray	Set of regular arranged spots of DNA recognition elements positioned on microscopic slides	Phylogenetic identification, semi-quantitative, fast	Low detection limit, Hybridization biases Novel species/strains unidentified	Paliy et al., 2009
16S rRNA gene sequencing	Massive parallel sequencing of partial 16S rRNA marker gene	Phylogenetic identification, quantitative, fast, identification of unknown bacteria	PCR bias, Extensive data analysis	McCartney et al., 2002; Dethlefsen et al., 2006
Metagenomic shotgun sequencing	Massive parallel sequencing of the whole genome	Phylogenetic identification, quantitative	Expensive, analysis of data is computationally intense	Breitbart et al., 2006; Manichanh et al., 2006

Abbreviations: DGGE, denaturing gradient gel electrophoresis; TGGE, temperature gradient gel electrophoresis; T-RFLP, terminal restriction fragment length polymorphism.

1.8 Problem statement

On a global scale, there are clear knowledge deficits regarding our understanding of the human gut microbiome. Moreover, the studies have focused on providing insights regarding the microbiomes of individuals from Europeans, and US. Cross-population studies have only focused on population from remote rural communities whose lifestyles resemble either the Neolithic revolution or traditional/cultural communities (hunter-gatherers or traditional farmers/agriculturists) compared to individuals from highly industrialized societies. These study designs are important for understanding the interaction of the host-microbes and may explain some of the environmental adaptations of the gut microbiome during the evolution of humans and their lifestyles. However, there is still a gap in knowledge about the composition of the gut microbiota of modern human populations from distinct geographical areas, with different lifestyles, yet similar genetic and ethnic background. Moreover, the transition from

rural to urban and the subsequent adaptations to western lifestyles and dietary habits, may alter and decrease the gut bacterial diversity of traditional communities.

1.8.1 Aim

To elucidate the gut microbiomes of healthy South African individuals from rural and urban localities.

1.8.2 Objectives

- To characterize the gut microbiota of rural and urban populations.
- To determine the relative influence of diet, geography and different lifestyles on the gut microbiota.

Chapter 2

Materials and methods

2.1 Ethical clearance

Ethical clearance for the study (EC160630-051) was approved on 26/04/2017 by the Faculty of Health Sciences Research Ethics Committee and Faculty of Natural and Agricultural Sciences: Ethics Committee at the University of Pretoria. All the study procedures and experiments were conducted in-line with the relevant ethical guidelines and regulations.

2.2 Study area

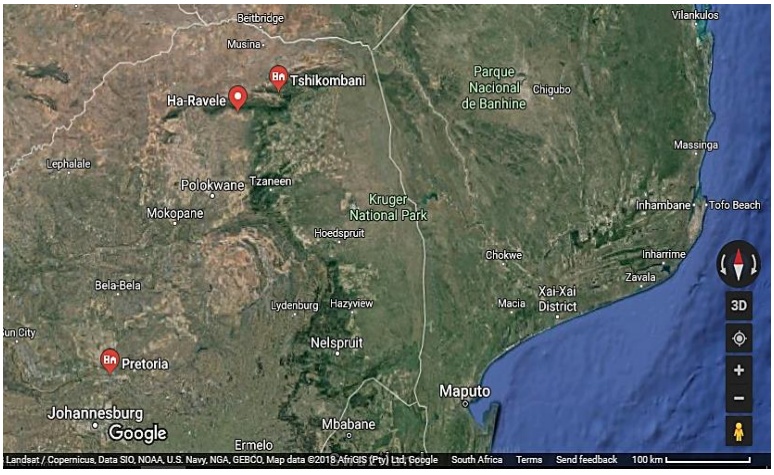


Figure 2.1. The geographic locations of Ha-Ravele, Tshikombani and Pretoria. (Source: Google Maps).

The City of Tshwane (Pretoria) one of the major metropolitans in Gauteng Province was identified as the urban location of the study. Ha-Ravele a village outside a small town called Makhado (Vhembe District in Limpopo Province) was the identified rural location. However, to get to the total number of required participants, additional number of participants were recruited from another village called Tshikombani located north-east of Makhado Town. The geographical locations of the three sampled sites are shown in **Figure 2.1**.

While the two selected rural communities are a modern human population, they still practice some of their traditional life ways especially when it comes to the food they eat and the preparation. They still cook food in open-fire (**Figure 2.2A**) and they grow and process their own maize meal (**Figure 2.2B**), a common staple food in rural areas. Subsistence farming is still a very common practice and some people continue to rely on it for their food supply. Most households have their own garden where they grow maize in summer seasons, and leafy vegetables (Chinese cabbage, *Amaranthus, Solanum retroflexum, Corchorus olitorius, Cucurbita pepo* and spinach) all year round. A typical lunch or supper meal contain maize meal (porridge), a portion of leafy vegetable and sometimes with protein (beef/chicken/fish) **Figure 2.2C**. Thus, their diet contains a substantially amount of dietary fiber. Additionally, participants from both rural areas have contact with the urban society when they go approximately twice a month to buy food and other household supplies.

Figure 2.2. Traditional lifestyle features of Ha-Ravele and Tshikombani populations. (A) Washing the maize grains before grinding them to make maize meal, (B) cooking food in open fire remains one of the most preferred methods to cook food in these communities and (C) a typical supper or lunch meal. (Images source: My Venda (@MyVenda) | Twitter, https://twitter.com/MyVenda)

2.2 Participant recruitment

Participants from Ha-Ravele were a group of soccer players and were recruited through their team manager. Participants from Pretoria (Hartfield and central business district) and Tshikombani were recruited randomly through direct approach. A total of 100 healthy volunteers equally divided between genders and locations (rural (25 males and 25 females) and urban (25 males and 25 females)) were recruited, respectively. Participants were aged between 18 – 50 years, had not taken antibiotics for at least 6 months prior to volunteering,

were healthy and have not been diagnosed with any inflammatory-related bowel diseases or gastrointestinal diseases (i.e stomach ulcers). Volunteers were given a briefly description of the study protocol and completed informed consent forms to participate including agreement with the study confidentiality statement.

2.3 Sample collection

A total of 100 stool samples were collected between the months of April to July 2017. Participants collected the first stool of the day, using the provided fecal/stool collection kit (Easy Sampler® Stool collection Kit, Hounisen Lab Equipment A/S, Skanderborg, Denmark). Volunteers were briefed on how to use and collect the sample with minimal contamination following the detailed instructions included in the sampling kit. After sample collection, samples were collected from participants within 2 hours and transported on dry ice in a cooler box to the laboratory at the University of Pretoria. Samples were stored at -80ºC prior DNA isolation. Samples will be stored at -80ºC for a maximum of 15 years for use in further research studies.

Upon sampling, volunteers were required to complete a questionnaire focusing on dietary, lifestyle, anthropometrics, Bristol Stool Chart (BSC) score, and medical/clinical information (**Figure 2.3**). The medical information included: blood type, usage of antibiotics and/or other medication six months prior sample collection and if participant has been diagnosed with gastrointestinal related diseases. The height and weight of each volunteer was recorded to calculate the body mass index (BMI) using a normal household bathroom scale and measuring tape (>2 meters).

Figure 2.3. A summary of the covariates collected during sampling. The hexagon radial chart reflects each of the metadata category. The covariates and responses are listed in a box of corresponding colour. The number in the brackets is the number of participant's responses.

2.2 DNA extraction

Stool samples were thawed for 30 min at room temperature prior to DNA extraction. DNA was isolated using MO BIO PowerSoil DNA Isolation Kit (MO BIO Laboratories Inc., Carlsbad, CA) following manufacturer's instruction, with minor modifications. Briefly, approximately 0.25 grams of stool sample was aliquoted into Power-Bead tubes using a sterile disposable wooden

spatula (Lasec Laboratories, RSA). The tubes (with samples) were gently vortexed to mix prior to adding 60µl of the lysis reagent. Samples were incubated in a water-bath for additional 30 min at 55°C to maximize lysis (additional step not specified in the protocol). The suspension was centrifuge at 10,000 x g for 30 seconds at room temperature prior to transferring 2 ml of supernatant into a clean collection tube. Then, 250 µl of inhibitor removal reagent (provided) was added into the samples and incubated on ice for 5 minutes. The inhibitor removal reagent removes organic and inorganic substances that may reduce DNA purity. About, 1.2 ml of highly salt concentrated solution was added to allow DNA to bind to the binding column. Next, 500µl of ethanol was added then centrifuged at room temperature for 1 minute at 10,000 x g to wash the bound DNA. The DNA was eluted into 100µl filter-sterilised autoclaved Millipore water. NanoDrop™ 2000/2000c Spectrophotometer (Thermo Scientific, Waltham, MA, USA) was used to perform DNA quantification following manufacturer's protocol.

2.3 Agarose gel preparation and electrophoresis

The agarose gels were prepared using SeaKem® LE Agarose powder (Sigma Aldrich, St Louis, USA). Two percent [w/v] agarose gels prepared in 100ml TAE buffer solution (0.2% [w/v] Tris, 0.5% [v/v] acetic acid, 1% [v/v] 5 M EDTA [pH 8]). Samples were mixed with GelRed loading dye (Bio-Rad Laboratories Ltd, RSA) in a ratio of 1:3. Next, samples were loaded and ran for 30 to 90 minutes at 90V. The DNA fragments were compared against a 1 kb DNA Ladder (KAPA Biosystems Inc., Wilmington, USA) under the UV light using digital imaging system, Molecular Imager® GelDocTM XR+ System (Bio-RAD, RSA) and Image LabTM Software.

2.3 16S rRNA gene amplification and sequencing

The 16S rRNA genes PCR amplification and sequencing were performed at MR DNA (www.mrdnalab.com, Shallowater, TX, USA).

2.3.1 16S rRNA gene PCR amplification

The V4 region of the 16S rRNA gene was amplified using the 515F/806R (Caporaso et al., 2011) primer set. Polymerase chain reaction (PCR) reactions were carried out in triplicates of 25µl, each containing 13.0 µl of PCR grade water (MOBIO Laboratories Inc., Carlsbad, CA), 10.0 µl of HotStarTaq Plus Master Mix kit (containing DNA polymerase, buffer, MgCl2 and dNTPs, Qiagen, USA), 0.5 µl of forward primer (10µM), 0.5 µl reverse primer (10µM) and 1.0 µl of DNA template. Thermal cycling conditions: initial denaturation step at 94°C for 3 minutes, followed by 28 cycles of final denaturation step at 94°C for 30 seconds, annealing at 53°C for 40 seconds and extension at 72°C for 1 minute and the final elongation step at 72°C for 5 minutes. PCR products were assayed in 2% agarose gel as described in section 2.3, the expected fragment bands were 300 – 350 bp. All samples (triplicates) were pooled together in equal proportions based on their molecular mass and DNA concentrations (standard amount: 240 ng of DNA per sample and 4nM). Next, pooled samples were purified using calibrated Ampure XP beads (Beckman Coulter Inc.).

2.3.2 16S rRNA gene sequencing

The pooled and purified PCR products were used to prepare DNA library using Illumina TruSeq DNA library (adds adaptors and sequencing primers) (http://www.illumina.com/) following manufacturer's protocol. Sequencing was performed on Illumina MiSeq Platform (Illumina, San Diego, CA, US) using MiSeq Reagent Kit version 3 (http://www.illumina.com/) to obtain 2 x 300 base pairs (bp) reads following manufacturer's protocol.

2.4 Data and statistical analysis

2.4.1 16S rRNA gene sequence data processing

The 16S rRNA gene sequences were processed (demultiplexing and filtered) using Quantitative Insights into Microbial Ecology (QIIME) pipeline (v1.9.1) (Caporaso et al., 2010). Barcodes, primers, adapter sequences, sequences with ambiguous base call and sequences less than 150 bp were removed. Chimeric sequences were removed through Uchime implemented in QIIME (Edgar, 2010). Operational Taxonomic Units (OTUs) were clustered at 97% sequence similarity. Sequences that failed to match the reference set were subsequently clustered de novo at 97% similarity using UCLUST algorithm (Edgar, 2010). The final OTUs were taxonomically classified using BLAST alignment method against the Greengenes 13_8 reference database (McDonald et al., 2012).

2.4.2 Bacterial community composition analysis

Abundance OTUs table and taxonomic assignments in each sample was generated from QIIME (Caporaso et al., 2010). Bacterial relative abundance and heat map plot at genus level analysis were performed using MicrobiomeAnalyst online software (http://www.microbiomeanalyst.ca.) (Dhariwal et al., 2017).

2.4.3 Analysis of the bacterial community metabolic potential

The microbiome functional profiling was predicted using phylogenetic investigation of communities by reconstruction of unobserved states (PICRUSt) (Langille et al., 2013). PICRUSt is a computational approach to predict the metabolic functional potential of the bacterial community by matching 16S rRNA maker gene sequences with a closely related reference genome annotation, Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathways database (Kanehisa & Goto, 2000). Using the clustered (at 79%) OTU file derived from QIIME, a reference phylogenetic tree is created from Greengenes database (DeSantis et al., 2006) and the gene contents are assigned to nodes. PICRUSt might also use ancestral state reconstruction algorithms if the sequenced genomes are unavailable (Langille et al., 2013). This creates an annotated table of gene counts. The gene counts were normalized per

sample and inferred in KEGG (Kanehisa & Goto, 2000) classifier implemented in PICRUSt. The relative abundance of functional pathways in the microbial community between groups was analysed and visualized using Statistical analysis of taxonomic and functional profiles (STAMP) software (Parks, 2014). KEGG functional analysis was performed at level three which shows different metabolisms and pathways (e.g carbohydrates, amino acids, etc) of the bacterial community.

2.4.4 Statistical analysis

Statistical analysis and visualizations were performed in R Studio version 3.2.3. (http://www.r-project.org) (R Core Team, 2016), using vegan (Oksanen et al, 2017) and phyloseq (McMurdie & Holmes, 2013) packages. Diversity parameters (Shannon and observed OTUs) of the bacterial community were compared using analysis of variance (ANOVA) and Welch t-test. Non-metric distance scaling (nNMDS) plots using the Bray-Curtis, and unweighted UniFrac distance measures and permutational multivariate analysis of variance (PERMANOVA) were performed in MicrobiomeAnalyst (Dhariwal et al., 2017). Differences in of abundance of bacterial phylum, genera and KEGG pathways between populations were performed in STAMP (Parks et al., 2014) using Welch's t-test with the false detection rate correction. Correlations between bacterial taxa abundance and environmental variables or metadata was determined by Pearson's test in R Studio (R Core Team, 2016), using the MicrobiomeSeq package (http://www.github.com/umerijaz/microbiomeSeq).

Chapter 3

Bacterial community diversity and composition

3.1 Introduction

The role of the gut microbiome and its importance to human health has become a prominent area of research (Lloyd-Price et al., 2016). In the past decade, several studies have demonstrated the influence of the gut microbiome on host health, revealing crucial insights regarding its non-homogenous nature across populations (Consortium, 2010; Ehrlich & Consortium, 2011; Human Microbiome Project, 2012; Nishijima et al., 2016). Other studies have suggested that the variations in community composition may be due to factors including different lifestyles and dietary habits (De Filippo et al., 2010), hygiene practices (Martinez et al., 2015), pathogen exposures (Morton et al., 2015), geographical locations (Suzuki & Worobey, 2014; Yatsunenko et al., 2012), cultural traditions and subsistence stratagems (Obregon-Tito et al., 2015). However, due to the high variability among individuals, key questions remain regarding the role played by each of these factors in structuring the microbiome and regulating associated functional processes.

While it is not yet clear which factor plays a dominant role in shaping the gut microbiome, several studies have attributed variations in gut microbiomes across populations to methodological biases such as primer choice (Brooks et al., 2015), DNA extraction and amplification (Gerasimidis et al., 2016). However, these reasons may not fully explain these variations as several studies using similar methodologies have also shown high variability across individuals (Turnbaugh et al., 2007; Human Microbiome Project, 2012; Huttenhower et al., 2012). Taken together, these studies suggest that the extent to which methodological biases skew observations in gut microbiome studies may be overestimated (Clooney et al., 2016).

To understand these differences and the high variability in the structure of the gut microbiomes of healthy individuals, several studies have assessed the microbiomes of individuals living in

urban and rural locations. These studies have assessed the combined effects of different lifestyles and environmental factors on the gut microbiome (Yatsunenko et al., 2012; Schnorr et al., 2014; Obregon-Tito et al., 2015). Geographic locality includes unique factors such as the local environment, host genetics, and cultural traditions or lifestyles, which may play a central role in the assembly of the gut microbiomes (Gupta, 2017). For instance, studies assessing rural and urban individuals have shown variations in the composition of the gut microbiomes between urban and rural individuals (De Filippo et al., 2010; Gomez et al., 2016). These differences were ascribed to high-fat and -sugar consumption in urban individuals in contrast to high fibre low-fat and sugar diets in rural populations (Nakayama et al., 2017; Kisuse et al., 2018). Recent studies have also shown that rural populations subject to urbanization or industrialization tend to harbour comparatively lower gut bacterial diversity (Ayeni et al., 2018).

A study by Martinez and colleagues (2015), demonstrated that western diet, lifestyles including hygiene practices were associated with reduced bacterial dispersal rates, gut microbiome structure alterations and higher inter-individual variations in the urban population (Martinez et al., 2015). Gomez and colleagues (2016) investigated two rural traditional communities (hunter-gatherers and Bantu agriculturalists) and demonstrated that their gut microbiome profiles follow a gradient of traditional subsistence strategies and lifestyles (Gomez et al., 2016). The disparate geographic locations provide an opportunity to understand factors contributing to the high variability and may clarify the influence of local environmental exposures on the gut microbiota (Clemente et al., 2015).

However, many studies on rural and urban individuals have focused on geographically and culturally distant populations, where confounding factors such as diverse host genetics and ethnicity cannot be separated. For instance, traditional rural individuals from Africa and South America are often compared to western individuals from Europe and North America (Yatsunenko et al., 2012; Morton et al., 2015; Rampelli et al., 2015). Host genetics has been shown to play a role in the composition of the gut microbiome (Goodrich et al., 2014; Bonder

et al., 2016), which may explain observed high variabilities in the gut microbiomes across populations from different ethnic groups (Dehingia et al., 2015; Fujio-Vejar et al., 2017). To reduce the effects of the different confounding factors, this study will characterize the gut microbiome of healthy black South African adults from rural and urban areas. We recruited participants from two villages in the Vhembe district of the Limpopo Province, representative of the rural cohort and other cohort from Pretoria, one of the major metropolitan areas in Gauteng Province, South Africa. To assess community composition and structure, 16S rRNA gene amplicon sequencing was used. By comparing the gut bacteria of these two populations, we predict that the rural and urban gut microbiome profiles will have low inter-population variation, due to low genetic variability between the cohorts. In addition, it is expected that the rural population may harbor distinct gut microbiota from urban individuals due to more traditional lifestyles and dietary regimens. It is also predicted that low inter-individual variation in the gut microbiomes will be observed in rural individuals and high inter-individual variations in urban individuals.

3.2 Results

3.2.1 Characteristics of participants

In total, 100 stool samples from healthy male and female adults were collected for this study. Five of those samples were excluded from further analysis due to sample quality. All 5 samples were retrieved from rural males. The 95 remaining samples, comprised of 50 urban (25 females and 25 males), and 45 rural (25 females and 20 males) were analyzed.

Metadata from participants recording the diet, lifestyle, anthropometrics, and medical or clinical characteristics of all 95 healthy adult participants are listed in **Table 3.1**. Participants ranged from 19 to 49 years (30,5 ± 8,77) in urban and 18-35 years (24 ±5,38) in the rural group. Ninety-one participants including rural and urban were obese with BMI values above 30kg/m^2, while the remaining four were overweight with BMI values greater than or equal to 25kg/m^2. All participants from both rural and urban areas were omnivorous. The mean intake

of total dietary fibre from fruits and vegetables was 2.22 grams and 1.52 grams per day (g/day) for rural and urban population, respectively. About 51% of rural participants consumed fast-foods at a frequency of between 2 to more than 5 times a month. In contrast, 50% of the urban cohort declared fast-food (or takeaways i.e burgers, chips, pizza, etc) consumption at a minimum of 3 times per month to daily. On average, 49% of rural participants consumed fast foods once a month. In terms of dietary fiber intake, 69% of rural participants indicated that they consume an average of 3 portions of fruits and/or vegetables per day, while most of the urban individuals consumed at least one portion a day. When the difference was compared using one-way ANOVA, the difference in fibre-rich food consumption between rural and urban cohorts was statistically significant ($p = 0.001$). Most of both rural and urban alcohol consumers indicated that they consume at least 4 drinks a month. These beverages in included beers, ciders, wines or a shot of liquor.

Table 3.2: Participants characteristics. Data are expressed by mean and standard deviation (SD). The p - values are determined by one-way ANOVA.

Characteristics	**Rural females**	**Rural males**	**Rural Group**	**Urban females**	**Urban males**	**Urban group**	**P-value (ANOVA)**
Anthropometrics							
Age mean ±SD	*25.52(5,11)*	*22,1(5,21)*	*24(5,38)*	*31,2(9,20)*	*29,8(8,45)*	*30,5(8,77)*	*<0,001*
Age range	*18-37*	*18-35*	*18-35*	*19-46*	*19-49*	*19-49*	*<0.001*
BMI, mean ± SD	*25,5(5,5)*	*40,04(4,16)*	*41,13(10,08)*	*46,1(11,44)*	*44(7,9)*	*45(9,8)*	*0,03 (Kruskal test)*
Clinical information							
Birth-mode	*v(20);c(5)*	*v(20)*	*v(40);c(5)*	*v(22);c(3)*	*v(24);c(1)*	*v(46);c(4)*	-
Birth-place	*hp25*	*hp20*	*hp45*	*hp25*	*hp24; h(1)*	*hp59;h(1)*	-
Blood type	-	-	-	-	-	-	-
Digestive diseases	-	-	-	-	-	-	-
Dietary information							
Pap (maize meal)	*d25*	*d20*	*5(0)*	*d(11);w(14)*	*d(17);w(8)*	*3,82(1,42)*	*0,001*

consumption (mean ±SD)							
Fast foods (mean ±SD)	-	-	*1,8(1)*	-	-	*2,56(1,4)*	*0,003*
Fruits and vegetables (mean ±SD)	-	-	*2,22(0,52)*	-	-	*1,52(0,61)*	*0,001*
Lifestyle							
Smoker	*0*	*4*	*1,91(0,28)*	*3*	*8*	*1,76(0,43)*	*0,05*
Drinker	*6*	*4*	*1,77(0,42)*	*13*	*12*	*1,5(0,50)*	*0,04*
Alcohol use p/m	*4 glass*	*4 glass*	*0,47(0,89)*	*4 glass*	*4 glass*	*1,02(1,22)*	*0,01*
Others							
Breastfeed	*24*	*20*	*1,02(0,14)*	*25*	*25*	*1,16(0,37)*	*0,02*
Breastfeed period	*> 6 month*	*> 6 month*	*3,68(0,51)*	*> 6 month*	*> 6 month*	*0,04(0,56)*	*0,002*
Antibiotics	*2*	*0*	*-*	*5*	*1*	*-*	*-*
Probiotics	*0*	*0*	*-*	*4*	*6*	*-*	*-*
Bistrol chart	*-*	*-*	*-*	*-*	*-*	*-*	*0,02 (chi-square)*

Abbreviations: v – vaginal, c – caesarean, hp – hospital, h – home, d – daily, w – weekly.

3.2.2 Most abundant phylotypes were shared between rural and urban Individuals

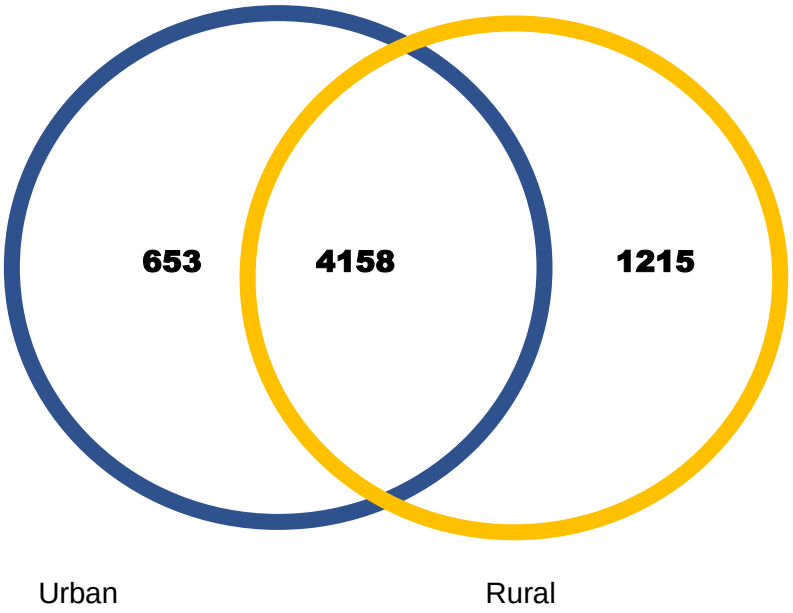


Figure 3.1 Venn diagram showing unique and shared OTUs between two populations.

The 16S rRNA gene sequencing resulted in a total of 7 701 098 reads, with an average of 297.38. Analysis of quality trimming resulted in 5 821 330 reads, grouped into 5 863 OTUs using a 97% cut-off threshold. Samples were rarefied to the lowest sample number of 31 924. The majority of OTUs were shared between rural and urban population, 4158 OTUs (69%), with 653 OTUs (11%) and 1215 OTUs (20%) assigned to urban and rural individuals, respectively (**Figure 3.1)**. This suggests higher gamma diversity in the rural population.

3.2.3 The gut microbiota between rural and urban individuals showed no significant difference in bacterial diversity

To determine bacterial richness between the two populations several alpha diversity measures were using (**Figure 3.2)**. Shannon diversity showed no significant difference in bacterial diversity between rural and urban individuals (Welch test: p = 0.3; **Figure 3.2A**). There was no difference in Shannon diversity was observed based on gender between rural and urban individuals (Welch test: p = 0.1 for males and p = 0.3 for females, respectively **Figure 3.2B-C**). Bacterial species richness based on Observed OTUs revealed a significant difference in species richness between urban and rural individuals (ANOVA: *p = 0.001* **Figure 3.2D**).

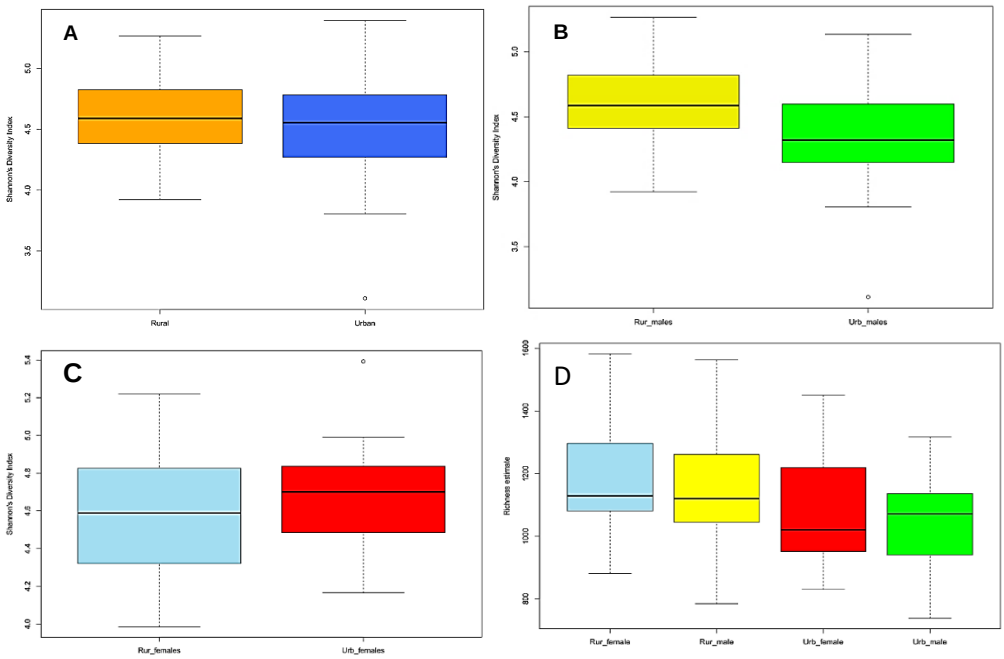


Figure 3.2 Boxplots of the alpha diversity indices. Plot A – C is comparison of rarefied Shannon diversity in the gut microbiota of rural and urban populations (plot A: rural-orange, urban-blue; plot B: rural males-yellow, urban males-green; plot C: rural females-blue, urban females-red). Plot D is the species richness estimate based on Observed OTUs compared between rural and urban cohorts (rural females-blue, rural males-yellow, urban females-red, urban males-green).

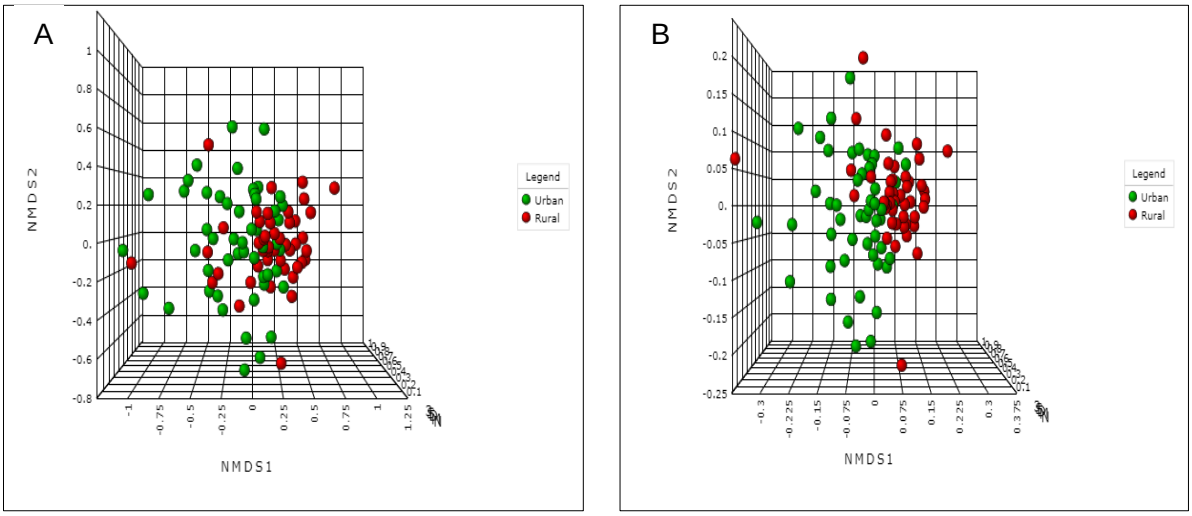


Figure 3.3: (A**)** Non-metric multidimensional scaling (nMDS) plot based on Bray-Curtis dissimilarity index shows separations between rural and urban populations. (B**)** nMDS plot based on unweighted Unifrac distance to show relative contributions of other environmental factors shaping the bacterial communities from rural and urban populations.

3.2.4 High dispersion between two populations consistent with geography

Beta-diversity based on Bray-Curtis dissimilarity matrix was used to assess the contribution of geography on the population structure of the gut microbial communities. The rural and urban gut bacterial communities were heterogeneous and clustered based on geographic location (PERMANOVA: $R^2 = 0.03$, $p < 0.001$, higher β-diversity; **Figure 3.3A**). To further assess differences in gut bacterial communities of the two populations, beta-diversity based on unweighted Unifrac distance was performed to disentangle the relative contributions of other environmental factors shaping the bacterial communities. The results suggest that rural bacterial community composition is similar and more homogenous whereas urban samples is more variable and heterogenous **Figure 3.3B**.

3.2.5 Bacterial community relative abundance

3.2.5.1 Difference in Faecal Bacterial Communities Between rural and urban individuals

Gut bacterial community abundance was compared at different taxonomic levels to determine bacterial taxa in rural and urban populations. In total, 11 phyla contributed more than 2% relative abundance of the gut bacteria in rural and urban populations (**Figure 3.5**). At phylum level, *Firmicutes* dominated the gut bacteria of rural and urban individuals with no significant difference (relative abundance, 66% and 62% in rural and urban respectively; t-test: $p = 0.33$). In both groups, *Bacteroidetes* were the second most dominant phylum and significantly higher in urban compared to rural population (relative abundance, 20% and 28% in rural and urban, respectively; t-test: $p = 4.83e^{-4}$). *Proteobacteria* were marginally higher in rural samples (relative abundance, 8% and 6% in rural and urban respectively; t-test: $p = 0.01$). The main discriminant phylum was *Spirochetes*, which was only present in rural samples (relative abundance, 3%). Moreover, less abundant phyla included *Actinobacteria*, *Verrocumicrobia*, *Lentisphaerae*, *Elusimicrobia*, *Tenericutes*, *Cyanobacteria*, *Euryarchaeota*, and unknown phyla which contributed about 3% and 4% relative abundance in rural and urban, respectively.

However, *Actinobacteria* and *Verrocumicrobia* were significantly higher in urban samples than in the rural group (t-test: p = 0.01 and p = 0.01 respectively).

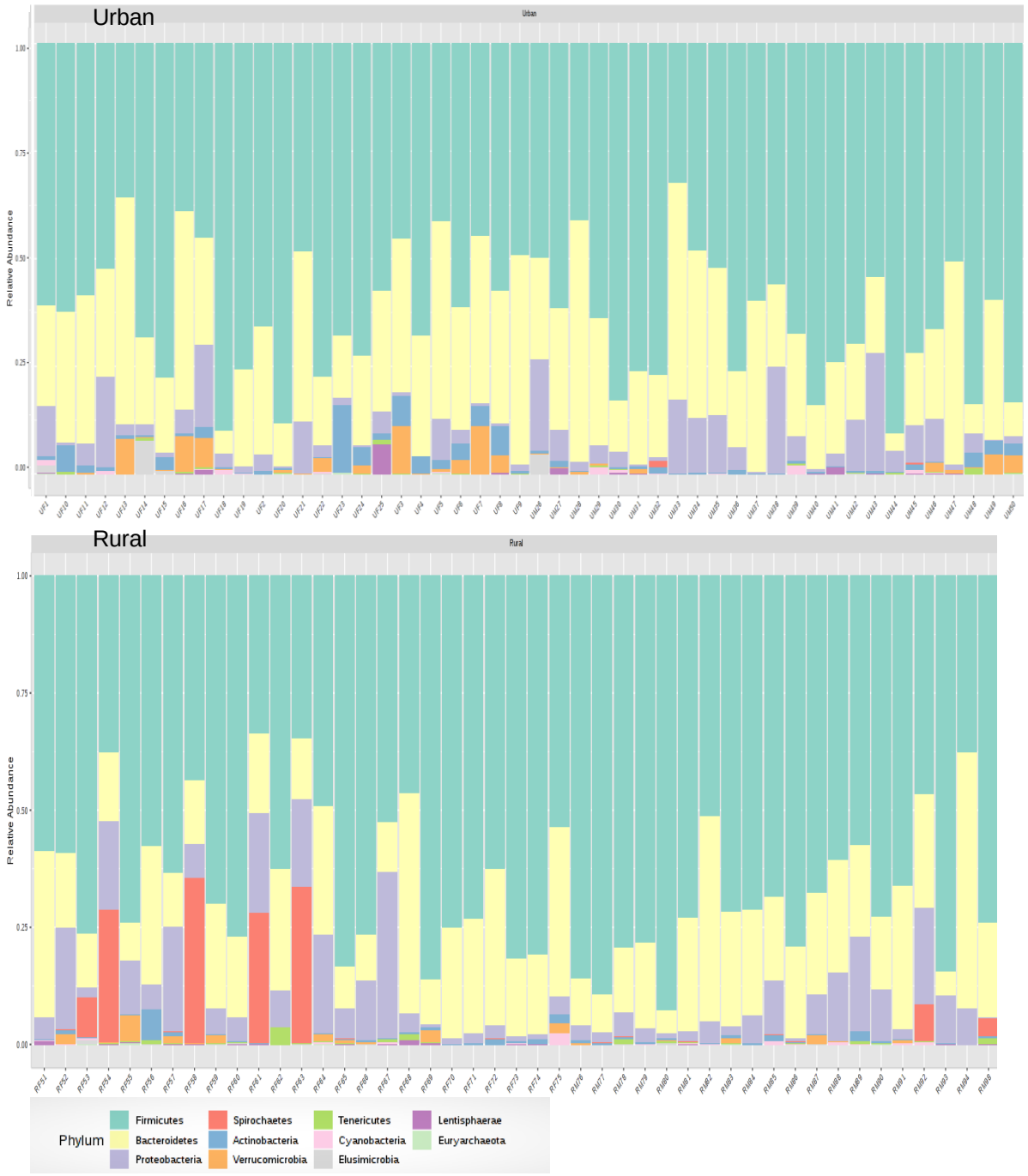


Figure 3.5. Phylum Composition of rural and urban gut bacterial communities that contribute more than 2% of the relative abundance of each sample. On the x-axis is the sample names and the y-axis the relative abundance

At genus level, most sequences were unclassified (**Figure 3.6**). The rural and urban populations were enriched with *Prevotella* with no significant difference between the two

groups (t-test: p= 0.38). However, *Bacteroides* were significantly higher in urban samples (t-test: p= $1.18e^{-4}$). *Blautia*, *Parabacteroides*, *Succinivibrio*, *Bifidobacterium*, *Akkermansia*, *Oscillospira*, and *Sutterella* were significantly abundant in urban samples, whereas only

Roseburia were significantly abundant in rural samples. Other genera that were present but showed no significant difference between the two groups included *Faecalibacterium, Ruminococcus, Clostridium, Coprococcus, Dorea, Haemophilus, Lachnospira, Eubacterium, Streptococcus, Veillonella, Catenibacterium, Desulfovibrio, Dialister,* and *Phascolarctobacterium*. Interestingly, when the taxa at genus level were adjusted to merge all small taxa with less than 10 000 counts, *Treponema* were detected only in rural samples.

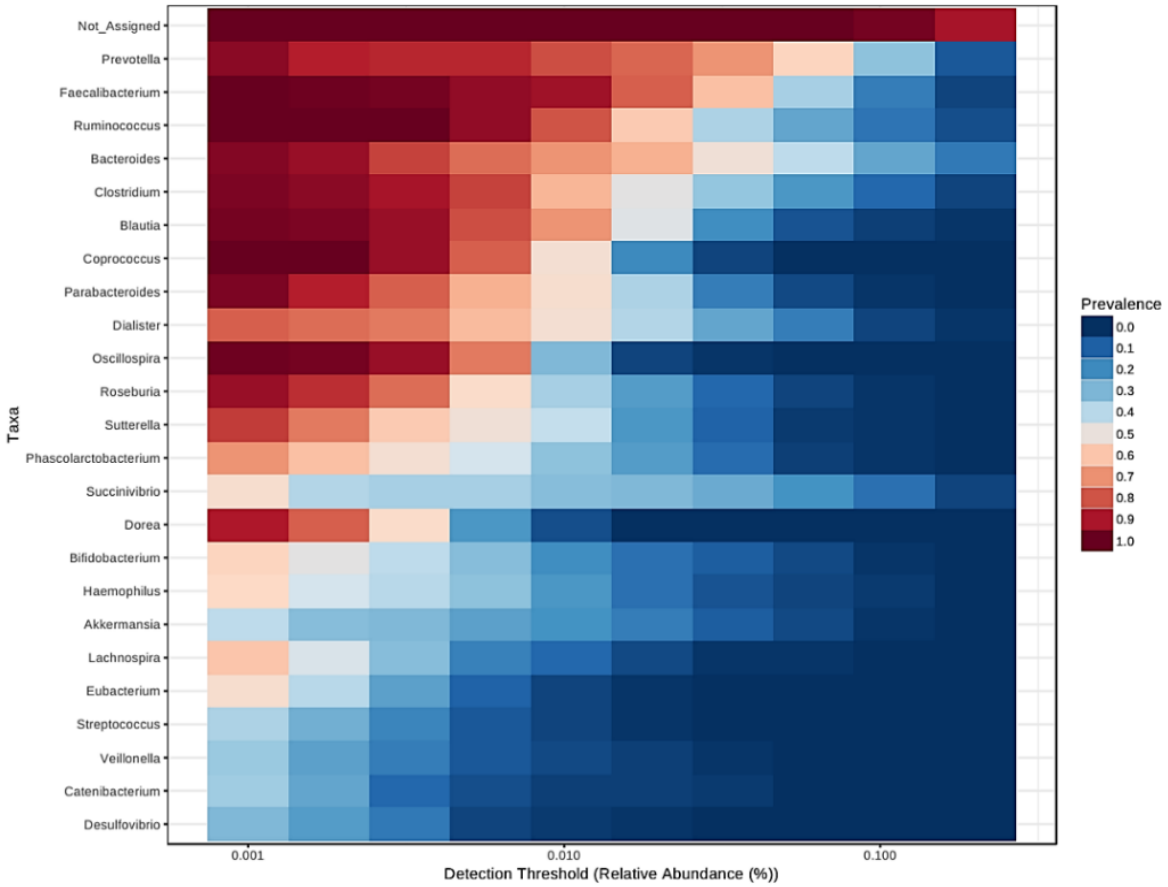


Figure 3.6 Heat-map of taxa at genus level that remained unchanged (core microbiome) in composition across all samples (rural and urban), based on relative abundance (>2%).

3.2.6 Bacterial taxa associations with metadata

Pearson's correlation (p-value < 0.05) analysis were performed to identify specific microbial taxa that correlated with metadata such as diet, smoking, BMI, and dietary factors between two populations (**Figure 3.7**). At phylum level, the rural group showed a strong association between breastfeeding and the abundance of *Euryarchaeota* ($p < 0.001$) at phylum level and *Methanobrevibacter* ($p < 0.001$) at the genus level (98% participants were breastfeeding). Our data also revealed positive correlations between the breastfeeding period and the abundance of *Firmicutes* in the rural group ($p < 0.01$). A total of 98% of the rural group indicated that they were breastfed for at least 6 months to over 1 year. Inversely, *Proteobacteria* showed a statistically significant ($p < 0.01$) negative correlation with the breastfeeding period in the rural group. A strong negative association ($p < 0.01$) between *Lentisphaerae* abundance and antibiotics was observed in the rural group (96% of participants who did not take any antibiotics 6-month prior sample collection). In the urban group, there was a statistically significant ($P < 0.01$) positive correlation between mode of birth and *Lentisphaerae* taxa. At genus level, fruit and vegetable consumption was strongly correlated with *Sutterella* and *Parabacteroides* relative abundance in the urban group, $P < 0.01$. Pap (maize meal), probiotics, birthplace, blood type, the frequency of stool per day/week, Bistro chart, menstrual cycle and pregnancy (between women) all showed some level of positive and negative correlations at phylum and genus level in the two populations (not shown in the heat-maps). However, these factors were not statistically significant.

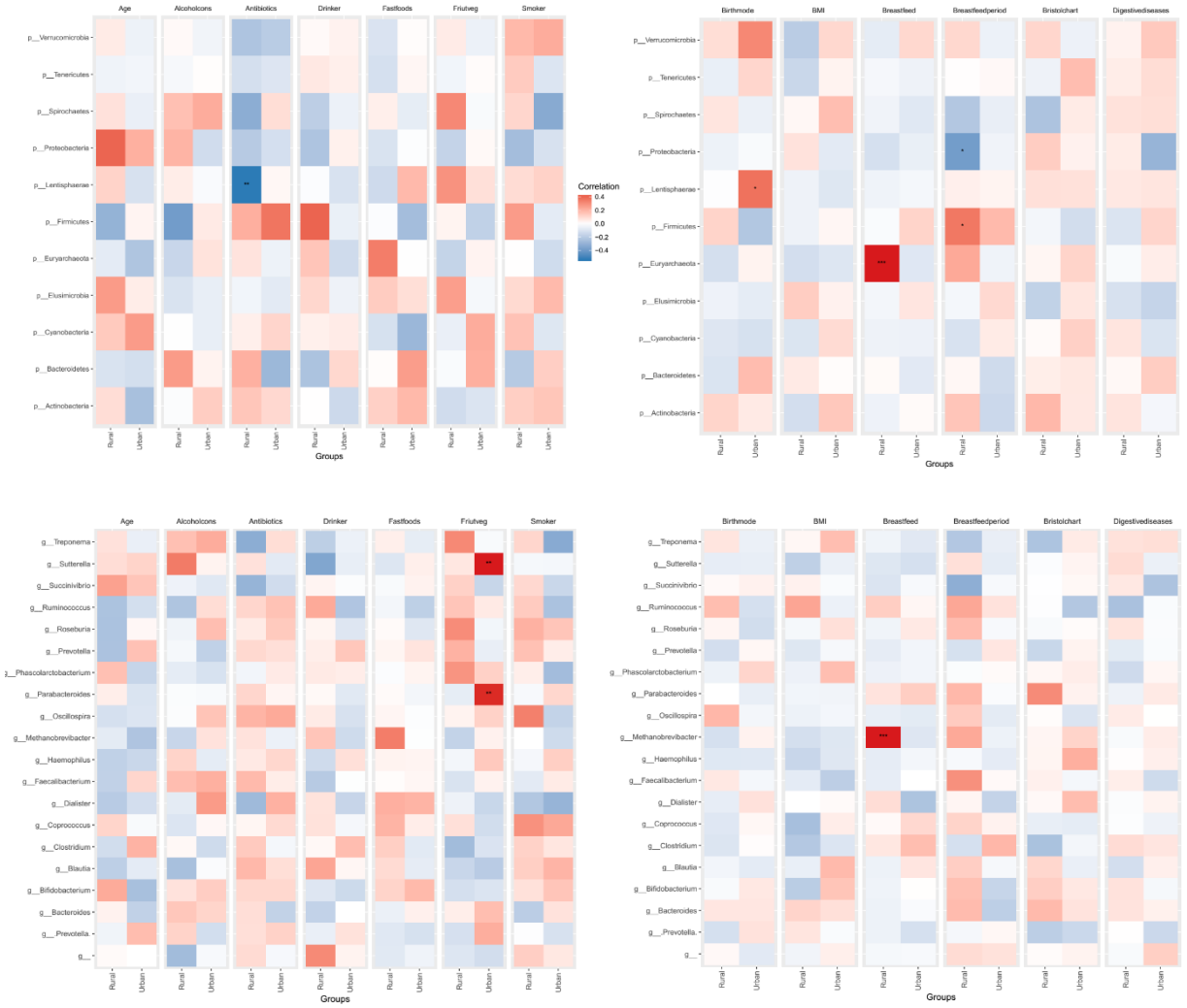


Figure 3.7 The heat-maps show Pearson's correlation (p-value less than 0.05) between bacterial community abundance and collected metadata for the two populations at phylum and genus level.

3.3 Discussion

This study characterized the intestinal bacterial community of 45 healthy adults from rural villages and 50 healthy adults living in Pretoria. Several previous studies showed that the gut microbiomes of rural individuals differ from individuals living in urban areas with western lifestyles (De Filippo et al., 2010; Yatsunenko et al., 2012; Clemente et al., 2015). These studies have reported that alpha diversity is higher in rural individuals (De Filippo et al., 2010; Yatsunenko et al., 2012; Clemente et al., 2015). These studies have also revealed diet as the primary driver of variations in the gut microbiome. Contrary to previous studies, the focus of this study was to compare the gut microbiomes of rural and urban individuals with the same

ethnicity. In this study, alpha diversity showed no significant difference between the two populations. Although current research suggests that diet exerts the largest effect on the gut microbiomes (Donaldson, 2016), these results suggest that when the gut microbiomes are not confounded by genetic variations, diet may have a minor influence on the gut microbiomes individuals.

Community structure visualized using non-metric multidimensional scaling (nMDS) of unweighted UniFrac distance revealed distinct clustering along nMDS1 between the two groups, indicating a strong core division in gut microbiome phylogeny between rural and urban individuals. The mean values of unweighted UniFrac distance also show a lower within-group variability among rural individuals than the urban group. This similarity in the phylogeny among rural cohort is probably a result of proximity community living and parallel lifestyles. It is likely that the homologous pattern in the microbiomes of rural individuals is also influenced by the consumption of the same foods. The rural cohort consisted of Vhavenda people who eat the same staple food (i.e pap and a side of meat, chicken or leafy-vegetables) and preparation is similar from household to household. Whereas, within the urban the structure of the gut microbiomes appears scattered, suggesting that there are several influences that are shaping the gut microbiomes. Similar results regarding high within-group variability in the gut microbiomes of urban cohorts have been reported (Ayeni et al., 2018; Obregon-Tito et al., 2015). Several factors can be attributed to the observed high within-group variability in urban. A study by Martinez and colleagues (2015) reported that observed heterogeneity in urban populations can arise from great variability in cultural backgrounds, lifestyle, and dietary habits when compared with non-urban populations (Martinez et al., 2015). In this study, the urban cohort consisted of individuals from a different cultural background with diverse lifestyle and dietary habits. Geography has a strong influence on the composition of gut microbiomes, results in this study showed that the gut microbiomes of rural and urban populations clustered according to geographic origin. Our findings support previous reports that geography plays a role in structuring gut bacterial communities (Yatsunenko et al., 2012; De Filippo et al., 2017)

The analysis of gut microbiomes was dominated by three major phyla, *Firmicutes, Bacteroidetes*, and *Proteobacteria* which is consistent with previously observed gut microbiomes of other human populations (The Human Microbiome Project et al., 2012). Previous studies have indicated that urban populations have less microbial richness than non-urban (or rural) populations (Schnorr et al., 2014; Rampelli et al., 2015). In this study, rural individuals had phylum *Spirochete* which was absent in urban individuals. A noteworthy feature of the fecal bacteria composition of the South African population is the equal dominance of *Firmicutes* in both rural and urban populations. These findings differ from previously reported urban and rural populations in that the *Firmicutes* to *Bacteroidetes* (F/B) ratio is high in urban populations while *Bacteroidetes* to *Firmicutes* (B/F) ratio is higher in rural population. Schnorr et al. (2014) found relatively high B/F ratio in Hadza, a rural hunt-gatherer community in Tanzania, while their urban (Italians) counterparts had a higher F/B ratio. De Filippo et al. (2010) also reported similar findings when children from Burkina Faso were compared with children living in an urban area in Italy. The F/B ratio relates to dietary habit (Portune et al., 2017) as well as host physiology (Mariat et al., 2009; Turnbaugh & Gordon, 2009). Populations consuming high amounts of animal protein/fats and high-fat-high-sugar diets (or western diets) tend to exhibit higher F/B ratios, whereas those consuming high amounts of vegetables (or plant-derived foods) and legume have higher B/F ratio (Gomez et al., 2016). Due to the modern lifestyles and socioeconomic status in many rural areas of South Africa, perhaps the rural cohort have adapted to the urban diets (including high consumption of animal protein, fats and sugar) which resembles western diets than that of other rural African areas from previous studies (Schnorr et al., 2014; Gomez et al., 2016).

Among minor phyla of gut microbiota, *Spirochaetes* and subsequently *Treponema* at genus level distinguished the rural from the urban population. *Treponema* like *Prevotella* is involved in degradation of resistant starch such as hemicellulose, and xylan (De Filippis & Ercolini, 2018). Previously, *Spirochaetes* were only reported from the intestinal microbiomes of non-human primates and ancient human populations (Ley et al., 2008). However, relatively high

abundances of *Spirochaetes* are observed among human populations with non-western lifestyles i.e., traditional, agriculturists and hunter-gatherer communities (Schnorr et al., 2014; Morton et al., 2015). As a result, few authors have suggested that the absence of *Spirochaetes* and *Treponema* represent a part of the human ancestral gut microbiome that gradually declined through adoptions to modern lifestyles (Obregon-Tito et al., 2015; Gomez et al., 2016). This explains why *Spirochaetes* and *Treponema* were absent in the urban populations.

The human gut microbiota can be assigned to one of three enterotypes, *Prevotella*, *Bacteroides* and *Ruminococcus* species which are driven by long-term diets (Arumugam et al., 2011). In a recent study, only *Bacteroides* and *Prevotella* enterotypes were associated with diets rich in animal protein and carbohydrate, respectively, but not *Ruminococcus* (Wu et al., 2011). Although the enterotype analysis was not performed like in previous studies, using the core gut microbiome analysis *Prevotella, Bacteroides* and *Ruminococcus* were among the core OTUs that remained unchanged across all 95 samples. These results suggest that the gut microbiome of rural and urban populations could be driven by three enterotype clusters influenced by their long-term dietary habits.

In addition, *Prevotella* were most prevalent in the composition in both rural and urban samples, while *Bacteroides* significantly dominated the urban samples. *Prevotella* contains a wide range of carbohydrate- acetate-, protein-fermenting and H_2-producing bacteria such as *Prevotella ruminicola* (Zhang et al., 2014). The dietary habits of the two populations are characterized by high amounts of starches such as maize, rice, and bread with low average daily consumption of vegetables and fruits in urban compared to rural. It is therefore not surprising that the *Prevotella* genera dominated the gut microbiomes of both rural and urban cohorts. On the other hand, *Bacteroides* have been associated with the metabolism of animal proteins, saturated fats and an array of amino acids.

Regarding correlations, *Sutterella* and *Parabacteroides* were strongly correlated with an average (p/d) consumption of fruits and vegetables in urban population. Although *Sutterella* species have been associated with human disorders such as, down syndrome, autism, and

inflammatory bowel disease, their role in the human gastrointestinal tract is less clear (Hiippala et al., 2016). *Parabacteroides* has been identified as one genus responsible for digesting high-fiber diets in the gut microbiome (Martínez et al., 2010), suggesting that there are some urbanites who consume fiber-rich diets. Studies have indicated that the dominance of members of *Methanobrevibacter* (prominent archaeon), particularly *M.smithii*, affects bacterial digestion of polysaccharides metabolism thereby influencing host calories to harvest and adiposity (Samuel & Gordon, 2006; Hansen et al., 2011). Here, phylum *Euryarchaeota* and subsequently the genus *Methanobrevibacter* were both strongly correlated ($P < 0.001$) with breastfeeding in rural populations. To our knowledge, we were unable to find the association of breastfeeding with relatively high levels of *Methanobrevibacter* or *Euryarchaeota* in literature. However, the high prevalence of genus *Methanobrevibacter* in infants (Grineet al., 2017) and adults is common and are involved in the production of methane through the reduction of CO_2 using H_2 or formate (Gaci et al., 2014).

3.4 Conclusion

The 16S rRNA gene analysis of the gut microbiome of rural and urban populations in South Africa revealed similar species diversity between two populations and are enriched with *Firmicutes*. We hypothesize that lack the of species diversity difference between rural and urban populations is due to less genetic variances between the cohorts. Several previous studies have demonstrated that human gut microbiomes are driven by long term diets and can be grouped into three enterotypes, the rural gut microbiome is driven by plant-based diet, thus *Prevotella* dominance, while the urban is driven by *Prevotella* and *Bacteroides*. Although geographic locations did not have any significant effect of the species diversity (Shannon index), several environmental factors such as geographic location and various life ways and dietary habits explained and separated the gut microbial community structures of rural and urban populations. Thus, in agreement with previous studies, the composition of the gut microbiomes of rural and urban populations varies.

Chapter 4

Predicted metabolic functional potential of the bacterial community

4.1 Introduction

The microbiome of healthy individuals varies substantially within and among individuals (The Human Microbiome Project et al., 2012). In addition to external environmental influences, ecological relationships among our microbial inhabitants have been shown to be important contributors to these variations (Faust et al., 2012). In nature, organisms coexist in complex ecological niches with various symbiotic relationships including; mutualism, commensalism (Woyke et al., 2006) and parasitism (Faust et al., 2012). Within these defined ecological niches, microorganisms may exchange or compete for nutrients, signalling molecules, or immune evasion mechanisms (Woyke et al., 2006). While such ecological interactions have been studied extensively in environmental microbiomes (Barberán, Bates, Casamayor, & Fierer, 2014; Cao et al., 2018; Eveillard et al., 2019), the range of normal interactions among human-associated microbes and how their occurrence influence host health or disease is less clear.

The intestinal microbiota interacts with the host in several ways with effects on the health and intestinal homeostasis (Parfrey et al., 2011). The human-associated microbiomes directly interact with the host through secreted metabolic products, immune modulation, and the capacity to perform various biochemical activities (Flint et al., 2012). Previous studies have shown that the gut bacterial composition is shaped by substrate availability (David et al., 2014), physiological status (Rampelli et al., 2013; Vesterbacka et al., 2017) and their response to dietary changes (Wu et al., 2011). For example, metabolic outputs of short-chain fatty acids (SCFA) by the gut microbiota may be influenced by the supply of dietary substrates and by diet-mediated changes in the composition of the microbiota (Richards et al., 2016). Thus, in several ways, intestinal bacteria have co-evolved with the human host to facilitate the

acquisition of nutrients and contribute to maintaining homeostasis in response to various human lifestyles.

The synthesis of gut bacterial metabolites is primarily influenced by diet, especially the proportion of proteins, fats, and non-digestible carbohydrates consumed. Metabolites are determined by the chemical composition or the structure of substrates and microbial metabolic pathways (Richards et al., 2016). Recently, whole-genome and/or amplicon approaches have shown that the metabolic functions of intestinal bacteria demonstrate extensive metabolic versatility (Abubucker et al., 2012; Huttenhower et al., 2012; The Human Microbiome Consortium, 2012). These studies have shown that the functional potential of the human gut microbiome is essential for understanding the role of the microbiota in health and disease (Huttenhower et al., 2012).

Several cross-population studies have compared the metabolic functional profiles of traditional societies with industrial populations, using shotgun metagenomics approaches (Yatsunenko et al., 2012, Rampelli et al., 2015, Obregon-Tito et al., 2015). These studies have provided further insights into functional adaptations of gut microbiota to various diet (De Filippo et al., 2010), cultural or traditions lifestyles (Rampelli et al., 2015), and subsistence practices (Obregon-Tito et al., 2015). The functional potential of rural populations includes enriched phosphotransferase and alpha-amylase pathways, typically associated with high consumption of foods containing complex plant polysaccharides (Obregon-Tito et al., 2015). The predicted functional potential of the gut microbiota of rural communities also indicated a lifestyle reliant on subsistence farming as prevalent in rural communities (Obregon-Tito et al., 2015). Meanwhile, predicted functional potential in the urban populations appear to be enriched in several metabolic pathways linked to the breakdown of sugars, bile acids and amino acids (Obregon-Tito et al., 2015). These metabolic pathways are consistent with western diets, which are enriched in simple sugars, animal proteins, and fats (Ou et al., 2013; David et al., 2014). Although functional profiles of industrialized and non-industrialized lifestyles highlight metabolic variations among populations with different lifestyles, it is still unclear whether the

observed metabolic functional trends are due to the types or quantities of foods consumed, traditional or cultural practices, geographic location, genetics, or other influencing factors. Furthermore, within African populations, the current available data on the functional aspects of the gut microbiome is based on populations with agricultural (Obregon-Tito et al., 2015) and hunter-gatherer (Rampelli et al., 2015) lifestyles.

Here, we present a comparison of the metabolic and functional potential of the gut microbial community from rural populations in the villages of Tshikombani and Ha-Ravele, and individuals living in Pretoria, a metropolitan area. The rural volunteers were comprised of individuals who partly relied on subsistence farming of vegetables, maize meal and some fruits. These individuals reported a diet with a high intake of plant-derived foods or carbohydrate-rich foods. The previous chapter showed that both populations are enriched with *Prevotella,* which indicates a high intake of carbohydrates-rich foods. Therefore, we predict that there will be no variation in the carbohydrate metabolism, due to the dominance of *Prevotella,* we predict that there may be high functional redundancy in the carbohydrate metabolisms between the two groups. Additionally, amino acids, bile acids, and lipids metabolism are predicted to be enriched in an urban cohort. This is because the urban population had significantly higher *Bacteroides* abundance analysis (Figure 3.6, chapter 3). A study by Ou and colleagues showed relatively higher counts of *Bacteroides* in subjects consuming diet high in animal protein and fats (Ou et al., 2013). Although the rural cohort maintains a predominately pastoral lifestyle, these individuals may have access to some processed foods. However, survey responses from the rural volunteers indicated generally low consumption of fast-foods on average compared to the urban cohort (**Table 3.1**). Fast foods or western diets typically consumed in urban areas are distinguished by high saturated- and trans-fats and low mono- and polyunsaturated fats (Singh et al., 2017). Several studies have demonstrated that a high fat diet increases the proportion of *Bacteroides* and other anaerobic microflora (Fava et al., 2013; Wu et al., 2011). Therefore, lipids metabolism and related pathways is predicted to be higher in urban samples.

4.2 Results

In the previous chapter, substantial inter-population variability in the gut microbiome composition was shown. To investigate whether the gut microbiomes of rural and urban South African populations have different functional capacity, we used the PICRUSt online software tool (Langille et al., 2013). The inferred functional analysis of the gut bacterial communities revealed several differences in gene abundance between the two populations. The KEGG pathway predictions revealed a high abundance of functional potential including metabolism, human diseases, environmental information processing, and genetic information processing (**Table 4.1**).

At level II KEGG, several differences in the predicted functional pathways between the two groups were observed (**Figure 4.1**). There was a low abundance of genes involved in the metabolism of carbohydrates in the rural cohort compared to urban. The lipids, amino acids, and bile acids metabolism-related genes were significantly enriched in the urban population. Other metabolic processes that were abundant in the urban cohort included glycan biosynthesis and biosynthesis of other secondary metabolites.

Table 4.3 Level 1 KEGG pathways as annotated by PICRUSt. Significant differences (Welch test; p-value < 0.05) in abundances were observed between the two populations.

Level_1 KEGG pathways	***Rural: mean rel. freq. (%)***	***Urban: mean rel. freq. (%)***	***p-values***
Cellular Processes	3.42021527411	3.18321166353	0.0595024135382
Cellular Processes and Signalling	4.04116068588	4.12859455229	0.0201829294729
Environmental Information Processing	14.1093535473	13.488192414	0.00470093555612
Genetic Information Processing	23.3506973617	23.1238660827	0.138340717064
Human Diseases	0.689882567621	0.69643698148	0.385411789044
Metabolism	48.8122220687	49.7833369749	4.37274879351e-05
Organismal Systems	0.784478070044	0.788954698452	0.742157061912
Poorly Characterized	4.79199042463	4.80740663265	0.40917008553

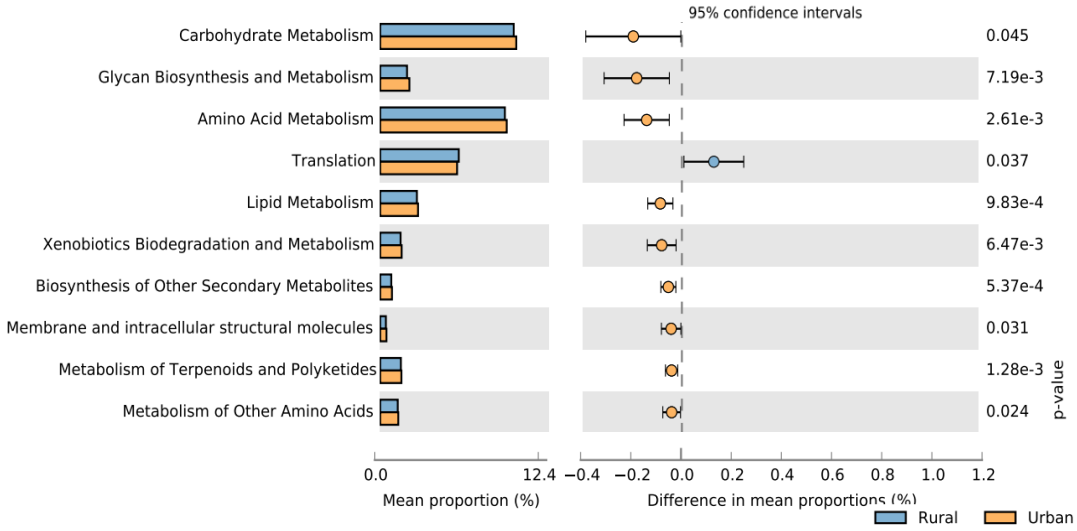


Figure 4.1 Predicted functional pathways of rural and urban populations inferred in KEGG pathways database from 16S rRNA gene sequences using PICRUSt and visualized in STAMP. The predicted functional pathways were significantly enriched in urban population compared to rural group, (Welch test; $P \leq 0.05$).

To further characterize some of the functional differences, additional analysis was carried out at level 3 KEGG pathways. Overall, 83 pathways were identified and showed significant differences between the populations (**Annexure 1**). Of these, 71 pathways (85%) were more abundant in the urban population, including several pathways associated with protein digestion and absorption, Tricarboxylic acid cycle (TCA - cycle) and pyruvate pathways. Other carbohydrate related pathways, such as pentose-glucuronate interconversions, galactose, pentose-phosphate pathway, and starch and sucrose metabolism, showed no significant difference between the two populations. However, amino sugar and nucleotide sugar metabolism were enriched in the urban cohort ($p = 0.03$). The pathways related to lipid metabolism were differentially distributed. Fatty acid elongation ($p = 0.02$), steroid biosynthesis ($p = 0.03$), secondary bile acids biosynthesis ($p = 0.02$), sphingolipids metabolism ($p = 1.69e-3$) and biosynthesis of unsaturated fatty acids ($p = 0.05$) were significantly enriched in urban population. Amino acids metabolic pathways such as phenylalanine and tyrosine metabolism were significantly enriched in the urban group ($p = 0.02$ and $p = 4.49e-3$, respectively) while tryptophan metabolism showed no differences between the populations.

Results showed that both populations had several pathways that are related to xenobiotic metabolism (**Annexure 2 and 3**). This included xenobiotics pathways which are metabolized by the cytochrome p450 family, such as benzoate and aminobenzoate. Other xenobiotic degradation pathways include atrazine, xylene, and ethylbenzene. While xenobiotic degradation metabolism was significantly higher for KEGG pathways assigned at level 1 in urban population, level 3 assigned pathways did not show significant difference between the two populations. In general, very few antibiotic synthesis pathways were observed including; streptomycin and vancomycin biosynthesis. All these pathways were higher in the urban cohort, ($p = 4.68e-4$, $p = 1.08e-3$, respectively; Annexure 2).

4.3 Discussion

Culture-independent 16S rRNA sequencing provides an essential overview regarding potential biological functions of bacterial communities. Here, 16S rRNA analysis of gut microbiomes from the South African rural and urban populations suggests enrichment and differences in various functional pathways including metabolic, genetics and environmental information processing that aligns with the dietary and environmental factors experienced by these communities. The results show that the gut microbial taxa of both populations were related to many physiological functions (carbohydrate, amino acid, and glycan metabolism). These functions may contribute to the host's wellbeing. For example, glycans have diverse biological functions which are distinct from nucleic acids and proteins, such as the maintenance of cell or tissue structure, molecular signal transduction, and cell recognition (Varki, 2016).

In contrast to our hypothesis, carbohydrate metabolism was significantly higher in the urban population. Predicted carbohydrate pathways suggests that the gut microbiota of urban individuals is more active with simple sugars pathways such as galactose, sucrose and no complex carbohydrates digestion pathways. These results agree with the dietary record investigated in this study, where the urban volunteers reported a higher consumption of fast-

foods and processed foods known to contain high amounts of added simple sugars (Sanders, 2016). The abundance of pathways involved in simple sugar metabolism has been reported in western diet-associated gut microbiota of the obesity-induced mice (Turnbaugh et al., 2008). Lower activity of carbohydrates metabolism within the rural gut microbiota differs from other rural African populations (De Filippo et al., 2010, Schnorr et al., 2014) with pathways involved in complex carbohydrates digestion or resistant starch. Amino acids, bile acids, and lipids metabolisms were enriched in the urban, reflecting a high intake of diets rich animal protein and fats. These findings agree with the initial hypothesis. Comparable results were reported in previous studies (O'keefe et al., 2007; Ou et al., 2013) where the urban population had high genes for secondary bile acid metabolism which was associated with animal-based diet. Although secondary bile acids biosynthesis pathway has a key function in cholesterol homeostasis, the resulting metabolites are also considered potentially carcinogenic (Vipperla & O'Keefe, 2012; O'Keefe et al., 2015).

The increased abundance of xenobiotic degradation pathways in both rural and urban population is noteworthy. In an ecosystem or niche, pathways associated with xenobiotic biodegradation and metabolism play a role in bioremediation (Singleton, 1994, Das et al., 2016). The presence of bacterial genes related to xenobiotic biodegradation pathways indicates the adaptation of the gut microbial community in the two populations to degrade environmental contaminants or toxins. Although xenobiotic metabolism is typically reported in natural environments (Zhou et al., 2016) and other animals (Hu et al., 2018), humans are exposed to numerous xenobiotics, and the majority of these are in the form of pharmaceuticals (Haiser & Turnbaugh, 2013). For instance, both populations exhibited higher abundances of pathways involving cytochrome P450. Cytochrome P450 enzymes metabolize external substances such as chemicals and drugs/medications which are ingested, as well as internal toxins substances that are formed within cells (Lynch & Price, 2007). These observations are consistent with the impact the metabolism of drugs including antibiotics has on the intestinal bacteria (Saad et al., 2012; Haiser and Turnbaugh, 2013). Additionally, these findings may

support the fact that both populations have easier access to therapeutic drugs/medications, which is a signature of industrialized lifestyles even in the rural population.

4.4 Conclusion

A detailed analysis of the functional attributes of the gut bacterial community of healthy rural and urban South Africans has been performed. Predicted functional pathways suggests that urban microbiota is enriched and metabolically more active with, carbohydrates, amino acids, lipids and bile acid biosynthesis, while rural gut microbiota had lower activity of these metabolisms. An increase in dietary diversity and in protein and lipid intake in the urban population could have caused an increase of the metabolic capabilities towards degradation of a wider range of carbohydrates and utilization of a higher number of available amino acids and lipids. Although the abundance of metabolic pathways differed between two populations, there were no predicted functional metabolic pathway specific or unique to one population despite difference in geographic location. Suggesting that the urban and rural diet could be similar.

Chapter 5

Summarizing research findings and providing future perspectives

Summary of the dissertation and future perspectives

Taken together, this exploratory study describes gut bacterial community data from an understudied population thus, providing a starting point for further comparative work on structure and functional potential of the gut microbiota of healthy South African population. A detailed analysis of the gut microbiota of healthy South Africans clearly reflects *Firmicutes* dominated gut microbiota in two populations living in two distinct geographical locations. Further analysis revealed distinct microbial signatures and a higher degree of Heterogeneity between populations, however, without significant differences in species diversity within the two populations. This study also provided notable evidence that there was no significant difference in species diversity between the two populations. It can be speculated that the lack of species diversity could be due to less genetic variation between the two populations. However, this speculation serves as a basis for further gut bacterial community research. It is important to pursue further research in this area with the use of additional populations or group either Xhosas, Zulus or Tsongas to improve our understanding on these observations.

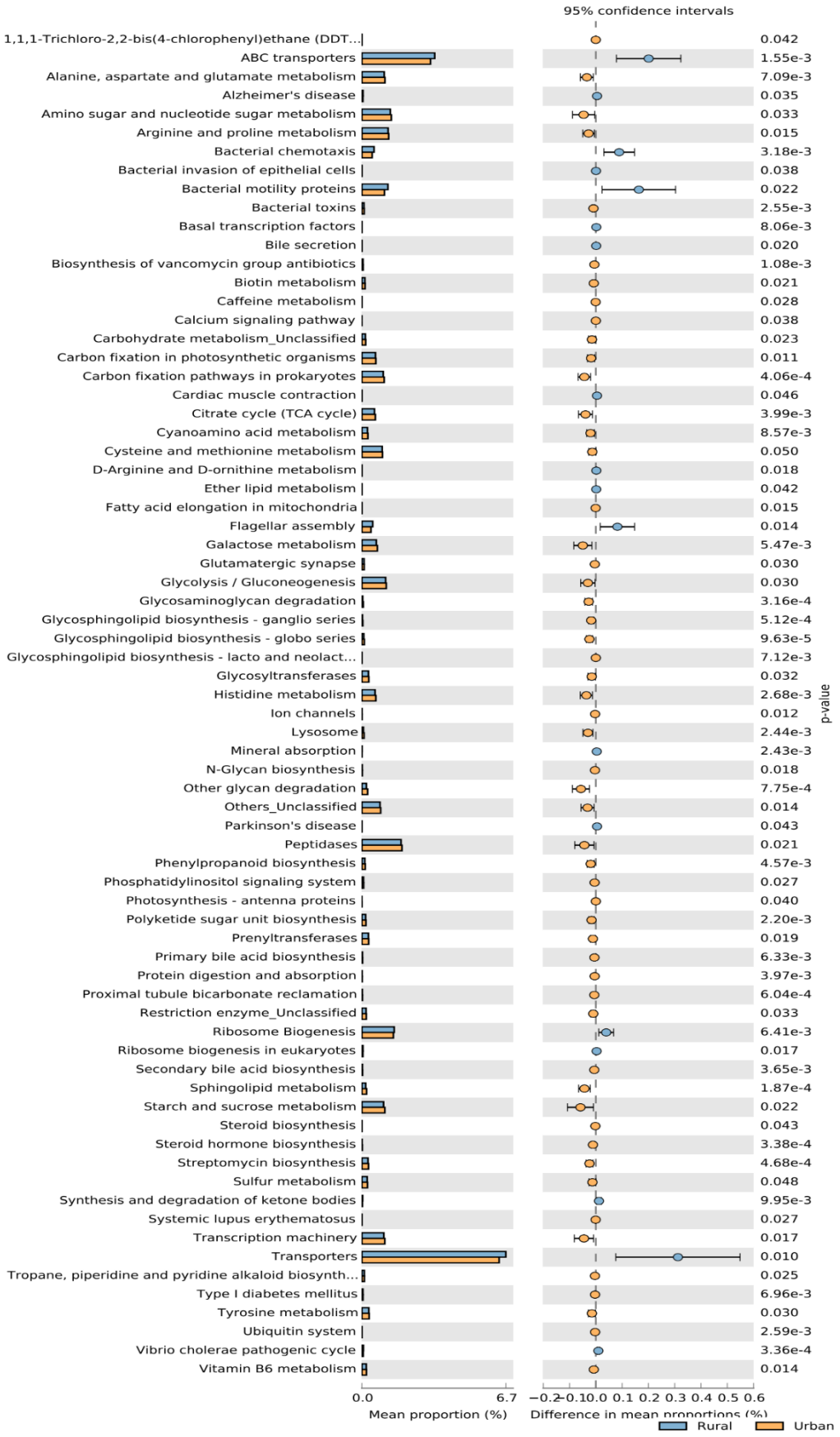
Moreover, variability of gut bacterial communities within and between populations could be due to differing dietary habits based on lifestyle and demographic factors. This is especially evident with phyla *Spirochete* that is only present in rural population. The comparison between rural and urban populations gut microbiota allowed the identification of taxa that seem to have been lost (phyla *Spirochete* and genus *Treponema*) in the urban corhot, possibly as a result of dietary variance or changes. In this context, the increase in dietary diversity and in protein and lipid intake in the urban population could have caused an increase of the metabolic capabilities towards degradation of a wider range of carbohydrates and utilization of a higher number of available amino acids and lipids.

The study of external influences on gut microbiota composition and function is especially pertinent because the human living environment is becoming rapidly urbanized. Such changes may interrupt the healthy development of the microbial community and increase risk of associated diseases, such as inflammatory bowel diseases. Moving forward, gut microbiota studies must be incorporated into a broader framework of environmental exposure, for a thorough understanding of how external factors contribute to gut microbial community assembly and affect the quality of human health.

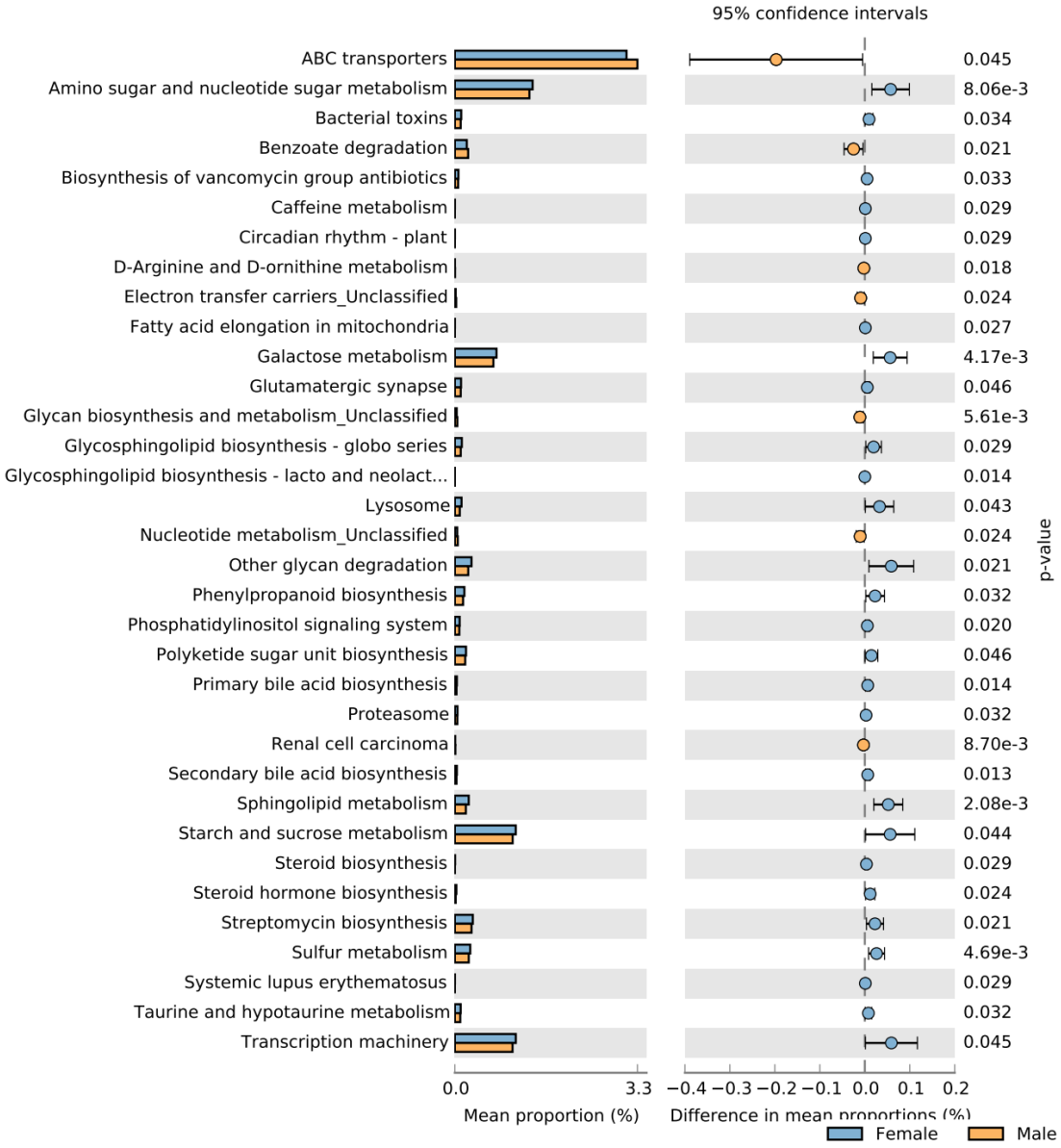
Limitations of the study

The predictive nature of PICRUSt as a computational tool is limited by the accuracy and comprehensiveness of the database of reference genomes (Langille et al., 2013). Environments which are less extensively covered in the reference database, such as the human gut microbiome, are less accurately described by predictive analysis (Nagpal et al., 2016). However, the use of PICRUSt as an affordable approach prior to more expensive 'omics analysis is suitable due to the thousands of uncultivated microbial communities for which only marker gene surveys are currently available (Petri et al., 2017).

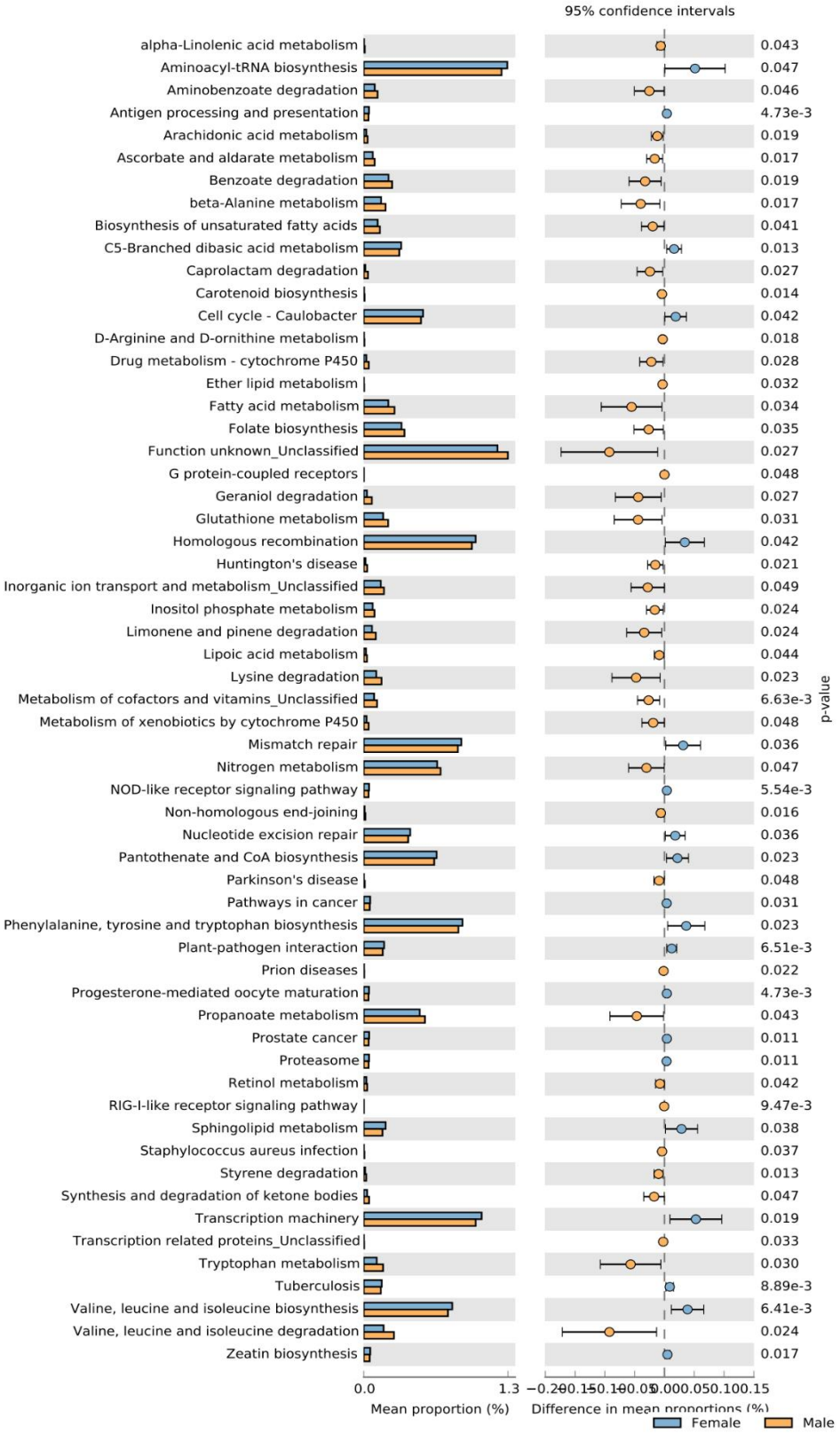
Annexures



Annexure 1: Level 3 KEGG pathways. A total of 83 various metabolic pathways that showed significant variations between the two cohorts (Welch test, $p < 0.05$).



Annexure 2: Level 3 KEGG pathways for the urban population. A total of 34 various metabolic pathways that showed significant variations within genders (Welch test, $p < 0.05$).



Annexure 3: Level 3 KEGG pathways for the rural population. A total of 34 various metabolic pathways that showed significant variations within genders (Welch test, $p < 0.05$).

References

ABUBUCKER, S., SEGATA, N., GOLL, J., SCHUBERT, A. M., IZARD, J., CANTAREL, B. L., HUTTENHOWER, C. (2012). Metabolic Reconstruction for Metagenomic Data and Its Application to the Human Microbiome. *PLOS Computational Biology*, *8*(6),

ALLABAND, C., MCDONALD, D., VAZQUENZ-BAEZA, Y., MINICH, J. J, TRIPATHI, A., BRENNER, D. A., & DORRESTEIN, P. (2018). Microbiome 101: Studying, Analyzing, and Interpreting Gut Microbiome Data for Clinicians. *Clincal Gastroenterology and Hepatology*. *17*(2), 218-230.

ARUMUGAM, M., RAES, J., PELLETIER, E., LE PASLIER, D., YAMADA, T., MENDE, D. R., BORK, P. (2011). Enterotypes of the human gut microbiome. *Nature, 473*(7346), 174-180.

AYENI, F. A., BIAGI, E., RAMPELLI, S., FIORI, J., SOVERINI, M., AUDU, H. J., CARELLI, V. (2018). Infant and Adult Gut Microbiome and Metabolome in Rural Bassa and Urban Settlers from Nigeria. *Cell Reports*, 23(10), 3056-3067.

BACKHED, F., ROSWALL, J., PENG, Y., FENG, Q., JIA, H., KOVATCHEVA-DATCHARY, P., WANG, J. (2015). Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe*, *17*(5), 690-703.

BAILEY, M. T., & COE, C. L. (1999). Maternal separation disrupts the integrity of the intestinal microflora in infant rhesus monkeys. *Developmental Psychobiology, 35*(2), 146–155.

BARBARA, G., FEINLE-BISSET, C., GHOSHAL, U. C., QUIGLEY, E. M., SANTOS, J., VANNER, S., ZOETENDAL, E. G. (2016). The Intestinal Microenvironment and Functional Gastrointestinal Disorders. *Gastroenterology, 150*(6), 1305-1318.

BARBERÁN, A., BATES, S. T., CASAMAYOR, E. O., & FIERER, N. (2014). Using network analysis to explore co-occurrence patterns in soil microbial communities. *The ISME Journal, 8*(4), 952.

BARCENILLA, A., PRYDE, S. E., MARTIN, J. C., DUNCAN, S. H., STEWART, C. S., & HENDERSON, C. (2000). Phylogenetic relationships of butyrate-producing bacteria from the human gut. *Applied and Environmental Microbiology*, *66*(4), 1654-1651.

BARLOW, G. M., YU, A., & MATHUR, R. (2015). Role of the gut microbiome in obesity and diabetes mellitus. *Nutrition Clinical Practice*, *30*.

BAUMLER, A. J., & SPERANDIO, V. (2016). Interactions between the microbiota and pathogenic bacteria in the gut. *Nature, 535*(7610), 85-93.

BELIZARIO, J. E., & NAPOLITANO, M. (2015). Human microbiomes and their roles in dysbiosis, common diseases, and novel therapeutic approaches. *Frontiers in Microbiology, 6*, 1050.

BENJAMIN, J. L., HEDIN, C. R., KOUTSOUMPAS, A., NG, S. C., MCCARTHY, N. E., PRESCOTT, N. J., WHELAN, K. (2012). Smokers with active Crohn's disease have a clinically relevant dysbiosis of the gastrointestinal microbiota. *Inflammatry Bowel Disorders*, *18*(6), 1092-1100.

BENÍTEZ-PÁEZ, A., DEL PULGAR, E. M. G., KJØLBÆK, L., BRAHE, L. K., ASTRUP, A., LARSEN, L., & SANZ, Y. (2016). Impact of dietary fiber and fat on gut microbiota re-modeling and metabolic health. *Trends in Food Science & Technology, 57*, 201-212.

BERTONI, A. G., BURKE, G. L., OWUSU, J. A., CARNETHON, M. R., VAIDYA, D., BARR, R. G., ROTTER, J. I. (2010). Inflammation and the incidence of type 2 diabetes: the Multi-Ethnic Study of Atherosclerosis (MESA). *Diabetes Care, 33*.

BIAGI, E., NYLUND, L., CANDELA, M., OSTAN, R., BUCCI, L., PINI, E., FRANCESCHI, C. (2010). Through ageing, and beyond: gut microbiota and inflammatory status in seniors and centenarians. *PLoS One, 5*(5), e10667.

BIK, E. M. (2009). Composition and function of the human-associated microbiota. *Nutritional Reviews, 67 Suppl 2*, S164-171.

BLASER, M. J. (2016). Antibiotic use and its consequences for the normal microbiome. *Science, 352*(6285), 544.

BODELIER, P. L. (2011). Toward understanding, managing, and protecting microbial ecosystems. *Frontiers in Microbiology, 2*.

BONDER, M. J., KURILSHIKOV, A., TIGCHELAAR, E. F., MUJAGIC, Z., IMHANN, F., VILA, A. V., ZHERNAKOVA, A. (2016). The effect of host genetics on the gut microbiome. *Nature Genetics, 48*(11), 1407-1412.

BOOTH, F. W., ROBERTS, C. K., & LAYE, M. J. (2012). Lack of exercise is a major cause of chronic diseases. *Comprehensive Physiology, 2*(2), 1143-1211.

BREITBART, M. et al. Metagenomic analyses of an uncultured viral community from human feces. (2003). *Journal of Bacteriology*, 185, 6220–6223

BROOKS, J. P., EDWARDS, D. J., HARWICH, M. D., RIVERA, M. C., FETTWEIS, J. M., SERRANO, M. G., GIRERD, P. (2015). The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiology, 15*(1), 66.

BROWNE HP, FORSTER SC, ANONYE BO, KUMAR N, NEVILLE BA, STARES MD, GOULDING D, LAWLEY TD. (2016). Culturing of ‘unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature,533*(7604):543–546.

BULL, M. J., & PLUMMER, N. T. (2014). Part 1: The Human Gut Microbiome in Health and Disease. Integrative Medicine: *A Clinician's Journal*, *13(*6), 17-22.

CAMPBELL, S. C., & WISNIEWSKI, P. J. (2017). Exercise is a novel promoter of intestinal health and microbial diversity. *Exercise and Sport Sciences Reviews*, 45(1), 41-47.

CÂNDIDO, F. G., VALENTE, F. X., GRZEŚKOWIAK, Ł. M., MOREIRA, A. P. B., ROCHA, D. M. U. P., & ALFENAS, R. D. C. G. (2018). Impact of dietary fat on gut microbiota and low-grade systemic inflammation: mechanisms and clinical implications on obesity. *International Journal of Food Sciences and Nutrition, 69*(2), 125-143.

CAO, X., ZHAO, D., XU, H., HUANG, R., ZENG, J., & YU, Z. (2018). Heterogeneity of interactions of microbial communities in regions of Taihu Lake with different nutrient loadings: *A network analysis. Scientific Reports, 8*(1), 8890.

CAPORASO, J. G., KUCZYNSKI, J., STOMBAUGH, J., BITTINGER, K., BUSHMAN, F. D., COSTELLO, E. K., KNIGHT, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods, 7*(5), 335-336

CAPORASO, J. G., LAUBER, C. L., WALTERS, W. A., BERG-LYONS, D., LOZUPONE, C. A., TURNBAUGH, P. J., KNIGHT, R. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proceedings of the National Academy of Sciences, 108*(supplimentary 1), 4516-4522.

CARBONETTO, B., FABBRO, M., SCIARA, M., SERAVALLE, A., MÉJICO, G., REVALE, S., et al. (2016). Human microbiota of Argentine population- a pilot study. *Frontiers in Microbiology. 7*:51

CARDING, S., VERBEKE, K., VIPOND, D. T., CORFE, B. M., & OWEN, L. J. (2015). Dysbiosis of the gut microbiota in disease. *Microbial Ecology in Health and Disease, 26*(1), 26191.

CHO, I., & BLASER, M. J. (2012). The human microbiome: at the interface of health and disease. *Nature Reviews Genetics, 13*(4), 260-270.

CHO, I., YAMANISHI, S., COX, L., METHE, B. A., ZAVADIL, J., LI, K., BLASER, M. J. (2012). Antibiotics in early life alter the murine colonic microbiome and adiposity. *Nature, 488*(7413), 621-626.

CLARKE, J. M., TOPPING, D. L., CHRISTOPHERSEN, C. T., BIRD, A. R., LANGE, K., SAUNDERS, I., & COBIAC, L. (2011). Butyrate esterified to starch is released in the human gastrointestinal tract. *The American Journal of Clinical Nutrition, 94*(5), 1276-1283.

CLEMENTE, J. C., PEHRSSON, E. C., BLASER, M. J., SANDHU, K., GAO, Z., & WANG, B. (2015). The microbiome of uncontacted Amerindians. *Science Advances, 1*(3), e1500183.

CLOONEY, A. G., FOUHY, F., SLEATOR, R. D., O'DRISCOLL, A., STANTON, C., COTTER, P. D., & CLAESSON, M. J. (2016). Comparing apples and oranges?: next generation sequencing and its impact on microbiome analysis. *PLoS One, 11*(2), e0148028.

COLE JR, WANG Q, CARDENAS E, FISH J, CHAI B, et al. (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Research 37*: D141–145.

CONLON, M. A., & BIRD, A. R. (2014). The impact of diet and lifestyle on gut microbiota and human health. *Nutrients, 7*(1), 17-44.

CONSORTIUM, H. M. J. R. S. (2010). A catalog of reference genomes from the human microbiome. *Science, 328*(5981), 994-999.

COX, I. J., MCPHAIL, M. J., & WILLIAMS, R. (2019). Assessing Liver Disease and the Gut Microbiome by NMR Metabolic Profiling of Body Fluids. *Applications of NMR Spectroscopy: 7*, 7, 40.

DAS, A., SRINIVASAN, M., GHOSH, T. S., & MANDE, S. S. (2016). Xenobiotic Metabolism and Gut Microbiomes. *PLoS One, 11*(10),

DAVID, L. A., MAURICE, C. F., CARMODY, R. N., GOOTENBERG, D. B., BUTTON, J. E., WOLFE, B. E., TURNBAUGH, P. J. (2014). Diet rapidly and reproducibly alters the human gut microbiome. *Nature, 505*(7484), 559-563.

DE FILIPPIS, F., & ERCOLINI, D. (2018). Microbiome and Diet. In The Gut Microbiome in Health and Disease. *Springer* (pp. 79-88)

DE FILIPPO, C., CAVALIERI, D., DI PAOLA, M., RAMAZZOTTI, M., POULLET, J. B., MASSART, S., 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, 107*(33), 14691-14696.

DE FILIPPO, C., DI PAOLA, M., RAMAZZOTTI, M., ALBANESE, D., PIERACCINI, G., BANCI, E., LIONETTI, P. (2017). Diet, environments, and gut microbiota. A preliminary investigation in children living in rural and urban Burkina Faso and Italy. *Frontiers in Microbiology*, 8, 1979.

DEHINGIA, M., DEVI, K. T., TALUKDAR, N. C., TALUKDAR, R., REDDY, N., MANDE, S. S., KHAN, M. R. (2015). Gut bacterial diversity of the tribes of India and comparison with the worldwide data. *Scientific Reports, 5,* 18563.

DeSANTIS TZ, HUGENHOLTZ P, LARSEN N, ROJAS M, BRODIE EL, et al. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology 72*: 5069–5072

DETHLEFSEN L, ECKBURG PB, BIK EM, RELMAN DA. Assembly of the human intestinal microbiota. (2006). *Trends in Ecology and Evolution, 21*:517–523.

DHARIWAL, A., CHONG, J., HABIB, S., KING, I. L., AGELLON, L. B., & XIA, J. (2017). MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. *Nucleic Acids Research, 45*(W1), W180-W188.

DICKSVED, J. (2008). Exploring the human intestinal microbiome in health and disease (Vol. 2008, No. 30)

DOMINGUEZ-BELLO, M. G., BLASER, M. J., LEY, R. E., & KNIGHT, R. (2011). Development of the human gastrointestinal microbiota and insights from high-throughput sequencing. *Gastroenterology, 140*(6), 1713-1719.

DOMINGUEZ-BELLO, M. G., COSTELLO, E. K., CONTRERAS, M., MAGRIS, M., HIDALGO, G., FIERER, N., & KNIGHT, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proceedings of the National Academy of Sciences, 107*(26), 11971-11975.

DONALDSON, G. P., LEE, S. M., & MAZMANIAN, S. K. (2016). Gut biogeography of the bacterial microbiota. *Nature Reviews. Microbiology, 14*(1), 20-32.

DUNCAN, S. H., LOBLEY, G. E., HOLTROP, G., INCE, J., JOHNSTONE, A. M., LOUIS, P., & FLINT, H. J. (2008). Human colonic microbiota associated with diet, obesity and weight loss. *International Journal of Obesity (Lond), 32*(11), 1720-1724.

ECKBURG, P. B., BIK, E. M., BERNSTEIN, C. N., PURDOM, E., DETHLEFSEN, L., SARGENT, M., RELMAN, D. A. (2005). Diversity of the human intestinal microbial flora. *Science, 308*(5728), 1635-1638.

EDGAR, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics, 26*.

EHRLICH, S. D., & CONSORTIUM, M. (2011). MetaHIT: The European Union Project on metagenomics of the human intestinal tract. In Metagenomics of the human body. *Springer,* (pp. 307-316)

EVEILLARD, D., BOUSKILL, N. J., VINTACHE, D., GRAS, J., WARD, B. B., & BOURDON, J. (2019). Probabilistic Modeling of Microbial Metabolic Networks for Integrating Partial Quantitative Knowledge Within the Nitrogen Cycle. *Frontiers in Microbiology, 9*(3298).

FALKOWSKI, P. G., FENCHEL, T., AND DELONG,E. F. (2008). The microbial engines that drive earth's biogeochemichal cycles. *Science 320*, 1034-1039

FAN, W., HUO, G., LI, X., YANG, L., DUAN, C., WANG, T., & CHEN, J. (2013). Diversity of the intestinal microbiota in different patterns of feeding infants by Illumina high-throughput sequencing. *World Journal of Microbiology and Biotechnology, 29*(12), 2365-2372.

FAUST, K., SATHIRAPONGSASUTI, J. F., IZARD, J., SEGATA, N., GEVERS, D., RAES, J., & HUTTENHOWER, C. (2012). Microbial Co-occurrence Relationships in the Human Microbiome. *PLOS Computational Biology, 8*(7),

FAVA, F., GITAU, R., GRIFFIN, B., GIBSON, G., TUOHY, K., & LOVEGROVE, J. (2013). The type and quantity of dietary fat and carbohydrate alter faecal microbiome and short-

chain fatty acid excretion in a metabolic syndrome ‘at-risk’population. *International Journal of Obesity, 37*(2), 216.

FLINT, H. J., SCOTT, K. P., LOUIS, P., & DUNCAN, S. H. (2012). The role of the gut microbiota in nutrition and health. *Nature Reviews of Gastroenterology and Hepatoogyl, 9*(10), 577-589.

FRANZOSA EA, HSU T, SIROTA-MADI A, SHAFQUAT A, ABU-ALI G, MORGAN XC, HUTTENHOWER C. (2015). Sequencing and beyond: Integrating molecular ‘omics' for microbial community profiling. *Nature Reviews of Microbiology,13*(6):360–372

FUJIO-VEJAR, S., VASQUEZ, Y., MORALES, P., MAGNE, F., VERA-WOLF, P., UGALDE, J. A., GOTTELAND, M. (2017). The gut microbiota of healthy Chilean subjects reveals a high abundance of the phylum *Verrucomicrobia*. *Frontiers in Microbiology, 8*, 1221.

GACI, N., BORREL, G., TOTTEY, W., O’TOOLE, P. W., & BRUGÈRE, J.-F. (2014). Archaea and the human gut: new beginning of an old story. *World Journal of Gastroenterology*, *20*(43), 16062.

GERASIMIDIS, K., BERTZ, M., QUINCE, C., BRUNNER, K., BRUCE, A., COMBET, E., IJAZ, U. Z. (2016). The effect of DNA extraction methodology on gut microbiota research applications. *BMC research notes, 9*(1), 365.

GOMEZ, A., PETRZELKOVA, K. J., BURNS, M. B., YEOMAN, C. J., AMATO, K. R., VLCKOVA, K., BLEKHMAN, R. (2016). Gut Microbiome of Coexisting BaAka Pygmies and Bantu Reflects Gradients of Traditional Subsistence Patterns. *Cell Report, 14*(9), 2142-2153.

GOODRICH, J. K., DAVENPORT, E. R., BEAUMONT, M., JACKSON, M. A., KNIGHT, R., OBER, C., LEY, R. E. (2016). Genetic determinants of the gut microbiome in UK twins. *Cell Host & Microbe, 19*(5), 731-743.

GOODRICH, J. K., WATERS, J. L., POOLE, A. C., SUTTER, J. L., KOREN, O., BLEKHMAN, R., LEY, R. E. (2014). Human genetics shape the gut microbiome. *Cell, 159*(4), 789-799.

GRINE, G., BOUALAM, M., & DRANCOURT, M. (2017). *Methanobrevibacter smithii*, a methanogen consistently colonising the newborn stomach. *European Journal of Clinical Microbiology & Infectious Diseases, 36*(12), 2449-2455.

GUARALDI, F., & SALVATORI, G. (2012). Effect of breast and formula feeding on gut microbiota shaping in newborns. *Frontiers in Cellular and Infection Microbiology, 2,* 94.

GUARNER, F., & MALAGELADA, J.-R. (2003). Gut flora in health and disease. *The Lancet, 361*(9356), 512-519.

GUPTA, V. K., PAUL, S., & DUTTA, C. (2017). Geography, ethnicity or subsistence-specific variations in human microbiome composition and diversity. *Frontiers in Microbiology, 8,* 1162.

HAGLAND, H. R., & SØREIDE, K. (2015). Cellular metabolism in colorectal carcinogenesis: Influence of lifestyle, gut microbiome and metabolic pathways. *Cancer Letters, 356*(2), 273-280.

HAISER, H. J., & TURNBAUGH, P. J. (2013). Developing a metagenomic view of xenobiotic metabolism. *Pharmacological Research, 69*(1), 21-31.

HANSEN, E. E., LOZUPONE, C. A., REY, F. E., WU, M., GURUGE, J. L., NARRA, A., GORDON, J. I. (2011). Pan-genome of the dominant human gut-associated archaeon, and *Methanobrevibacter smithii* studied in twins. *Proceedings of the National Academy of Sciences, 108*(1), 4599.

HANSEN, M. E. B., RUBEL, M. A., BAILEY, A. G., RANCIARO, A., THOMPSON, S. R., CAMPBELL, M. C., TISHKOFF, S. A. (2019). Population structure of human gut bacteria in a diverse cohort from rural Tanzania and Botswana. *Genome Biology, 20*(1), 16.

HAYASHI, H., SAKAMOTO, M., KITAHARA, M. & BENNO, Y. Molecular analysis of fecal microbiota in elderly individuals using 16S rDNA library and T-RFLP. (2003). *Microbioogy and Immunology, 47,* 557–570.

HENTGES, D. J., MAIER, B. R., BURTON, G. C., FLYNN, M. A., & TSUTAKAWA, R. K. (1977). Effect of a high-beef diet on the fecal bacterial flora of humans. *Cancer Research, 37*(2), 568-571.

HIIPPALA, K., KAINULAINEN, V., KALLIOMÄKI, M., ARKKILA, P., & SATOKARI, R. (2016). Mucosal prevalence and interactions with the epithelium indicate commensalism of *Sutterella spp*. *Frontiers in Microbiology, 7*, 1706-1706.

HOFFMANN, C., DOLLIVE, S., GRUNBERG, S., CHEN, J., LI, H., WU, G. D., BUSHMAN, F. D. (2013). Archaea and Fungi of the Human Gut Microbiome: Correlations with Diet and Bacterial Residents. *PLoS One, 8*(6)

HOLD, G. L., PRYDE, S. E., RUSSELL, V. J., FURRIE, E., & FLINT, H. J. (2002). Assessment of microbial diversity in human colonic samples by 16S rDNA sequence analysis. *FEMS Microbiology Ecology*, 39.

HONGFEI CUI, Y. L., XUEGONG ZHANG. (2016). An overview of major metagenomic studies on human microbiomes in health and disease. *Quantitative Biology, 4*(3), 192-206.

HU, Z., CHEN, X., CHANG, J., YU, J., TONG, Q., LI, S., & NIU, H. (2018). Compositional and predicted functional analysis of the gut microbiota of Radix auricularia (Linnaeus) via high-throughput Illumina sequencing. *Peer Journal, 6,* e5537

HUGON, P., DUFOUR, J.-C., COLSON, P., FOURNIER, P.-E., SALLAH, K., RAOULT, D. (2015) A comprehensive repertoire of prokaryotic species identified in human beings. *Lancet Infectious Diseases. 15*, 1211–1219

HUMAN MICROBIOME PROJECT, C. (2012). Structure, function and diversity of the healthy human microbiome. *Nature, 486*(7402), 207-214.

HUTTENHOWER, C., GEVERS, D., KNIGHT, R., ABUBUCKER, S., BADGER, J. H., CHINWALLA, A. T., FULTON, R. S. (2012). Structure, function and diversity of the healthy human microbiome. *Nature, 486*(7402), 207.

INGHAM, C. J., SPRENKELS, A., BOMER, J., MOLENAAR, D., VAN DEN BERG, A., VAN HYLCKAMA VLIEG, J. E., & DE VOS, W. M. (2007). The micro-Petri dish, a million-well growth chip for the culture and high-throughput screening of microorganisms. *Proceedings of the National Academy of Sciences*, 104(46), 18217-18222.

JACOBS, D. M., DELTIMPLE, N., VAN VELZEN, E., VAN DORSTEN, F. A., BINGHAM, M., VAUGHAN, E. E., & VAN DUYNHOVEN, J. (2008). 1H NMR metabolite profiling of feces as a tool to assess the impact of nutrition on the human microbiome. NMR in Biomedicine: *An International Journal Devoted to the Development and Application of Magnetic Resonance In vivo, 21*(6), 615-626.

JAKOBSSON, H. E., JERNBERG, C., ANDERSSON, A. F., SJÖLUND-KARLSSON, M., JANSSON, J. K., & ENGSTRAND, L. (2010). Short-Term Antibiotic Treatment Has Differing Long-Term Impacts on the Human Throat and Gut Microbiome. *PLoS One, 5*(3)

JERNBERG, C., LÖFMARK, S., EDLUND, C., & JANSSON, J. K. (2007). Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. *The ISME Journal, 1*(1), 56.

JERNBERG, C., SULLIVAN, Å., EDLUND, C., & JANSSON, J. K. (2005). Monitoring of antibiotic-induced alterations in the human intestinal microflora and detection of probiotic strains by use of terminal restriction fragment length polymorphism. *Applied and Environmental Microbiology, 71*(1), 501-506.

JI, B., & NIELSEN, J. (2015). From next-generation sequencing to systematic modeling of the gut microbiome. *Frontiers in Genetics, 6*, 219.

JOST, T., LACROIX, C., BRAEGGER, C., & CHASSARD, C. (2013). Assessment of bacterial diversity in breast milk using culture-dependent and culture-independent approaches. *British Journal of Nutrition, 110*(7), 1253-1262.

JOVEL, J., PATTERSON, J., WANG, W., HOTTE, N., O'KEEFE, S., MITCHEL, T., WONG, G. K. (2016). Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. *Frontiers in Microbiology, 7*, 459.

KABEERDOSS, J., DEVI, R. S., MARY, R. R., & RAMAKRISHNA, B. S. (2012). Faecal microbiota composition in vegetarians: comparison with omnivores in a cohort of young women in southern India. *British Journal of Nutrition, 108*(6), 953-957.

KANEHISA, M., & GOTO, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research, 28*(1), 27-30.

KARLSSON, C. L., ÖNNERFÄLT, J., XU, J., MOLIN, G., AHRNÉ, S., & THORNGREN-JERNECK, K. (2012). The microbiota of the gut in preschool children with normal and excessive body weight. *Obesity, 20*(11), 2257-2261.

KEIM, N. L., & MARTIN, R. J. (2014). Dietary whole grain–microbiota interactions: insights into mechanisms for human health. *Advances in Nutrition, 5*.

KIM, C. H., PARK, J., & KIM, M. (2014). Gut Microbiota-Derived Short-Chain Fatty Acids, T Cells, and Inflammation. *Immune Network, 14*(6), 277-288.

KIM, M. S., HWANG, S. S., PARK, E. J., & BAE, J. W. (2013). Strict vegetarian diet improves the risk factors associated with metabolic diseases by modulating gut microbiota and reducing intestinal inflammation. *Environmental Microbiology Reports, 5*(5), 765-775.

KINROSS, J. M., DARZI, A. W., & NICHOLSON, J. K. (2011). Gut microbiome-host interactions in health and disease. *Genome Medicine, 3*(3), 14.

KISUSE, J., LA-ONGKHAM, O., NAKPHAICHIT, M., THERDTATHA, P., MOMODA, R., TANAKA, M., NAKAYAMA, J. (2018). Urban diets linked to gut microbiome and metabolome alterations in children: a comparative cross-sectional study in Thailand. *Frontiers in Microbiology, 9*, 1345.

KOENIG, J. E., SPOR, A., SCALFONE, N., FRICKER, A. D., STOMBAUGH, J., KNIGHT, R., LEY, R. E. (2011). Succession of microbial consortia in the developing infant gut microbiome. *Proceedings of the National Academy in Sciences, 108*.

KOETH, R. A., WANG, Z., LEVISON, B. S., BUFFA, J. A., ORG, E., & SHEEHY, B. T. (2013). Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nature Med, 19.*

KONOPKA, A. (2009). What is microbial community ecology? *The ISME journal, 3*, 1223.

KONTUREK, P. C., HAZIRI, D., BRZOZOWSKI, T., HESS, T., HEYMAN, S., & KWIECIEN, S. (2015). Emerging role of fecal microbiota therapy in the treatment of gastrointestinal and extra-gastrointestinal diseases. *Journal of Physiology and Pharmacology, 66*.

LANGILLE, M. G., ZANEVELD, J., CAPORASO, J. G., MCDONALD, D., KNIGHTS, D., REYES, J. A., HUTTENHOWER, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature Biotechnology, 31*(9), 814-821.

LAUDADIO, I., FULCI, V., PALONE, F., STRONATI, L., CUCCHIARA, S., & CARISSIMI, C. (2018). Quantitative assessment of shotgun metagenomics and 16s rDNA amplicon sequencing in the study of human gut microbiome. OMICS: *A Journal of Integrative Biology, 22*(4), 248–254.

LEY, R. E., HAMADY, M., LOZUPONE, C., TURNBAUGH, P. J., RAMEY, R. R., BIRCHER, J. S., KNIGHT, R. (2008). Evolution of mammals and their gut microbes. *Science, 320*(5883), 1647-1651.

LI, J., JIA, H., CAI, X., ZHONG, H., FENG, Q., SUNAGAWA, S., NIELSEN, T. (2014). An integrated catalog of reference genes in the human gut microbiome. *Nature Biotechnology, 32*(8), 834.

LIMON, J. J., SKALSKI, J. H., & UNDERHILL, D. M. (2017). Commensal fungi in health and disease. *Cell Host & Microbe, 22(*2), 156-165.

LLOYD-PRICE, J., ABU-ALI, G., & HUTTENHOWER, C. (2016). The healthy human microbiome. *Genome Medicine, 8(*1), 51.

LÖFMARK, S., JERNBERG, C., JANSSON, J. K., & EDLUND, C. (2006). Clindamycin-induced enrichment and long-term persistence of resistant Bacteroides spp. and resistance genes. *Journal of Antimicrobial Chemotherapy, 58*(6), 1160-1167.

LYNCH, S. V., & PEDERSEN, O. (2016). The human intestinal microbiome in health and disease. *New England Journal of Medicine, 375*(24), 2369-2379.

LYNCH, T., & PRICE, A. (2007). The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects. *Am Fam Physician, 76*(3), 391-396.

MACFARLANE, G. T., CUMMINGS, J. H., & ALLISON, C. (1986). Protein degradation by human intestinal bacteria. *Journal of Genetics and Microbiology*, 132(6), 1647-1656.

MACKOS, A. R., VARALJAY, V. A., MALTZ, R., GUR, T. L., & BAILEY, M. T. (2016). Role of the intestinal microbiota in host responses to stressor exposure. *International Reviews of Neurobiology, 131*, 1-19.

MAIER, T. V., LUCIO, M., LEE, L. H., VERBERKMOES, N. C., BRISLAWN, C. J., BERNHARDT, J., & MORTON, J. T. (2017). Impact of dietary resistant starch on the human gut microbiome, metaproteome, and metabolome. *MBio, 8*(5), e01343-17.

MANICHANH, C., RIGOTTIER-GOIS, L., BONNAUD, E., GLOUX, K., PELLETIER, E., FRANGEUL, L., & ROCA, J. (2006). Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut, 55(2), 205-211.

MARIAT, D., FIRMESSE, O., LEVENEZ, F., GUIMARĂES, V., SOKOL, H., DORÉ, J., FURET, J. (2009). The *Firmicutes/Bacteroidetes* ratio of the human microbiota changes with age. *BMC Microbiology, 9*(1), 123.

MARQUES, T. M., WALL, R., ROSS, R. P., FITZGERALD, G. F., RYAN, C. A., & STANTON, C. (2010). Programming infant gut microbiota: influence of dietary and environmental factors. *Current Opinion in Biotechnology, 21*(2), 149-156.

MARSH, T.L. Terminal restriction fragment length polymorphism (T-RFLP): an emerging method for characterizing diversity among homologous populations of amplification products. (1999). *Current Opinion in Microbiology, 2*, 323–327

MARTINEZ, I., STEGEN, J. C., MALDONADO-GOMEZ, M. X., EREN, A. M., SIBA, P. M., GREENHILL, A. R., & WALTER, J. (2015). The gut microbiota of rural papua new guineans: composition, diversity patterns, and ecological processes. *Cell Reports, 11*(4), 527-538.

MARTÍNEZ, I., KIM, J., DUFFY, P. R., SCHLEGEL, V. L., & WALTER, J. (2010). Resistant starches types 2 and 4 have differential effects on the composition of the fecal microbiota in human subjects. *PLoS One, 5*(11), e15046-e15046.

MATIJAŠIĆ, B. B., OBERMAJER, T., LIPOGLAVŠEK, L., GRABNAR, I., AVGUŠTIN, G., & ROGELJ, I. (2014). Association of dietary type with fecal microbiota in vegetarians and omnivores in Slovenia. *European Journal of Nutrition, 53*(4), 1051-1064.

McCartney AL. Application of molecular biological methods for studying probiotics and the gut flora. (2002). *British Journal of Nutrition, 88(1):*S29–S37

MCDONALD, D., PRICE, M. N., GOODRICH, J., NAWROCKI, E. P., DESANTIS, T. Z., & PROBST, A. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME Journal, 6*(3), 610.

MCDONALD JA, FUENTES S, SCHROETER K, HEIKAMPDEJONG I, KHURSIGARA CM, DE VOS WM, ALLENVERCOE E. (2015). Simulating distal gut mucosal and luminal communities using packed-column biofilm reactors and an in vitro chemostat model. *Journal of Microbiology Methods, 108*:36–44

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.

MELNIK, A. V., DA SILVA, R. R., HYDE, E. R., AKSENOV, A. A., VARGAS, F., BOUSLIMANI, A., & KNIGHT, R. (2017). Coupling targeted and untargeted mass spectrometry for metabolome-microbiome-wide association studies of human fecal samples. *Analytical chemistry, 89*(14), 7549-7559.

MOREIRA, A. P. B., TEXEIRA, T. F. S., FERREIRA, A. B., PELUZIO, M. D. C. G., & ALFENAS, R. D. C. G. (2012). Influence of a high-fat diet on gut microbiota, intestinal permeability and metabolic endotoxaemia. *British Journal of Nutrition, 108*(5), 801-809.

MORGAN XC, HUTTENHOWER C (2012) Chapter 12: Human Microbiome Analysis. *PLoS Computational Biology 8*(12) e1002808.

MORTON, E. R., LYNCH, J., FROMENT, A., LAFOSSE, S., HEYER, E., PRZEWORSKI, M., SEGUREL, L. (2015). Aariation in rural African gut microbiota is strongly correlated with colonization by Entamoeba and subsistence. *PLoS Genetics, 11*(11)

MUELLER, S., SAUNIER, K., HANISCH, C., NORIN, E., ALM, L., MIDTVEDT, T., VERDENELLI, M. C. (2006). Differences in fecal microbiota in different European study

populations in relation to age, gender, and country: a cross-sectional study. *Applied and Environmental Microbiology, 72*(2), 1027-1033.

MULLISH, B. H., PECHLIVANIS, A., BARKER, G. F., THURSZ, M. R., MARCHESI, J. R., & MCDONALD, J. A. (2018). Functional microbiomics: evaluation of gut microbiota-bile acid metabolism interactions in health and disease. *Methods, 149*, 49-58.

MUTAVHADSINDI, M. A., & MEIRING, P. G. (2014). Church planting in South Africa: The role of the Reformed Church Tshiawelo. *Verbum et Ecclesia, 35(*1), 01-07.

MUYZER, G. DGGE/TGGE a method for identifying genes from natural ecosystems. (1999). *Curruent Opinion in Microbiology, 2,* 317–322

NAGAO-KITAMOTO, H., KITAMOTO, S., KUFFA, P., & KAMADA, N. (2016). Pathogenic role of the gut microbiota in gastrointestinal diseases. *Intestinal Research, 14(*2), 127-138.

NAGPAL, S., HAQUE, M. M., & MANDE, S. S. (2016). Vikodak-a modular framework for inferring functional potential of microbial communities from 16S metagenomic datasets. *PLoS One, 11*(2)

NAKAYAMA, J., YAMAMOTO, A., PALERMO-CONDE, L. A., HIGASHI, K., SONOMOTO, K., TAN, J., & LEE, Y.-K. (2017). Impact of westernized diet on gut microbiota in children on Leyte Island. *Frontiers in Microbiology, 8*, 197.

NICHOLSON, J. K., HOLMES, E., KINROSS, J., BURCELIN, R., GIBSON, G., JIA, W., & PETTERSSON, S. (2012). Host-Gut Microbiota Metabolic Interactions. *Science, 336*(6086), 1262.

NISHIJIMA, S., SUDA, W., OSHIMA, K., KIM, S. W., HIROSE, Y., MORITA, H., & HATTORI, M. (2016). The gut microbiome of healthy Japanese and its microbial and functional uniqueness. *DNA Research, 23*(2), 125-133.

OBREGON-TITO, A. J., TITO, R. Y., METCALF, J., SANKARANARAYANAN, K., CLEMENTE, J. C., URSELL, L. K., GAFFNEY, P. M. (2015). Subsistence strategies in traditional societies distinguish gut microbiomes. *Nature Communications, 6*, 6505.

O'HARA, A. M., & SHANAHAN, F. (2006). The gut flora as a forgotten organ. *EMBO Rep, 7*(7), 688-693.

O'KEEFE, S. J. (2008). Nutrition and colonic health: the critical role of the microbiota. *Current Opininion in Gastroenterology, 24*(1), 51-58.

O'KEEFE, S. J., CHUNG, D., MAHMOUD, N., SEPULVEDA, A. R., MANAFE, M., ARCH, J., VAN DER MERWE, T. (2007). Why do African Americans get more colon cancer than Native Africans? *The Journal of Nutrition, 137*(1), 175S-182S.

O'KEEFE, S. J., LI, J. V., LAHTI, L., OU, J., Carbonero, F., Mohammed, K., Zoetendal, E. G. (2015). Fat, fibre and cancer risk in African Americans and rural Africans. *Nature Communications, 6,* 6342.

OKSANEN, J., BLANCHET, F. G., FRIENDLY, M., KINDT, R., LENGENDRE, P., MCGLINN, D., SOLYMOS, P. (2017). Vegan: community ecology package.

OU, J., CARBONERO, F., ZOETENDAL, E. G., DELANY, J. P., WANG, M., NEWTON, K., O'KEEFE, S. J. (2013). Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans. *The American Journal of Clinical Nutrition, 98*(1), 111-120.

OU, J., DELANY, J. P., ZHANG, M., SHARMA, S., & O'KEEFE, S. J. (2012). Association between low colonic short-chain fatty acids and high bile acids in high colon cancer risk populations. *Nutrition and Cancer, 64*(1), 34-40.

PALIY, O., KENCHE, H., ABERNATHY, F. & MICHAIL, S. High-throughput quantitative analysis of the human intestinal microbiota with a phylogenetic microarray. (2009). *Applied and Environmental Microbiology. 75*, 3572–3579.

PALMER, C., BIK, E. M., DIGIULIO, D. B., RELMAN, D. A., & BROWN, P. O. (2007). Development of the human infant intestinal microbiota. *PLOS Biology, 5*(7), e177.

PARFREY LW, WALTERS WA, KNIGHT R (2011) Microbial eukaryotes in the human microbiome: ecology, evolution, and future directions. *Frontiers in Microbiology 2*: 153.

PARKS, D. H., TYSON, G. W., HUGENHOLTZ, P., & BEIKO, R. G. (2014). STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics, 30*(21), 3123-3124.

PENDERS, J., THIJS, C., VINK, C., STELMA, F. F., SNIJDERS, B., KUMMELING, I., STOBBERINGH, E. E. (2006). Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics, 118*(2), 511-521.

PETRI, R. M., POURAZAD, P., KHIAOSA-ARD, R., KLEVENHUSEN, F., METZLER-ZEBELI, B. U., & ZEBELI, Q. (2017). Temporal dynamics of in-situ fiber-adherent bacterial community under ruminal acidotic conditions determined by 16S rRNA gene profiling. *PloS One, 12*(8)

PORTUNE, K. J., BENÍTEZ-PÁEZ, A., DEL PULGAR, E. M. G., CERRUDO, V., & SANZ, Y. (2017). Gut microbiota, diet, and obesity-related disorders—The good, the bad, and the future challenges. *Molecular Nutrition & Food Research, 61*(1), 1600252

PRUESSE E, QUAST C, KNITTEL K, FUCHS BM, LUDWIG W, et al. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Reseach 35*: 7188–7196

PRYDE, S. E., DUNCAN, S. H., HOLD, G. L., STEWART, C. S., & FLINT, H. J. (2002). The microbiology of butyrate formation in the human colon. *FEMS Microbiology Letters, 217*(2), 133-139.

PURCHIARONI, F., TORTORA, A., GABRIELLI, M., BERTUCCI, F., GIGANTE, G., IANIRO, G., GASBARRINI, A. (2013). The role of intestinal microbiota and the immune system. *European Review for Medical and Pharmacological Sciences, 17*(3), 323-333.

QIN, J., LI, R., RAES, J., ARUMUGAM, M., BURGDORF, K. S., MANICHANH, C., WANG, J. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature, 464*(7285), 59-65.

RAJOKA, M. S. R., SHI, J., MEHWISH, H. M., ZHU, J., LI, Q., SHAO, D., YANG, H. (2017). Interaction between diet composition and gut microbiota and its impact on gastrointestinal tract health. *Food Science and Human Wellness*, *6*(3), 121-130

RAMPELLI, S., CANDELA, M., TURRONI, S., BIAGI, E., COLLINO, S., FRANCESCHI, C., BRIGIDI, P. (2013). Functional metagenomic profiling of intestinal microbiome in extreme ageing. *Aging (Albany NY), 5.*

RAMPELLI, S., SCHNORR, STEPHANIE L., CONSOLANDI, C., TURRONI, S., SEVERGNINI, M., PEANO, C., CANDELA, M. (2015). Metagenome sequencing of the Hadza hunter-gatherer gut microbiota. *Current Biology, 25*(13), 1682-1693.

RAMPELLI, S., SOVERINI, M., D'AMICO, F., BARONE, M., TAVELLA, T., MONTI, D., & FRANCESCHI, C. (2019). Shotgun metagenomics of human gut microbiota up to extreme longevity and the increasing role of xenobiotics degradation. *Current-Biology 19*-01035.

RANJAN, R., RANI, A., FINN, P. W., & PERKINS, D. L. (2018). Multiomic Strategies Reveal Diversity and Important Functional Aspects of Human Gut Microbiome. *BioMed Research International, 2018*

RANJAN, R., RANI, A., METWALLY, A., MCGEE, H. S., & PERKINS, D. L. (2016). Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochemical and Biophysical Research Communications, 469*(4), 967-977.

RASTALL, R. (2004). Bacteria in the gut: friends and foes and how to alter the balance. *The Journal of Nutrition, 134*(8), 2022S-2026S.

RAVEL, J., GAJER, P., ABDO, Z., SCHNEIDER, G. M., KOENIG, S. S., & MCCULLE, S. L. (2011). Vaginal microbiome of reproductive-age women. *Proceedings of the National Academy of Sciences, 108*.

REYES, A., SEMENKOVICH, N. P., WHITESON, K., ROHWER, F., & GORDON, J. I. (2012). Going viral: next-generation sequencing applied to phage populations in the human gut. *Nature Reviews Microbiology., 10*.

RICHARDS, J. L., YAP, Y. A., MCLEOD, K. H., MACKAY, C. R., & MARINO, E. (2016). Dietary metabolites and the gut microbiota: an alternative approach to control inflammatory and autoimmune diseases. *Clinical Translational Immunology, 5*(5),

RODRÍGUEZ, J. M., MURPHY, K., STANTON, C., ROSS, R. P., KOBER, O. I., JUGE, N., COLLADO, M. C. (2015). The composition of the gut microbiota throughout life, with an emphasis on early life. *Microbial Ecology in Health and Disease*, *26*(1), 26050

SAAD, R., RIZKALLAH, M. R., & AZIZ, R. K. (2012). Gut Pharmacomicrobiomics: the tip of an iceberg of complex interactions between drugs and gut-associated microbes. *Gut Pathogens, 4*(1), 16.

SAKUMA, K., FUNABASHI, H., MATSUOKA, H., & SAITO, M. (2013). Potential use of *Lactobacillus* cell density in feces as a non-invasive bio-indicator for evaluating environmental stress during mouse breeding. *Biocontrol Science, 18*(2), 101–104.

SAMUEL, B. S., & GORDON, J. I. (2006). A humanized gnotobiotic mouse model of host-archaeal-bacterial mutualism. *Proceedings of the National Academy of Science, 103*(26), 10011-10016.

SANDERS, L M. (2016). Carbohydrates: Digestion, absoption and metabolism. *Elsevier*, 643-650.

SCHIPPA, S., & CONTE, M. P. (2014). Dysbiotic events in gut microbiota: impact on human health. *Nutrients, 6*(12), 5786-5805.

SCHNORR, S. L., CANDELA, M., RAMPELLI, S., CENTANNI, M., CONSOLANDI, C., BASAGLIA, G., CRITTENDEN, A. N. (2014). Gut microbiome of the Hadza hunter-gatherers. Nature Communications, 5, 3654.

SCHWIERTZ, A., TARAS, D., SCHAFER, K., BEIJER, S., BOS, N. A., DONUS, C., & HARDT, P. D. (2010). Microbiota and SCFA in lean and overweight healthy subjects. *Obesity (Silver Spring)*, *18*(1), 190-195.

SIMPSON, H. L., & CAMPBELL, B. J. (2015). Review article: dietary fibre-microbiota interactions. *Alimenaryt Pharmacology and Therapeutics, 42*(2), 158-179.

SINGH, R. K., CHANG, H.-W., YAN, D., LEE, K. M., UCMAK, D., WONG, K., LIAO, W. (2017). Influence of diet on the gut microbiome and implications for human health. *Journal of Translational Medicine, 15*(1), 73.

SINGLETON, I. (1994). Microbial metabolism of xenobiotics: fundamental and applied research. Journal of Chemical Technology & Biotechnology: *International Research in Process, Environmental and Clean Technology, 59*(1), 9-23.

SOERGEL DA, DEY N, KNIGHT R, BRENNER S.E. (2012). Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. *ISME Journal, 6*(7):1440-4

SONG, E. J., LEE, E. S., & NAM, Y. D. (2018). Progress of analytical tools and techniques for human gut microbiome research. *Journal of Microbiology, 56*(10), 693-705.

STEARNS, J. C., ZULYNIAK, M. A., DE SOUZA, R. J., CAMPBELL, N. C., FONTES, M., SHAIKH, M., ANAND, S. S. (2017). Ethnic and diet-related differences in the healthy infant microbiome. *Genome Medicine, 9*(1), 32.

STEYNA, N. P. (2006). Chronic diseases of lifestyle in South Africa. *South African Medical Research Council 54*(67), 33.

SUZUKI, T. A., & WOROBEY, M. (2014). Geographical variation of human gut microbial composition. *Biology Letters, 10*(2)

TAN, J., MCKENZIE, C., POTAMITIS, M., THORBURN, A. N., MACKAY, C. R., & MACIA, L. (2014). The role of short-chain fatty acids in health and disease. *Advances in Immunology, 121*, 91-119.

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.

TEAM, R. C. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2016.

THE HUMAN MICROBIOME PROJECT, C., METHÉ, B. A., NELSON, K. E., POP, M., CREASY, H. H., GIGLIO, M. G., WHITE, O. (2012). A framework for human microbiome research. *Nature, 486*, 215.

THORBURN, A. N., MACIA, L., & MACKAY, C. R. (2014). Diet, metabolites, and "western-lifestyle" inflammatory diseases. *Immunity, 40*(6), 833-842.

THURSBY, E., & JUGE, N. (2017). Introduction to the human gut microbiota. *Biochemical Journal, 474*(11), 1823-1836.

TREMAROLI, V., & BACKHED, F. (2012). Functional interactions between the gut microbiota and host metabolism. *Nature*, 489(7415), 242-249.

TURNBAUGH, P. J., & GORDON, J. I. (2009). The core gut microbiome, energy balance and obesity. *Journal of Physiology, 587*(Pt 17), 4153-4158.

TURNBAUGH, P. J., BACKHED, F., FULTON, L., & GORDON, J. I. (2008). Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host Microbe, 3*(4), 213-223

TURNBAUGH, P. J., HAMADY, M., YATSUNENKO, T., CANTAREL, B. L., DUNCAN, A., LEY, R. E., GORDON, J. I. (2009). A core gut microbiome in obese and lean twins. *Nature, 457*(7228), 480-484.

TURNBAUGH, P. J., LEY, R. E., HAMADY, M., FRASER-LIGGETT, C. M., KNIGHT, R., & GORDON, J. I. (2007). The human microbiome project. *Nature, 449*(7164), 804.

TURNBAUGH, P. J., RIDAURA, V. K., FAITH, J. J., REY, F. E., KNIGHT, R., & GORDON, J. I. (2009). The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Science Translational Medicine, 1*(6)

TURPIN, W., ESPIN-GARCIA, O., XU, W., SILVERBERG, M. S., KEVANS, D., SMITH, M. I., CROITORU, K. (2016). Association of host genome with intestinal microbial composition in a large healthy cohort. *Nature Genetics, 48,* 1413.

VAN DER MERWE, M. T., & PEPPER, M. S. (2006). Obesity in South Africa. *Obesity Reviews, 7*(4), 315-322.

VARKI, A. (2016). Biological roles of glycans. *Glycobiology, 27*(1), 3-49.

VERNOCCHI, P., DEL CHIERICO, F., & PUTIGNANI, L. (2016). Gut microbiota profiling: metabolomics based approach to unravel compounds affecting human health. *Frontiers in Microbiology, 7*, 1144.

VESTERBACKA, J., RIVERA, J., NOYAN, K., PARERA, M., NEOGI, U., CALLE, M., NOWAK, P. (2017). Richer gut microbiota with distinct metabolic profile in HIV infected elite controllers. *Scientific Reports, 7*(1), 6269.

VIPPERLA, K., & O'KEEFE, S. J. (2012). The microbiota and its metabolites in colonic mucosal health and cancer risk. *Nutrition in Clinical Practice, 27*(5), 624-635.

VOREADES, N., KOZIL, A., & WEIR, T. L. (2014). Diet and the development of the human intestinal microbiome. *Frontiers in Microbiology, 5.*

WALSH, C. J., GUINANE, C. M., O'TOOLE, P. W., & COTTER, P. D. (2014). Beneficial modulation of the gut microbiota. *FEBS Letters, 588*(22), 4120-4130.

WALTERS, W. A., XU, Z., & KNIGHT, R. (2014). Meta-analyses of human gut microbes associated with obesity and IBD. *FEBS Letters, 588*(22), 4223-4233.

WEST, C. E., RENZ, H., JENMALM, M. C., KOZYRSKYJ, A. L., ALLEN, K. J., VUILLERMIN, P. (2015). The gut microbiota and inflammatory noncommunicable diseases: associations and potentials for gut microbiota therapies. *Journal of Allergy Clinical Immunology, 135*(1), 3-13; quiz 14.

WEST, N. P., CHRISTOPHERSEN, C. T., PYNE, D. B., CRIPPS, A. W., CONLON, M. A., & TOPPING, D. L. (2013). Butyrylated starch increases colonic butyrate concentration but has limited effects on immunity in healthy physically active individuals. *Exercise Immunology Review, 19*.

WOODMANSEY, E. J. (2007). Intestinal bacteria and ageing. *Journal of Applied Microbiology, 102*(5), 1178-1186.

WOYKE, T., TEELING, H., IVANOVA, N. N., HUNTEMANN, M., RICHTER, M., GLOECKNER, F. O., DUBILIER, N. (2006). Symbiosis insights through metagenomic analysis of a microbial consortium. *Nature, 443*(7114), 950-955.

WU, G. D., BUSHMANC, F. D., & LEWIS, J. D. (2013). Diet, the human gut microbiota, and IBD. *Anaerobe, 24*, 117-120.

WU, G. D., CHEN, J., HOFFMANN, C., BITTINGER, K., CHEN, Y. Y., KEILBAUGH, S. A., LEWIS, J. D. (2011). Linking long-term dietary patterns with gut microbial enterotypes. *Science, 334*(6052), 105-108.

XIE, H., GUO, R., ZHONG, H., FENG, Q., LAN, Z., QIN, B., CHEN, X. (2016). Shotgun metagenomics of 250 adult twins reveals genetic and environmental impacts on the gut microbiome. *Cell Systems, 3*(6), 572-584. e573.

YANG, I., CORWIN, E. J., BRENNAN, P. A., JORDAN, S., MURPHY, J. R., & DUNLOP, A. (2016). The infant microbiome: implications for infant health and neurocognitive development. *Nursing Research, 65*(1), 76.

YATSUNENKO, T., REY, F. E., MANARY, M. J., TREHAN, I., DOMINGUEZ-BELLO, M. G., CONTRERAS, M., GORDON, J. I. (2012). Human gut microbiome viewed across age and geography. *Nature, 486*(7402), 222-227.

ZHANG, J., GUO, Z., LIM, A. A. Q., ZHENG, Y., KOH, E. Y., HO, D., ZHANG, H. (2014). Mongolians core gut microbiota and its correlation with seasonal dietary changes. *Scientific Reports, 4*, 5001.

ZHOU, C., ONTIVEROS-VALENCIA, A., WANG, Z., MALDONADO, J., ZHAO, H. P., KRAJMALNIK-BROWN, R., & RITTMANN, B. E. (2016). Palladium recovery in a H2-based membrane biofilm reactor: formation of Pd (0) nanoparticles through enzymatic and autocatalytic reductions. *Environmental Science & Technology, 50*(5), 2546-2555.

ZOETENDAL, E. G., AKKERMANS, A. D. L., AKKERMANS-VAN VLIET, W. M., VISSER, J. A., & VOS, W. M. (2001). The host genotype affects the bacterial community in the human gastrointestinal tract. *Microbial Ecology in Health and Disease, 13,* 129-134.