Temporal change in the Namibian Ju|'hoansi intestinal microbiome

by

Tiffany du Plessis

Submitted in partial fulfilment for the degree

Magister Scientiae Bioinformatics

in the Faculty of Natural & Agricultural Sciences

University of Pretoria

Pretoria

April 2023

Supervisor: Dr Riaan Francois Rifkin

Co-Supervisor: Prof Fourie Joubert

Declaration of Authorship

I, Tiffany du Plessis, declare that the dissertation, which I hereby submit for the degree MSc Bioinformatics at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE: 

DATE: 16 July 2023

Summary

Faculty of Natural and Agricultural Sciences

Magister Scientiae Bioinformatics

By Tiffany du Plessis

The microbes which reside in our digestive tract comprise the human intestinal microbiome (IM). Many historical events have influenced the composition of our IM, such as the invention of cooking, the advent of farming and the industrial revolution. Although ancestral IMs could provide insight into the evolution of the human IM, pristine ancestral hunter-gatherer IM samples are near-impossible to acquire. Alternatively, studying contemporary hunter-gatherers might allow us to gain insight into the pre-Industrial human IM. The Ju|'hoansi hunter-gatherers of Namibia are in transition to an increasingly Western lifestyle, providing an opportunity to study the effects of westernization on the human IM. Various factors impact our IM, including diet and our interaction with the environment. In this regard, the influence of drinking water on our IM is largely unknown and necessitates further exploration. This study aims to detect temporal changes in the Ju|'hoansi IM, while also exploring the impact of water sources on their IM, and the levels of the IM volatile fatty acids. In order to do so, faecal samples were collected and subsequently sequenced from a total cohort of 40 participants taken 2 years apart in 2019 and 2021. Based on our study, we were able to identify changes in the composition of microbial communities over a 2-year period. Statistically significant differences were found in the taxonomic composition of the Ju|'hoansi IM after 2 years. Within each year, the village of residency of each participant proved significant. The richness of the Ju|'hoansi IM did not change during this time period and still reflected that of a traditional hunter-gatherer IM. From baseline to follow-up, numerous bacterial genera differentially decreased, and the core microbiome significantly decreased in species. Only one volatile fatty acid was found to have significant correlations with the Ju|'hoansi IM, decanoic acid. The nitrate concentration, Langelier saturation index and the total oxidised nitrogen in the water sources impacted the IM composition of the Ju|'hoansi, causing differentially abundant genera. Chapter 2 is a draft manuscript that will be sent to Cell Press Community Review intended for publication a in Cell Press Journal.

Acknowledgements

To begin with, I would like to thank my supervisor, Dr Riaan F. Rifkin, for giving me the opportunity to be a part of this research project as well as for his unwavering support and guidance. I am grateful to my co-supervisors, Dr Jessica E. Koopman for getting me set up and started, and Prof Fourie Joubert for seeing this project finished. As a result of the power outages during the last two years, I wish to acknowledge the efforts of the Centre for Bioinformatics and Computational Biology at the University of Pretoria in keeping its servers running. I extend my gratitude for the financial assistance received from my funders, the Benjamin R. Oppenheimer Trust, which enabled me to complete my MSc. During my MSc, I received invaluable support from Dr Diego Castillo, a co-supervisor from my Honours year, and Dr Karen Jordaan, who were instrumental in helping me with my practical analyses. I am thankful for my friends at the University of Pretoria, Adri Grobler and Nyasha Mafumo. Supportive academic friends are a necessity for any postgraduate student. Last but not least, I would like to thank my family for supporting me throughout my academic endeavours and for encouraging me to pursue further education. I could not have achieved this without the love and support of my family.

Table of Contents

Declaration of Authorship ii

Summary iii

Acknowledgements iv

List of Figures vii

List of Abbreviations viii

Chapter 1: Importance of studying the traditional intestinal microbiome 1

Historical progression of the human intestinal microbiome 2

Relevance of hunter-gatherer microbiomes 4

Comparison of Western and traditional intestinal microbiomes 6

Ju|'hoansi 7

Conclusion 9

Aim and Objectives 11

Aim 11

Objectives 11

Chapter 2: Temporal change in the Namibian Ju|'hoansi intestinal microbiome 12

Summary 13

Introduction 14

Results 15

Cohort description 15

Ju'|hoansi IM composition 15

Temporal changes in IM microbial diversity 17

Discussion 20

Conclusion 23

Materials and Methods 24

Sampling 24

Bioinformatic pipeline 25

Analysis in R 26

Acknowledgements 27

Author Contributions 27

Declarations of Interest 27

Code Availability 27

Chapter 3: The effect of unfiltered borehole drinking water on the Namibian Ju|'hoansi intestinal microbiome 28

Summary 29

Introduction 30

Results 32

Correlation analysis 32

Water microbiome analysis 32

Discussion 34

Conclusion 36

Materials and Methods 37

Data collection 37

Bioinformatic pipeline for water samples 37

Analysis in R 37

Chapter 4: Conclusions 38

Summary 39

Temporal change in the Namibian Ju|'hoansi intestinal microbiome 39

The effect of unfiltered borehole drinking water on the Namibian Ju|'hoansi intestinal microbiome 39

Future prospects 40

Conclusion 40

Supplementary Information 41

Supplementary R Code 63

References 75

List of Figures

Figure 1. Geographic locations and populations represented in existing studies on lifestyle effects on the human IM......8
Figure 2. Map showing the location of the Nyae Nyae Conservancy in Namibia. 9
Figure 3. The bacterial core microbiome of the Ju|'hoansi from A) baseline to B) follow-up at differing prevalence and detection thresholds......16
Figure 4. Per-year α-diversity indicated with A) Observed, B) Shannon and C) Simpson indices. Baseline is represented in red and follow-up in blue. Statistical significance shown above the boxplots indicating either not significant (ns) or significant at <0.01 (**). 17
Figure 5. Combination violin and boxplots indicating A) Shannon diversity indices for baseline in red and follow-up in blue and B) Simpson diversity indices for baseline in red and follow-up in blue. 17
Figure 6. Principal coordinate analysis (PCoA) plot with Bray-Curtis dissimilarity between sampling years. Baseline is represented in red and follow-up in blue. 18
Figure 7. Bar plot showing the log2-fold-change of the differentially abundant genera and correlating phylum. Points on the left-hand side of the zero line show genera that have decreased in abundance in follow-up since baseline, those on the right-hand side have increased since......19
Figure 8. Map showing the sampling areas in the Nyae Nyae Conservancy, Tsumkwe, Namibia. 24

List of Abbreviations

ASV	Amplicon sequence variant
IM	Intestinal microbiome
ITS	Internal transcribed spacer
JUTA	The Ju|'hoansi Traditional Authority
NCRST	The Namibian National Commission on Research and Technology
NNC	The Nyae Nyae Conservancy in Namibia
OUT	Operational taxonomic unit
rRNA	Ribosomal ribonucleic acid
SCFA	Short chain fatty acid
SI	Supplementary information
TB	Tuberculosis
UK	United Kingdom
VFA	Volatile fatty acid

Chapter 1: Importance of studying the traditional intestinal microbiome

Historical progression of the human intestinal microbiome

The human microbiome comprises all the microbial cells that occupy the human body, including bacteria, viruses, and fungi. There are conflicting sources describing the estimated number of bacterial species inhabiting one's intestines from 200 to 1000 species, this forms the intestinal microbiome (IM) (Lloyd-Price *et al.*, 2016; Gilbert *et al.*, 2018). A metagenome refers to the collective genes and combined genomes of a community, such as a microbiome (Walter & Ley, 2011). Our own genome contains at least 100 times less genes than that of the metagenome concerning our intestinal microbes (Gill *et al.*, 2006; De Filippo *et al.*, 2010; Qin *et al.*, 2010).

The IM holds significant implications with regards to maintaining homeostasis in the human body, and the metabolic functionality of the taxonomic communities present in the IM has a noteworthy influence on the health of its host (Davenport *et al.*, 2017; Selber-Hnatiw *et al.*, 2017; Das & Nair, 2019). Changes occur in IM composition throughout an individual's lifetime, impacted by changes in host health or diseases (Odamaki *et al.*, 2016; Chen *et al.*, 2018). Other factors that could affect the composition of the IM are diet, the use of medication and even seasonality (Schnorr *et al.*, 2014; Smits *et al.*, 2017; Kurilshikov *et al.*, 2021).

Dominant bacterial IM phyla generally comprise Bacteroidota, Firmicutes, Actinobacteria and Proteobacteria but the quantity of each vary from host to host (Davenport *et al.*, 2017; Das & Nair, 2019). These play an important role in the IM, as has been extensively studied. Bacteroidota, Proteobacteria, and certain Firmicutes are all involved in carbohydrate digestion (Russell *et al.*, 2014; Berry, 2016; Selber-Hnatiw *et al.*, 2017). Moreover, Bacteroidota produce acetate and propionate, which are two key short chain fatty acids (SCFAs) (Feng *et al.*, 2018). It has been shown that Firmicutes are crucial in producing butyrate and a number of Firmicutes have also been observed to ferment complex carbohydrates (Dinan & Cryan, 2017; Selber-Hnatiw *et al.*, 2017; Feng *et al.*, 2018; Das & Nair, 2019). According to a recent study, having increased abundance of Actinobacteria in the gut may be associated with a healthy IM (Kurilshikov *et al.*, 2021).

Commonly referred to as the virome, the complex viral microbiome of the human intestinal tract plays an important role in the ecosystem that is our intestines. The virome encompasses a wide range of viruses from eukaryotic viruses to bacteriophages (Mukhopadhya *et al.*, 2019; Liang & Bushman, 2021). Recent studies have looked at the link of viruses in our intestines to the development of certain diseases such as obesity, metabolic syndrome, and Crohn's disease (Norman *et al.*, 2015; Bikel *et al.*, 2021; Shareefdeen

& Hill, 2022). Looking at the IM of children there has been an association found between a higher diversity and richness of phages and the presence of obesity and or metabolic syndrome (Bikel *et al.*, 2021). Similarly with Crohn's disease there was an association found between increased richness of *Caudovirales* and the development of Crohn's disease and ulcerative colitis (Norman *et al.*, 2015). As is with the bacterial IM, there isn't one intestinal virome that can represent a healthy virome in the human IM. What does seem consistent with the intestinal virome is the dominance of *Caudovirales* bacteriophages (Mukhopadhya *et al.*, 2019; Liang & Bushman, 2021). With regards to this study, we will only be looking at the gut bacterial microbiome of the Ju|'hoansi.

Aside from the bacterial and viral microbiome that exists in the gut, there is also the mycobiome made up of fungi. The mycobiome is the fungal microbiome and describes the complex fungal community present in an environment (Gillevet *et al.*, 2009). Similarly to the bacterial and viral IM, there is no definitive core gut mycobiome (Pérez, 2021). The most common fungus found in humans' faeces is *Candida albicans*, which is therefore considered to be part of the normal mycobiome in the intestines. While there is a limited amount of research on gut fungi and their association with intestinal immunity, it has been demonstrated that fungal members of intestinal microbiota can lessen or trigger inflammatory responses (Iliev *et al.*, 2012; Wheeler *et al.*, 2016). In order to gain information on the impact fungi have on out gut, it would require more large-scale studies to be performed. Although ITS sequencing was successfully done on samples from our baseline in 2019, we were unable to perform ITS sequencing on the follow-up samples from 2021, and thus cannot explore temporal changes to the fungal mycobiome of the bushmen in Namibia for this research project.

The overall taxonomic IM composition differs significantly between populations as well as between individuals, and there is a high inter-individuality with regards to the human IM. The age of an individual is one factor contributing to the high variability of IM composition between individuals. From birth to weaning, the IM undergoes many fluctuations, but after solid foods are introduced, it resembles the adult state of the IM (Quercia *et al.*, 2014). Lifestyle, such as that in traditional or urban societies, and health status affect IM composition dramatically. This presents difficulties in defining the healthy human IM (Walter & Ley, 2011).

The loss of microbial homeostasis in the IM leads to dysbiosis as well as disproportion in microbial taxonomic communities, which may be harmful to the human body. Various diseases and disorders have been linked to a loss of taxonomic diversity of the IM, including obesity, inflammatory bowel disease, colon cancer and even depression (Das & Nair, 2019; Ding *et al.*, 2019). There is strong evidence that

dysbiosis in the major taxa present in the IM is associated with health issues including obesity, which has been reported to be associated with an abundance of Firmicutes and a decreased abundance of Bacteroidota (Selber-Hnatiw *et al.*, 2017; Das & Nair, 2019). The ratio of Firmicutes to Bacteroidota can serve as an indicator of dysbiosis (Selber-Hnatiw *et al.*, 2017; Ding *et al.*, 2019). The loss of Firmicutes and Bacteroidota, along with the increase of Proteobacteria, can be a microbial signature of inflammatory bowel disease (Das & Nair, 2019). In patients with irritable bowel syndrome, Firmicutes can be enriched and Bacteroidota depleted (Selber-Hnatiw *et al.*, 2017). The microbial signature of undernutrition is enhanced Proteobacteria abundance as well as reduced Firmicutes abundance (Das & Nair, 2019). Due to the health implications of changes in the microbiome, a comprehensive understanding of the factors that affect the diversity, composition and functioning of the human microbiome is essential to understanding the relationship between human health and the microbiome (Rosas-Plaza *et al.*, 2022).

Our IM holds significance with regards to our health and it is important to acknowledge the historical events that have shaped our IM. Although the lifestyle changes associated with the invention of cooking and the Neolithic Revolution significantly impacted hunter-gatherer IM taxonomic composition and metabolic capacity (Walter & Ley, 2011; Quercia *et al.*, 2014), the impact of the Industrial Revolution, and the process of global 'westernization' on the human IM, is particularly marked (Blaser & Falkow, 2009; Gillings & Paulsen, 2014; Segata, 2015; Schnorr *et al.*, 2016; Rifkin *et al.*, 2020). In order to understand the impact of such historical events, it is important to study the evolution of our IM, which can be done by examining the IM of contemporary hunters and gatherers, such as the Hadza hunter-gatherers of Tanzania (Schnorr *et al.*, 2014). The possibility of obtaining and analysing pristine intestinal or faecal samples that represent our hunter-gatherer ancestors is extremely unlikely, thus present-day hunter-gatherers, such as the Ju|'hoansi of Namibia, may provide the closest resemblance to the ancestral IM. Accordingly, in this study, we present a long-term follow-up analysis of the IM of 37 Ju|'hoansi participants in which we compare samples taken two years apart, 2019 and 2021.

Relevance of hunter-gatherer microbiomes

Studying contemporary hunter-gatherers can enhance our current understanding of how people lived before the Industrial Revolution, due to being exposed to minimal Western influences like new types of food, or modern medicine. Several studies concerning the human IM have explored direct or indirect implications of westernization on our IM, at times along with a comparison of traditional societies to Western societies (De Filippo *et al.*, 2010; Walter & Ley, 2011; Gomez *et al.*, 2016; Davenport *et al.*, 2017; Mancabelli *et al.*, 2017; Jha *et al.*, 2018). Here, the term 'traditional' refers to rural subsistence-

based lifestyles that are non-industrialised. The term 'Western' is used here to refer to urban-industrialised lifestyle. Rural or 'traditional' non-industrialized communities typically consume low-fat, low-sugar diets, and medical access is limited. This type of 'traditional' context typically involves people living close to one another, their pets, livestock, and wildlife, as well as exposing them to more microbes in their environment. Contrary to this, the Western diet, which is common in urban-industrialized societies, tends to include high-fat, low-fibre processed foods, increased sedentarism and increased contact with modern medicine. Natural environment exposure and microbe exposure tend to be less prevalent in Western communities (Rook *et al.*, 2014; Thorburn *et al.*, 2014).

There is a lot of available research and reviews on the impact of westernization on the IM and its influence on human health (Thorburn *et al.*, 2014; Segata, 2015; Schnorr *et al.*, 2016; Ayeni *et al.*, 2018; Vangay *et al.*, 2018). One study cohort saw "microbiome westernization" occurring within 9 months of immigration to the US from Thailand (Vangay *et al.*, 2018).

Traditional hunter-gatherer communities such as the Hadza in Tanzania (Schnorr *et al.*, 2014) or the BaAka in the Central African Republic (Gomez *et al.*, 2016), who have both been well documented, have not been completely immune to Western influence and cannot be considered pure hunter-gatherers. As previously mentioned, it is near-impossible to obtain and analyse pristine intestinal or faecal samples that may accurately represent our hunter-gatherer ancestors (Walter & Ley, 2011). The closest resemblance of the ancestral IM may however be provided by contemporary hunter-gatherers, such as the Ju|'hoansi of Namibia.

Although the Ju|'hoansi have thus far had contact with Western society, their remote geographic location, unique lifestyle, and dependence on hunting and gathering may have resulted in minimal Western impact. They are nevertheless 'in transition' to a more Western lifestyle, and it is therefore important to gain information about their IM composition and metabolic capacity, while the Western influence remains comparatively minimal. Studying a traditional society 'in transition' to a more Western lifestyle, provides the opportunity to study the effects westernization may have had on early human IMs, allowing us to explore the consequences we may be experiencing today in the Western world.

The analysis of traditional societies can elucidate the evolutionary history of our IM. An earlier change in our dietary habits came about with the invention of cooking our food, followed by the Neolithic Revolution, when agriculture and livestock practices developed (Walter & Ley, 2011; Quercia *et al.*,

2014; Rifkin *et al.*, 2020). A subsequent historical event impacting our IM came around with the industrial revolution in the late 18th century, increasing the availability of processed flour and sugar (Gillings & Paulsen, 2014). These historical events are held to have significantly impacted our IM (De Filippo *et al.*, 2010; Walter & Ley, 2011; Quercia *et al.*, 2014).

One significant contribution of our IM is the substantial implication in the development of non-communicable diseases, such as inflammatory bowel disease and autoimmune diseases (Ochoa-Reparaz *et al.*, 2010; Berer *et al.*, 2011; Devkota *et al.*, 2012; Trompette *et al.*, 2014; Martínez *et al.*, 2015). Non-communicable diseases are increasing in prevalence in Western countries (Devkota *et al.*, 2012; Trompette *et al.*, 2014; Martínez *et al.*, 2015; Jacobson *et al.*, 2021). As a result of the importance of our IM in the development of non-communicable diseases, this further proves that studying societies that resemble early hunter-gatherers may provide insight into today's rise of non-communicable diseases. Therefore, comparative studies involving traditional and Western microbiomes are warranted.

Comparison of Western and traditional intestinal microbiomes

People living in traditional societies rely on foraging, limited agriculture, and pastoralism to sustain themselves, which creates a diet that is high in fibre, as well as low in fats and sugars. These populations have very limited access to modern medicine. Communities with traditional dwellings live closer to one another, have pets, livestock, and wildlife nearby, which all have an impact on their collective and individual IMs. Western diets and lifestyles, on the other hand, consist of more fats and sugars and less fibre, but with better access to modern medicine. With westernization came a reduction in exposure to microbes associated with the environment (Rook *et al.*, 2014; Thorburn *et al.*, 2014). These different lifestyles vary in terms of their culture, dietary habits and their interaction with the environment around them, these varying factors influence the composition and functionality of the IM (Gomez *et al.*, 2016; Davenport *et al.*, 2017; Mancabelli *et al.*, 2017; Jha *et al.*, 2018; Vangay *et al.*, 2018). Traditional populations tend to have taxonomically diverse IMs with a higher abundance of bacteria that produce SCFAs, such as *Prevotella*, *Succinivibrio* and *Treponema* (De Filippo *et al.*, 2010; Davenport *et al.*, 2017). These microbes may be beneficial in preventing some health implications such as diarrhoea that were seen less of in Western populations (De Filippo *et al.*, 2010).

Various studies have suggested that transitioning to a Western lifestyle leads to an increase of certain taxa such as *Barnesiella* or the loss of others, for example *Treponema* (Mancabelli *et al.*, 2017). Taxa like *Treponema* are associated with high fibre consumption, which is characteristic of traditional dietary habits (De Filippo *et al.*, 2010; Ayeni *et al.*, 2018; Rosas-Plaza *et al.*, 2022). *Bacteroidaceae*,

Lanchospiraceae, and *Rickenellaceae* have seen an increase in transition to Western lifestyles (Rosas-Plaza *et al.*, 2022). The IM exhibits distinct beta diversity from hunter-gatherer to urban populations, according to a recent study (Rosas-Plaza *et al.*, 2022). In a study comparing the Western IM with BaAka hunter-gatherers and Bantu-speaking agriculturalists, a loss of microbial diversity within the Western IM was observed (Gomez *et al.*, 2016). By studying the traditional Hadza IM and westernized Italian IM, differences in functionality were observed between the two IMs (Rampelli *et al.*, 2015). The Hadza IM exhibited an increased capacity for carbohydrate metabolism, whereas the Italian IM emphasized sugar metabolism. A loss of taxonomic diversity was also demonstrated for South-east Asian immigrants in the US proportionally related to the time they have spent there (Vangay *et al.*, 2018). In contrast to these findings, the IMs of Himalayans across a lifestyle gradient were examined, all within similar geographic locations, and it was found that lifestyle has a minimal effect on the alpha diversity of populations (Jha *et al.*, 2018). The reasoning for the disparity between findings is that previous comparisons were those in Africa and South America being compared to those in the United States and Europe, representing vastly different geographic locations which may account for the significant differences in alpha diversity. Another study concerning individuals originating from close geographical locations observed similar alpha diversity distinct from their communities, i.e., Bassa or Urban Nigerian (Ayeni *et al.*, 2018). Individuals who consumed water from wells was reported as exhibiting higher alpha diversity compared to those who consumed bottled, tap or filtered water (Vanhaecke *et al.*, 2022). A study looking at IM differences across different lifestyles found that urban populations reported the lowest alpha diversity and that there were no significant differences between alpha diversities amongst the traditional populations (Rosas-Plaza *et al.*, 2022). Composition, alpha diversity, and functionality are just some aspects in which the traditional IM may differ from the Western IM.

Ju|’hoansi

Limited microbiome research has been carried out within the southern African context, and the Ju|’hoansi presents an opportunity to expand our knowledge in this region. Several studies have been conducted on the Ju|'hoansi bushmen of Namibia to better understand their hunting practices and to conserve as much game as possible (Lee, 1968; Koot, 2019). Their traditional practices that they preserved from their ancestors have been the subject of some studies, as well as their archaeology (Smith, 2001; Wiessner, 2014; Ninkova, 2020). In regard to their microbiomes, there is a knowledge gap.

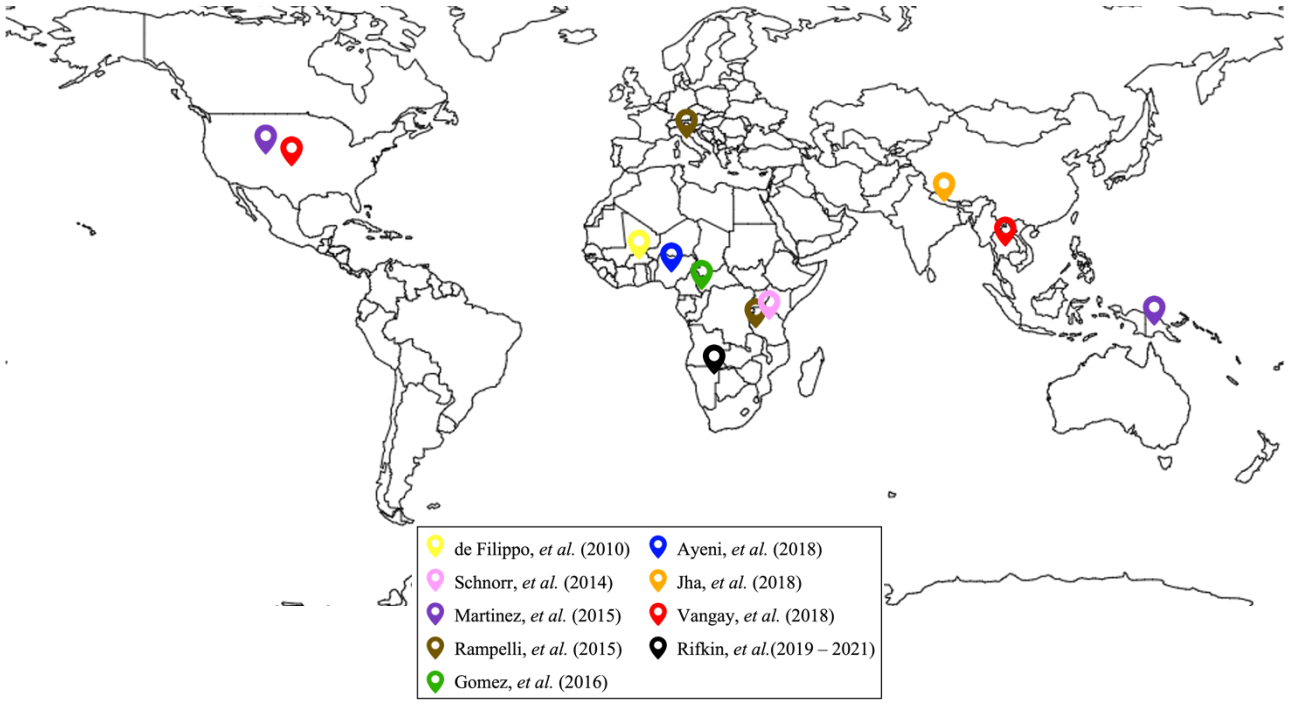


Figure 1. Geographic locations and populations represented in existing studies on lifestyle effects on the human IM.

The Ju|'hoansi hunter-gatherers of the Nyae Nyae Conservancy (NNC) in Namibia are a largely traditional community that survive on a non-industrialised rural subsistence-based lifestyle. The minimal Western influence seen here presents itself in the form of few shops, a school, and the occasional use of antibiotics when members of the community contract tuberculosis. In a southern African context, the Ju|'hoansi are uniquely positioned to contribute to the study of the evolution of the IM *in situ*. This community therefore provides a distinctive opportunity to detect the influence of the cumulative consumption of Western food sources, the use of antibiotics, increasing exposure to toxic pollutants and the consumption of unfiltered borehole water on their largely traditional IM. This transition to a more Western way of living renders the Ju|'hoansi suitable for a temporal study on the effect of westernization on the traditional human IM. Studying the transition over a time period may allow insight into any effects that it may have on future IM composition or functionality. A more Western population may see these changes in the IM that have been brought about by past westernization. There are some from the Ju|'hoansi population that have crossed Namibia's borders and therefore have experienced international travel. Should they travel to a high prevalence antimicrobial resistance zone, such as South Africa in some cases, they could be responsible for bringing back resistant bacteria and spread it to the rest of their respective village (Frost *et al.*, 2019). This is just one impact international travel could have on the Ju|'hoansi.

Figure 2. Map showing the location of the Nyae Nyae Conservancy in Namibia.

Our study focuses on the faecal analyses collected in 2019 and in 2021 from 40 participants from four Ju|'hoansi villages in NNC, i.e., Duinpos, Den/ui, Mountain Pos and !Om!o!o,. During the dry season (winter) in July 2019 we collected faecal samples, we sampled a second time in July 2021. As a result of the global COVID-19 pandemic, our sampling season in 2020 was interrupted, resulting in a two-year hiatus. Thus, we analysed two sets of faecal samples, for a total of 77 samples, collected from 40 Ju|'hoansi, living in the NNC. First, we aimed to determine whether any changes could be detected in the taxonomic composition of the core Ju|'hoansi IM, over a period of two years. In this regard, we discern between our original 'baseline' (2019) and 'follow-up' (2021) sampling periods, interrupted by a two-year hiatus. Second, we aimed to determine whether these observed changes could relate to five variable biological and abiotic environmental categories, namely 1) biological sex, 2) chronometric age, 3) residential mobility, 4) medical history and 5) the use of distinct unfiltered water sources.

Conclusion

There has been a multitude of research to show that the transition to a Western diet and lifestyle has historically had an impact on the composition and functionality of our IM. Whether or not this impact has been beneficial or detrimental, remains to be investigated. The vast number of studies concerned with which factors make significant contributions to the shaping of our IM have all had a narrow field

of interest, regarding diet, geographical location or, to a lesser extent, drinking water sources. One way to increase our knowledge of the repercussions of westernization on our IM is to analyse traditional hunter-gatherer communities undergoing westernization, and the effects on their IM. Those societies transitioning to a more Western lifestyle presents a window of opportunity to study that transition and its effects in the human body. These studies need to be conducted presently before the Western influence reaches remote traditional societies making research of this variety impossible. Research conducted in sub-Saharan Africa is severely lacking with regards to this field of interest, proving a study of this nature a valuable contribution to the field of human intestinal microbiomics.

Aim and Objectives

Aim

The primary aims of this project is to 1) detect temporal changes in the taxonomic composition of the intestinal microbiome of the Ju|'hoansi and 2) to determine whether water sources might contribute to the differences in the taxonomic composition of the Ju|'hoansi intestinal microbiome by using bioinformatics.

Objectives

1. Record changes in the IM of the Ju|'hoansi from a baseline date of July of 2019 to follow-up in July 2021,
2. Explore the correlation between bacterial taxonomic abundance and measured volatile fatty acid (VFA) levels, and
3. Examine the association between the consumption of unfiltered borehole drinking water and observed variability in the Ju|'hoansi IM.

Chapter 2: Temporal change in the Namibian Ju|'hoansi intestinal microbiome

Tiffany du Plessis[1], Jessica E. Koopman[1], Karen Jordaan[1], Jean-Baptiste Ramond[1,2], Diego Castillo[1], Leon ≠Oma Tsamkxao[3], Simon J. Underdown[1,4], Riaan F. Rifkin[1]

[1]Centre for Microbial Ecology and Genomics, Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Hatfield, South Africa

[2]Department of Molecular Genetics and Microbiology, Pontificia Universidad Católica de Chile, Santiago, Chile

[3]Ju|'hoan Traditional Authority (JUTA), Tsumkwe, Otjozondjupa Region, Namibia

[4]Department of Anthropology and Geography, Human Origins and Palaeoenvironmental Research Group, Oxford Brookes University, Oxford, UK

Summary

The microbes that reside in our digestive tract comprise the human intestinal microbiome (IM). Many historical events have influenced the composition of our IM, such as the invention of cooking, the advent of farming, and the Industrial Revolution. Although ancestral IMs could provide insight into the evolution of the human IM, pristine ancestral hunter-gatherer IM samples are near-impossible to acquire. Alternatively, studying contemporary hunter-gatherers might allow us to gain insight into the pre-industrial human IM. The Ju|'hoansi hunter-gatherers of Namibia are transitioning to an increasingly Western lifestyle, providing an opportunity to study the effects of westernization on human IM. This study aims to detect temporal changes in the Ju|'hoansi IM. In order to do so, faecal samples were collected and subsequently sequenced from a total cohort of 40 participants taken 2 years apart in 2019 and 2021. As a result of our study, we were able to identify changes in the composition of microbial communities over a 2-year period. Statistically significant differences were found in the taxonomic composition of the Ju|'hoansi IM after 2 years. Within each year, the village of residency of each participant proved significant. The richness of the Ju|'hoansi IM did not change during this time period and still reflected that of a traditional hunter-gatherer IM. From baseline to follow-up, numerous bacterial genera differentially decreased, and the core microbiome significantly decreased in number of species. This draft manuscript will be sent to Cell Press Community Review intended for publication a in Cell Press Journal.

Introduction

In the human body, the human microbiome consists of all microbes, including bacteria, viruses, and fungi (Gilbert *et al.*, 2018; Shanahan *et al.*, 2021). The human intestinal microbiome (IM) is estimated to comprise of 200 to 1000 bacterial species per individual (Lloyd-Price *et al.*, 2016; Gilbert *et al.*, 2018).

Human homeostasis depends upon maintaining the IM, and the metabolic functionality of the taxonomic communities present in the IM affects the health of its host in notable ways (Bäckhed *et al.*, 2012; Davenport *et al.*, 2017; Das & Nair, 2019). Individuals and populations display significant differences in taxonomic composition, as there is high inter-individuality in the IM. As a result, it is difficult to define a healthy human IM (Walter & Ley, 2011).

Although the lifestyle changes associated with the Neolithic Revolution significantly impacted hunter-gatherer IM taxonomic composition and metabolic capacity (Walter & Ley, 2011; Quercia *et al.*, 2014), the impact of the Industrial Revolution, and the process of global 'westernization' on the human IM, is particularly marked (Blaser & Falkow, 2009; Gillings & Paulsen, 2014; Segata, 2015; Schnorr *et al.*, 2016; Rifkin *et al.*, 2020). IM composition has been assessed over time, allowing researchers to gain a better understanding of health and disease as well as the effects of host-microbiome interactions (Faith *et al.*, 2013; Mehta *et al.*, 2018). It has been found that members of the phyla Bacteroidota and Actinobacteria remain more stable compared to Firmicutes and Proteobacteria (Faith *et al.*, 2013). It has also been shown that alpha diversity and IM composition differ significantly and that diverse microbial communities seem to be more stable over the long term (Chen *et al.*, 2021b). Numerous factors influence both the composition and functionality of our IM, and these should be explored in more detail over time and among different cultures and populations.

The analysis of traditional hunter-gatherers' IMs that are undergoing westernization may enable us to gain a better understanding of the consequences of westernization on our IM. In order to record changes in the IM of the Ju|'hoansi from a baseline date of July of 2019 to a follow-up in July 2021, we analysed faecal samples collected from the same 40 Ju|'hoansi participants from four villages in Tsumkwe, i.e. Duinpos, Den/ui, Mountain Pos and !Om!o!o. This study examined changes in taxonomic composition between baseline and follow-up Ju|'hoansi IMs as well as the influence of biological and geographic factors on their IM.

Results

Cohort description

Faecal samples were collected in 2019 from 40 Ju|'hoansi hunter-gatherers in Namibia, and 37 participants in 2021, to determine the IM variability over 2 years. Participants consented to participate in this study and agreed to collect the samples. Of the 40 participants, 20 were male, and 20 were female, ranging in age from 19 to 71 years ($\bar{x} = 39.65$ years). Three participants withdrew from the 2021 sampling period, leaving 37 participants for the follow-up analysis. These samples were processed, and the sequences were analysed with the QIIME2 pipeline (Bolyen *et al.*, 2019). In order to understand the results of the samples, a questionnaire was completed by participants in 2019 and 2021. The questionnaire was designed to determine any potential changes in lifestyle that may have occurred over the 2-year period. Interestingly, changes in lifestyle did occur. For example, a number of participants have travelled internationally since 2019, and several participants changed their dietary habits to include less store-bought food during winter by 2021 (SI Table 1). The data collected from the questionnaire and the IM analysis of the samples allowed for a more comprehensive understanding of the results. The breakdown of lifestyle habits and changes that may have occurred over the 2-year period are available in (SI Table 1).

Ju'|hoansi IM composition

A baseline taxonomic IM characterization was established in 2019, and a follow-up analysis was conducted in 2021. The total number of reads from our baseline was 4,679,902, whereas our follow-up had 3,037,238 reads. This resulted in an average number of 111,426 and 77,878 reads per sample in baseline and follow-up analyses, respectively. After quality control with DADA2 (Callahan *et al.*, 2016), we were left with 4,523 ASVs from baseline and 3,925 ASVs from follow-up. In Piquer-Esteban (2022), multiple core thresholds were defined to explore less prevalent core taxa, these were hard-core at a prevalence of 90%, medium-core with a prevalence of 70% and a soft-core at 50%. Using core statistics employed by Piquer-Esteban (2022), no hard-core microbiome was defined at a 90% prevalence of bacterial genera in either baseline or follow-up. Regarding a medium-core microbiome with a 70% prevalence, 6 taxa were present in the baseline but none in the follow-up. For a soft-core microbiome with a 50% prevalence, 25 taxa were present in baseline and 2 taxa in follow-up (SI Tables 3.1 - 3.2). Using the same relative abundance threshold of 1×10^{-4} from Piquer-Esteban (2022), we set a detection threshold of 0.01%. Figure 3 shows the soft-core microbiome for 2019 and 2021.

In our baseline sampling, Firmicutes (24.73%) and Bacteroidota (12.51%) were the dominant taxa with an approximate Firmicutes:Bacteroidota ratio of 2:1. Following Firmicutes and Bacteroidota, the top three subsequent dominant taxa comprised Proteobacteria (3.6%), Spirochaetota (0.37%), and

Actinobacteriota (0.32%). A total of 142 genera were identified in 2019, with Prevotella (8.67%), Blautia (2.58%), and Succinivibrio (2.19%) being the most abundant. During our follow-up sampling, Firmicutes (15.91%) and Bacteroidota (14.17%) were the dominant taxa, with an approximate Firmicutes:Bacteroidota ratio of 1:1. The top 3 subsequent dominant taxa were Proteobacteria (4.62%), Fusobacteriota (2.16%), and Patescibacteria (1.11%). One hundred forty-four genera were identified in 2019, with Prevotella (9.71%), Streptococcus (3.31%), and Veillonella (3.27%) being the most abundant (SI Tables 2.1.1 - 2.6.2).

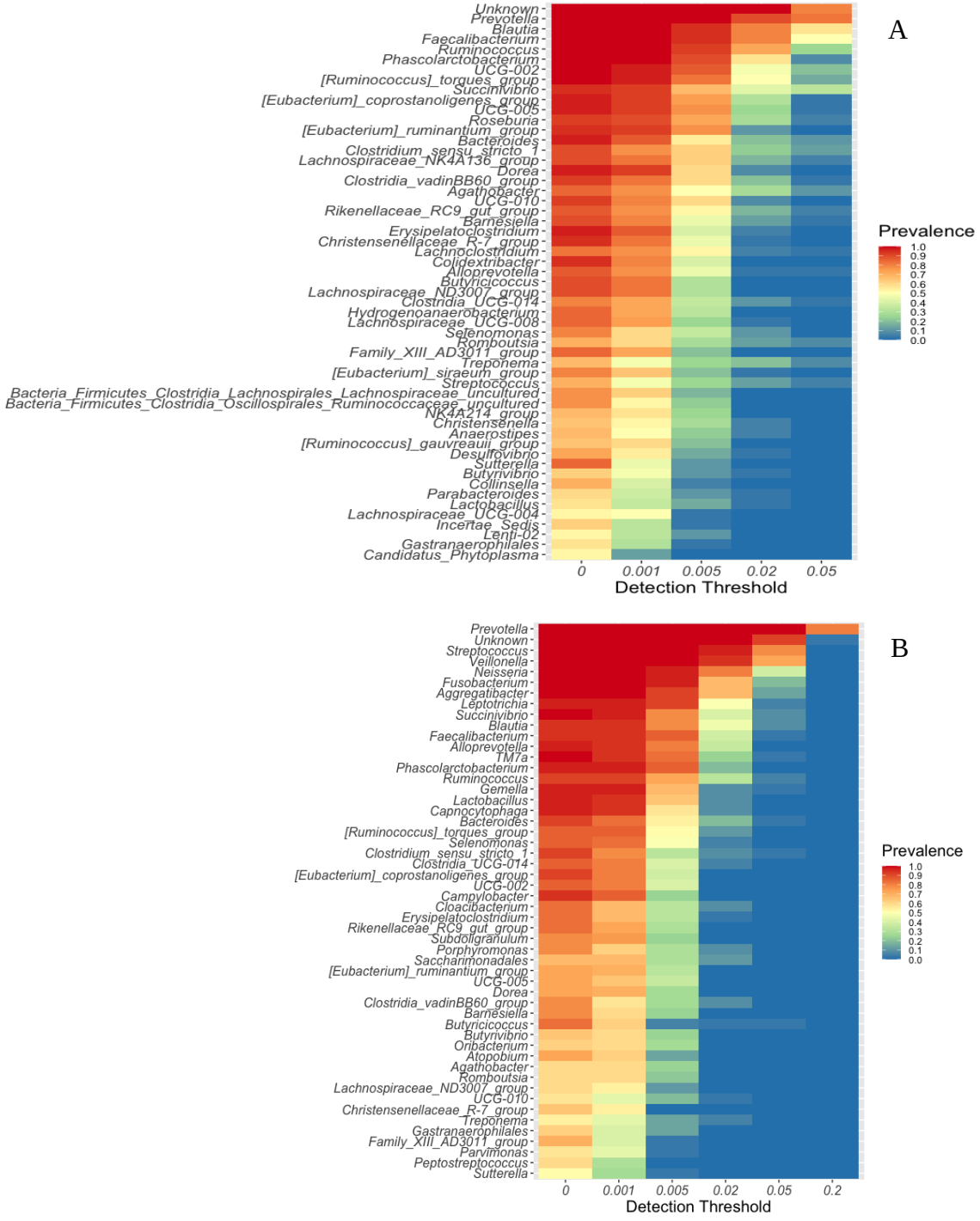


Figure 3. The bacterial core microbiome of the Ju|'hoansi from A) baseline to B) follow-up at differing prevalence and detection thresholds.

Temporal changes in IM microbial diversity

A microbial composition and diversity analysis was conducted to characterize the changes in the Ju|'hoansi IM over a 2-year period. As per the Shannon diversity index, the microbial species richness did not change from the baseline to the follow-up data ($p_{\text{Wilcoxon}} = 0.219$), yet there were significant differences in taxonomic abundances between the years as indicated by the Simpson diversity ($p_{\text{Wilcoxon}} = 0.005$), as is evident in Figures 4 and 5.

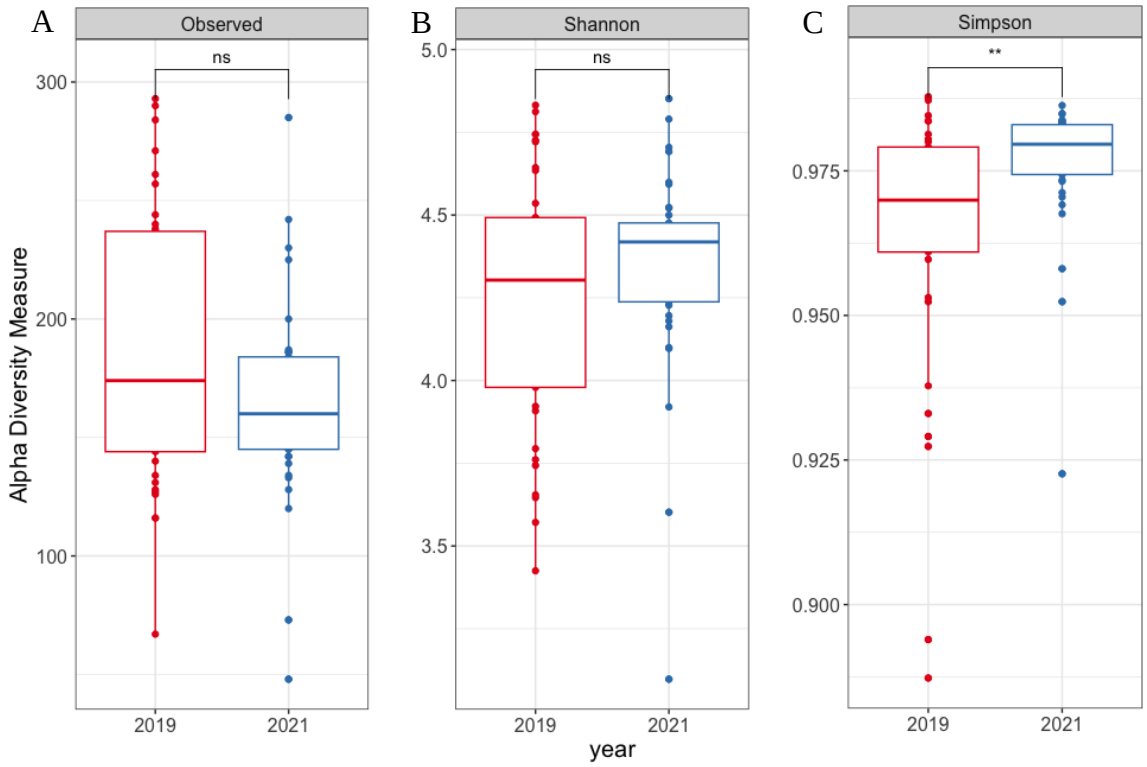


Figure 4. Per-year α-diversity indicated with A) Observed, B) Shannon and C) Simpson indices. Baseline is represented in red and follow-up in blue. Statistical significance shown above the boxplots indicating either not significant (ns) or significant at <0.01 (**).

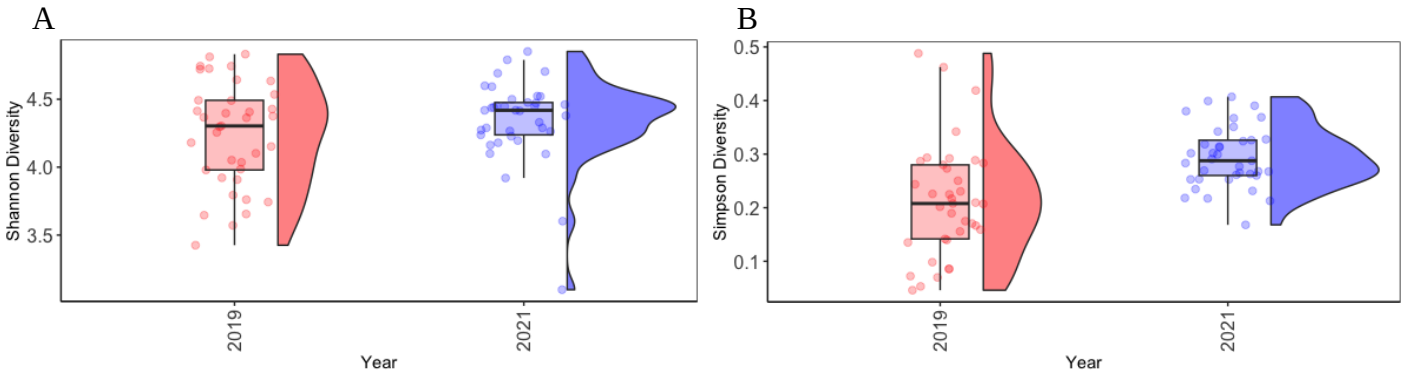


Figure 5. Combination violin and boxplots indicating A) Shannon diversity indices for baseline in red and follow-up in blue and B) Simpson diversity indices for baseline in red and follow-up in blue.

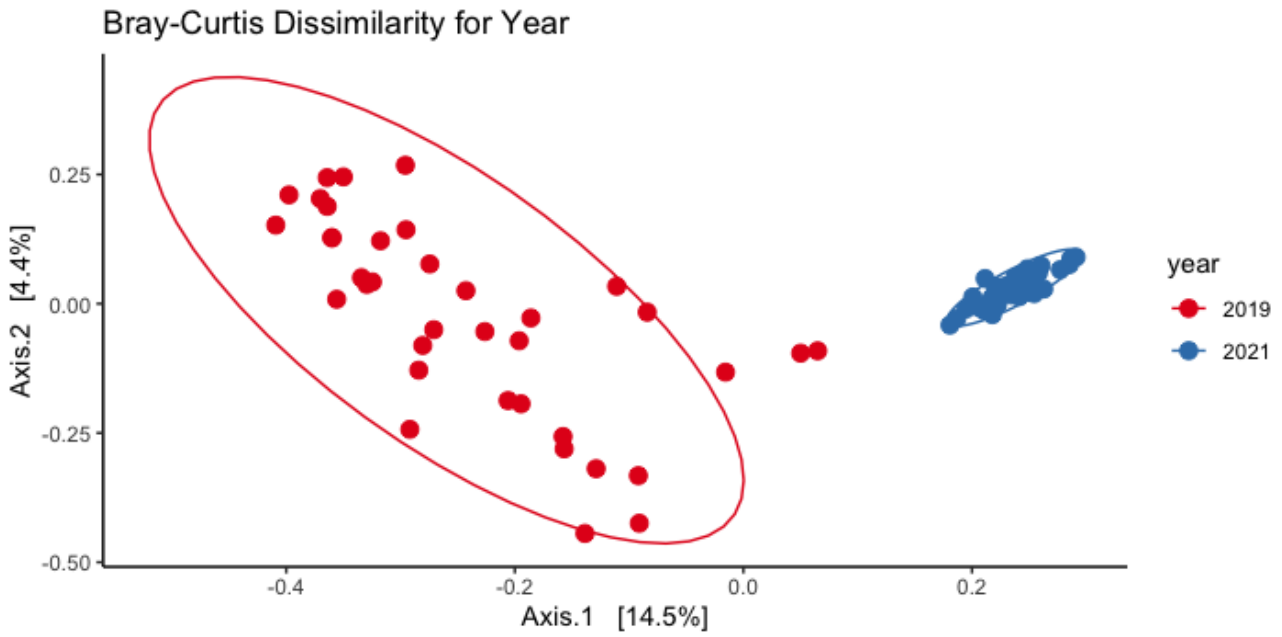


Figure 6. Principal coordinate analysis (PCoA) plot with Bray-Curtis dissimilarity between sampling years. Baseline is represented in red and follow-up in blue.

When looking at the sampling from baseline and follow-up together, many factors showed statistical significance with regards to Bray-Curtis dissimilarity. The year that sampling was done proved very statistically significant ($p = 1 \times 10^{-4}$) (Figure 6), along with travel (p = 0.0032), dietary preferences in winter ($p = 1 \times 10^{-4}$), treatment of intestinal infections ($p = 4 \times 10^{-4}$), contraction of tuberculosis (TB) (p = 0.0014), as well as use the of antibiotics ($p = 1 \times 10^{-4}$) (SI Table 5.1). Noteworthy is that the different villages and history of intestinal infections did not prove significant when looking at β-diversity. The different villages did prove as statistically significant ($p = 8 \times 10^{-4}$) when considering the β-diversity during our baseline study in 2019. Not as significant in baseline was the participants' dietary preferences in winter (p = 0.0306), whether or not they experienced intestinal infections (p = 0.0423) and treatment therefore (p = 0.033), as well as if they had ever contracted TB (p = 0.0441) (SI Table 5.2). In the analysis of our follow-up sampling for 2021, only 2 factors turned up significant when looking at the β-diversity. Village (p = 0.0473) and contraction of TB (p = 0.0387) showed some significance but less in comparison to the baseline (SI Table 5.3). Aside from the above-mentioned factors, none other proved significant when looking at β-diversity.

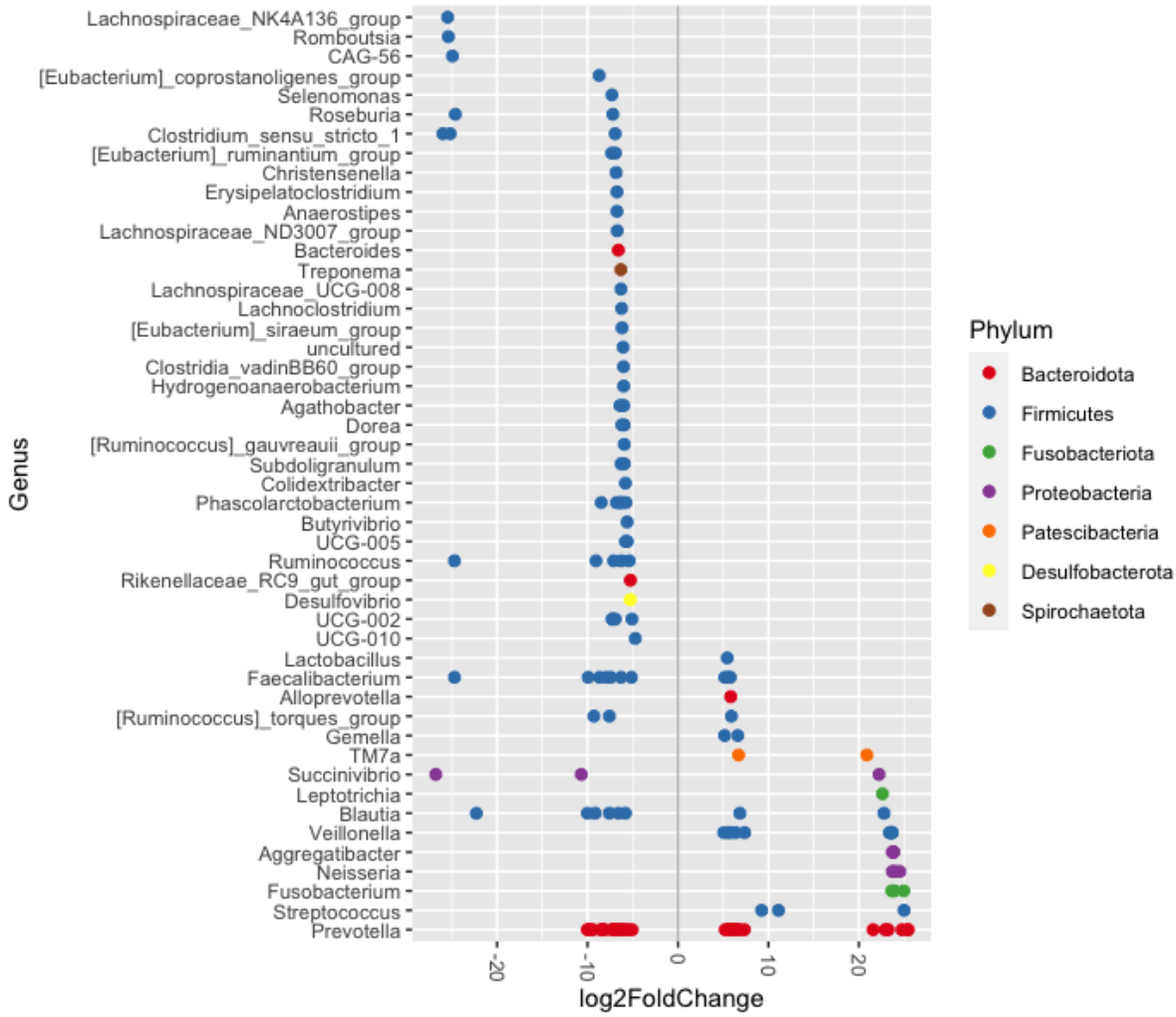


Figure 7. Bar plot showing the log2-fold-change of the differentially abundant genera and correlating phylum. Points on the left-hand side of the zero line show genera that have decreased in abundance in follow-up since baseline, those on the right-hand side have increased since.

DeSeq2 (Love *et al.*, 2014) was used to determine statistically significant differentially abundant genera between factors. Some genera were statistically different between participants from different villages, of different biological sex, travel habits, dietary preferences in winter, history of intestinal infections, malaria, TB, and antibiotics use (SI Tables 6.2 - 6.9). Since there were statistically significant differences in α- and β-diversity between the different sampling years, I focused on the differential abundance genera between years (Figure 7, SI Table 6.1). The differentially abundant bacterial genera included *Prevotella*, *Bacteriodes*, *Faecalibacterium*, *Ruminococcus*, *Roseburia*, *Blautia*, *Fusobacterium*, *Streptococcus*, *Christensenella*, *Veillonella*, *Clostridium_sensu_stricto_1*, *Romboutsia*, *Succinivibrio*, *Aggregatibacter*, and *Neisseria*.

Discussion

In order to detect temporal changes in the Ju|'hoansi IM, faecal samples were collected and subsequently sequenced from a total cohort of 40 participants taken 2 years apart in 2019 and 2021. From the analysis of the IM sequences, we detected no difference in species richness over the sampled time period. However, a change in species evenness was observed. Additionally, other factors proved statistically significant when looking at the diversity of the Ju|'hoansi IM, including the village in which participants resided.

Species richness did not change from baseline to follow-up over the 2-year time period, yet evenness did show a statistically significant change. Species evenness looks at the abundance of species within a specific community. In contrast to this, species richness refers to the number of different species within the community and does not look at the abundances thereof (Lloyd-Price *et al.*, 2016). Looking at the differentially abundant genera, we can see how taxa evenness changed from baseline to follow-up (SI Table 6.1). The core microbiome taxa found in baseline to follow-up decreased dramatically as well as the abundance thereof (SI Tables 3.1 - 3.2). Using the core statistics employed in Piquer-Esteban (2022) as well as the same relative abundance threshold of 1×10^{-4}, a soft-core microbiome at a 50% prevalence was present with 25 taxa in baseline and 2 taxa in follow-up. A large proportion of the baseline core microbiome, i.e., 9 out of 25 taxa, were part of the *Prevotella* genus. In the same regard, the 2 taxa present in the follow-up core microbiome were also part of the *Prevotella* genus. Some studies have seen a loss of *Prevotella* in Industrial or Western IMs and suggest this may contribute to the rise in non-communicable diseases (Jacobson *et al.*, 2021). Thus, the prominence of *Prevotella* in traditional IMs such as the Ju|'hoansi, Hadza or the Burkina Faso could be a positive factor (De Filippo *et al.*, 2010; Schnorr *et al.*, 2014; Ayeni *et al.*, 2018).

Fifteen of the 25 core taxa found in baseline belonged to the Firmicutes phylum (SI Table 3.1). No Firmicutes were found in the follow-up core microbiome which is further supported by the fact that the overall relative abundance of Firmicutes decreased from baseline to follow-up (SI Table 2.1.2, SI Table 3.2). The loss of Firmicutes abundances could be a result of the decrease in store-bought food preference in the Ju|'hoansi from 2019 to 2021. Studies have seen a similar trend with lower Firmicutes to Bacteroidota ratios in rural or traditional IMs (De Filippo *et al.*, 2010; Schnorr *et al.*, 2014; Ayeni *et al.*, 2018). This Firmicutes to Bacteroidota ratio differs from those with Western or urban IMs which presents a higher Firmicutes to Bacteroidota ratio (Gomez *et al.*, 2016; Ayeni *et al.*, 2018; Jha *et al.*, 2018). Firmicutes have been proposed as having a greater ability to extract energy from food than Bacteroidota, resulting in a more efficient calorie absorption (Krajmalnik-Brown *et al.*, 2012). This could explain why patients with obesity have a higher Firmicutes to Bacteroidota ratio (De Bandt *et al.*, 2011; Krajmalnik-Brown *et al.*, 2012). The subsequent dominant phyla after Firmicutes, Bacteroidota,

and Proteobacteria during baseline to follow-up changed from Spirochaetota and Actinobacteriota to Fusobacteriota and Patescibacteria.

Looking at Figure 7, we can see that the genera *Succinivibrio* has differentially decreased over the 2 years. *Succinivibrio* have the ability to ferment complex polysaccharides and are found in abundance in traditional IMs (Schnorr *et al.*, 2014; Angelakis *et al.*, 2019; Rosas-Plaza *et al.*, 2022); and decreased *Succinivibrio* has been linked to some diseases such as environmental enteric dysfunction (Chaima *et al.*, 2021) and Trichuris trichiura infection (Chen *et al.*, 2021a). Thus, this loss may have implications regarding the Ju'|hoansi health. In contrast, when looking at *Neiserria*, we see the genera was differentially enriched from baseline to follow-up (Figure 7). *Neisseria*'s role in traditional IMs is not well documented in literature, but an increase in *Neisseria* has been associated with psoriasis (Olejniczak-Staruch *et al.*, 2021) and chronic gastritis (Jiang *et al.*, 2021).

The different impacting factors between participants' IM differed from baseline to follow-up, village, dietary preferences in winter, history of intestinal infections, treatment thereof, contraction of TB, and use of antibiotics proved statistically significant between participants during baseline sampling. Whereas in follow-up, only 2 factors were significant, village and contraction of TB, and the different villages were not as statistically significant as in baseline. When looking at the data as one dataset with a total of 77 samples, the sampling year was statistically significant, along with travel, dietary preferences in winter, treatment of intestinal infections, contraction of TB, and use of antibiotics (SI Table 5.1). Although treatment of intestinal infections was statistically significant, the actual history of intestinal infections was not of statical importance, and neither was the village where each participant resided.

A possible reason for the villages to remain statistically significant from baseline to follow-up could be reliant on the villages' water sources. Each village has access to borehole drinking water in different locations. There have been a few studies concerning the impact of drinking water on our IM (Sofi *et al.*, 2014; Ayeni *et al.*, 2018; Bowyer *et al.*, 2020; Lugli *et al.*, 2022; Vanhaecke *et al.*, 2022). A study looking at the consumption of UK tap water suggested that the composition of the human IM may serve as mediators between tap water and human health (Bowyer *et al.*, 2020). Lifestyle gradients have been reported to influence the composition of the gut microbiota in the Himalayas, but researchers have been unable to explain why this is the case (Jha *et al.*, 2018). The role that drinking water has on our IM needs further exploration.

The significance of a history of intestinal infections, treatment thereof, contraction of TB, use of antibiotics in the baseline, and the contraction of TB in follow-up is not a novel finding. Several studies have found a link between health, treatment, and the composition of the IM (Hu *et al.*, 2019; Elvers *et*

al., 2020; Gomaa, 2020). A gut-lung axis has been suggested, involving cross-talk between the respiratory tract microbiome and the IM (Budden *et al.*, 2017; Hu *et al.*, 2019). One study found that certain bacterial species were differentially enriched in healthy subjects compared to participants who had contracted TB (Hu *et al.*, 2019). Antibiotics have been shown to affect the IM, and research continues on this topic in light of antibiotic resistance (Elvers *et al.*, 2020; Pennycook & Scanlan, 2021). Based on a review of studies regarding antibiotics and the IM, it was found that the IM recovered after the treatment caused a loss of diversity (Pennycook & Scanlan, 2021).

It is important to note that between 2019 and 2021, the global COVID-19 pandemic occurred in 2020. This may have impacted several factors concerning the Namibian Ju|'hoansi, including, but not limited to, their access to imported store-bought food and medication, as well as regional and international travel. We recognize this would be an important variable to consider, yet we also lack insight into the Ju|'hoansi during this time. From our sampling expeditions, we have no record of whether any of the participants had contracted COVID-19 over the time period. As previously mentioned, due to the global pandemic, we could not sample during 2020. Literature is also still exploring the effect that COVID-19 had on the human gut microbiome; thus, we cannot associate any changes we have found with that research.

As evidenced by their increasing access to shops, imported food, travel, and medicine, the Ju|'hoansi are transitioning to a more Western lifestyle. In turn, this results in a change in their IM, one that might start to mimic the IM of Western populations. In our study, we aimed to provide insight into this transition; however, further studies over longer time periods would contribute to a more comprehensive understanding of westernization. Understanding the effects of westernization could help determine what shifts the Western IM might have undergone due to such a lifestyle, providing insight into the rise of certain diseases in Western societies, such as inflammatory bowel disease, obesity, or autoimmune diseases.

Conclusion

Faecal samples from a cohort of 40 participants were collected and sequenced to detect changes in the Ju|'hoansi IM between 2019 and 2021. In order to do this, background information was collected about each participant from their biological data, i.e., age, and sex, to history, i.e., travel, medical history, and the village of residency. According to our investigation into the temporal changes of Ju|'hoansi IM, species richness over the 2-year period did not change, but species evenness did. Furthermore, the village where participants lived had a statistically significant impact on the composition of their IM from year to year. The core microbiome decreased substantially in species from baseline to follow-up, and many bacterial genera were differentially decreased. This study examined the possible implications of westernization, but a global pandemic may have affected its findings between the sampling years. The findings of our study demonstrated that changes in microbial composition could be identified over a 2-year period. This study was part of an ongoing study that aimed to characterize the Ju|'hoansi IM and intended to identify any changes that may have occurred to the Ju|'hoansi IM over a 2-year period.

Materials and Methods

Sampling

Data for this research project was obtained in July 2019 and July 2021, during the dry winter season, from four Ju|'hoansi villages in Nyae Nyae, i.e., Duinpos, Den/ui, Mountain Pos and !Om!o!o, which are located 18 km to 28 km ($\bar{x} = 23.3$ km) from Tsumkwe, Otjozondjupa's largest urban center.

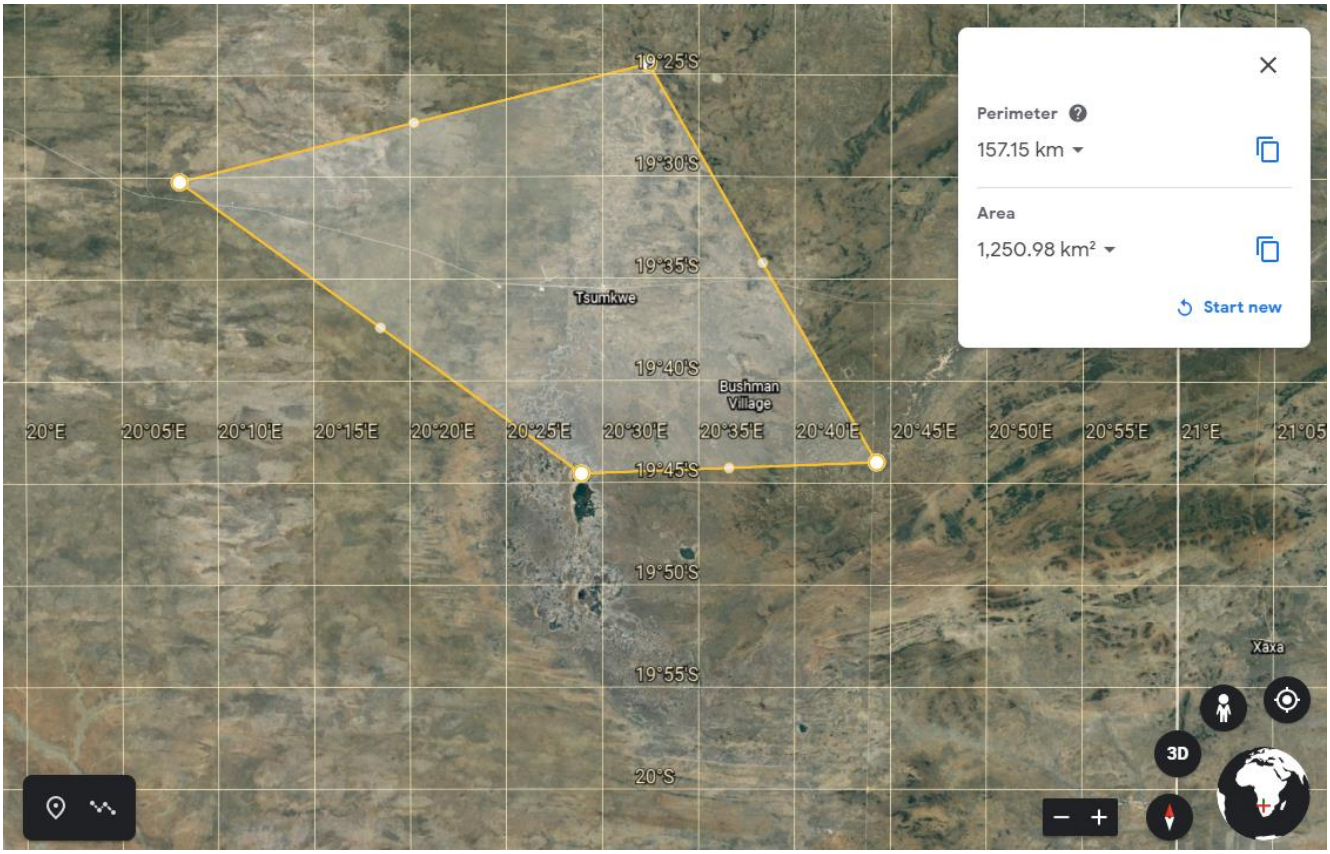


Figure 8. Map showing the sampling areas in the Nyae Nyae Conservancy, Tsumkwe, Namibia.

The COVID pandemic interrupted the initial plan to collect samples bi-annually, i.e., in summer and in winter, which is why this study focused on the Ju|'hoansi IM in the winter season from the available samples. Following informed consent, faecal samples were acquired from a total participant cohort of 40 adults ($n = 20$ males, $n = 20$ females) ranging in age from 19 to 71 years ($\bar{x} = 39.65$ years). Of the 40 participants in 2019, only 37 participated in 2021. Our co-researcher, Leon ≠Oma Tsamkxao, who is fluent in Ju|'hoansi, Afrikaans, and English, assisted in facilitating the study, recruiting research participants, and obtaining written consent. Consent was provided for the publication of results, agreeing that all information acquired, excluding names, may be disclosed. This study was approved by the Research Ethics Committee, Faculty of Health Sciences at the University of Pretoria, South Africa (Protocol number TEMP 2017-01469). The research permit for this study was approved by the Namibian National Commission on Research and Technology (NCRST) (RPIV00692019). The Ju|'hoansi Traditional Authority (JUTA) provided consent for the enrolment of project participants from the Nyae

Nyae Conservancy. It has been approved that all information necessary for the study, except for the participants' names, may be disclosed in this study. Participants provided their consent to publish the study results of the obtained biomaterials. The methods were all followed according to the Helsinki Declaration. The metadata collected from participants comprised data about their age, use of antibiotics in the treatment of TB, biological sex, experiences of diarrhoea in consumption of certain foods, intestinal infections, use of malaria medication, local, regional, and international travels, and the village of residency for each participant. During the winter season, when hunting and foraging are less important, the Ju|'hoansi purchase food from various shops available to them in Tsumkwe. Starches like maize, rice, and pasta have therefore become an integral part of their diet, as well as meats like beef and goat.

The faecal samples were collected in collection tubes containing 9 mL DNA/RNA ShieldTM (Zymo Research Corp, Irvine, CA, USA) and stored at 4ºC until further processing. In order to sequence the V3-V4 regions of 16S rRNA, samples were sent to Applied Biological Materials Inc., Richmond, B.C. Canada for paired-end (2 x 300bp) sequencing using the MiSeq Illumina platform (Illumina, San Diego, CA, USA). Additionally, 2 controls were included with each year, namely KIT-CTRL and CON-CTRL. CON-CTRL consisted of DNA/RNA ShieldTM (Zymo Research Corp, Irvine, CA, USA). KIT-CTRL contained the DNeasy® PowerLyzer® PowerSoil® Kit (Qiagen GmbH, Hilden, Germany).

Bioinformatic pipeline

The raw sequence reads (from Applied Biological Materials Inc., Richmond, B.C. Canada) were imported into QIIME 2 v.2022.2 (Bolyen *et al.*, 2019) and demultiplexed using the q2-demux plugin. Subsequently the paired-end sequences were denoised using DADA2 (Callahan *et al.*, 2016) with the q2-alignment plugin at varying truncation lengths; f250/r250, f200/r200, f270/r270, f230/r230, f250/r230. The trim lengths were kept the same throughout, f26 and r24. This generated representative sequences, stats, and a table for each output. In the same script, the metadata file was tabulated for each year, 2019 and 2021, generating a visual QIIME 2 file (.qzv) for stats. A feature table was created using the previous table output and metadata.txt. In the same script, a .qzv file was generated for representative sequences. These steps were also done on merged data where both years were included. The sequences were aligned using mafft (Katoh *et al.*, 2002), and subsequently, a phylogenetic tree was produced using fasttree2 (Price *et al.*, 2010) with the q2-phylogeny plugin for downstream analysis. In order to produce taxonomy, a naïve Bayes taxonomic classifier was trained using q2-feature-classifier (Bokulich *et al.*, 2018) classify-sklearn against the SILVA database (Quast *et al.*, 2012) reference sequences. Once the data were taxonomically classified, the CSV files from the taxa bar plots output were used to identify which taxa were present from domain to the species level. The relative abundances were calculated for the data from phyla to species level. The phylogenetic trees were rerun using IQ-Tree (Nguyen *et al.*,

2015) to produce a more accurate phylogeny. Feature tables were normalized to relative abundance in R (R Core Team, 2021; Rstudio Team, 2022), and contaminants were then removed using the Decontam package (Davis *et al.*, 2017) at a prevalence threshold of 0.1. Decontam determined whether reads were contaminants by comparing their prevalence between the controls, CON-CTRL and KIT-CTRL, and the rest of the samples, also known as the true samples. No contaminants were identified in both the 2019 and 2021 datasets.

Analysis in R

Once the data was imported into R as a phyloseq object (Mcmurdie & Holmes, 2013), the data were subsequently rarefied to 5469 reads, which was the minimum number of reads detected in a sample. After rarefaction, only 4.55% of OTUs were removed from the merged dataset. Any reads identifying as belonging to mitochondria or chloroplasts were filtered out. Sequencing reads only occurring once in the dataset were discarded, as well as the 3 participants that withdrew from the 2021 sampling because of the small cohort. We rarefied the data because of the discrepancy between the 2019 raw sequences and the 2021 sequences. Using rarefaction, we normalized the data by randomly subsampling the sequences from each year to a specific library size. Using set.seed, this random subsampling was kept consistent with each run of the R script. The core microbiome was determined using the microbiome package in R (Lahti *et al.*, 2017) and the full cohort of 40 participants for 2019 and the 37 remaining participants for 2021. Multiple packages in R (Rstudio Team, 2022) were used in order to explore the α- and β-diversity predictions for the data. α-diversity was determined using the Shannon and Simpson diversity indices based on the genera and using R packages; phyloseq (Mcmurdie & Holmes, 2013), ggpubr (Kassambara & Kassambara, 2020), microbiomeutilities (Shetty *et al.*, 2018), ggplot2 (Wicham, 2016) and stats (R Core Team, 2021). β-diversity was determined using phyloseq, ggplot2, microbiome (Lahti & Shetty, 2017), and vegan (Dixon, 2003). Figures for Results were made using pheatmap (Kolde, 2019), ggplot2, phyloseq, ggpubr, microbiomeutilities, and microbiome.

Acknowledgements

We would like to thank the participants of this study as well as the Ju|'hoansi Traditional Authority for consenting to the research. In issuing the research permit, we thank Loide Uahengo and Edgar Mowa (Namibian National Commission on Research and Technology) and Benetus Nangombe (Namibian Ministry of Health and Social Services). Thank you to Helvi Elago (National Heritage Council of Namibia) and Alma Nankela (Archaeology Unit, National Heritage Council of Namibia) for providing guidance and supporting documentation for our permit application. We thank the members of the Research Ethics Committee, Faculty of Health Sciences at the University of Pretoria, for providing clearance to conduct the research. Our appreciation goes out to our funding source, the Benjamin R. Oppenheimer Trust for financially supporting this research.

Author Contributions

Conceptualisation, ethics approval, sample collection, interview conduction, RFR, LOT, and SU; DNA extraction, JEK; Bioinformatics analysis, TD, and KJ; Writing, TD, RFR, and KJ; Figure creation, TD; Supervision, FJ, and RFR.

Declarations of Interest

None of the authors had any conflicts of interest.

Code Availability

All R code is available in Supplementary Information and Scripts used to analyse the data on QIIME2 are available at https://github.com/enchantiff/Kalahari-IM-Comparison

Chapter 3: The effect of unfiltered borehole drinking water on the Namibian Ju|’hoansi intestinal microbiome

Tiffany du Plessis[1], Jessica E. Koopman[1], Karen Jordaan[1], Jean-Baptiste Ramond[1,2], Leon ≠Oma Tsamkxao[3], Simon J. Underdown[1,4], Riaan F. Rifkin[1]

[1]Centre for Microbial Ecology and Genomics, Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Hatfield, South Africa

[2]Department of Molecular Genetics and Microbiology, Pontificia Universidad Católica de Chile, Santiago, Chile

[3]Ju|'hoan Traditional Authority (JUTA), Tsumkwe, Otjozondjupa Region, Namibia

[4]Department of Anthropology and Geography, Human Origins and Palaeoenvironmental Research Group, Oxford Brookes University, Oxford, UK

Summary

The human intestinal microbiome (IM) plays an important role in maintaining homeostasis in the human body. The composition of the IM can influence its host's health. The IM is known to have high inter-individuality when it comes to the taxonomic composition, and many factors can influence this from diet to geographic location. In an effort to explain the observed variability of the Ju|'hoansi hunter-gatherers of Namibia, we analysed their IM in conjunction with the microbiome and chemical analysis of their unfiltered borehole drinking water. Furthermore, the volatile fatty acid (VFA) levels were measured with the aim of exploring a potential correlation with bacterial taxonomic abundance of the Ju|'hoansi IM. Based on our findings, only one VFA proved significant in addition to multiple chemical composition factors of the borehole drinking water. This study forms part of an ongoing study characterizing the Ju|'hoansi IM.

Introduction

It is estimated that anywhere from 200 to 1000 bacterial species reside in one's intestine, where they are part of the intestinal microbiome (IM) (Lloyd-Price *et al.*, 2016; Gilbert *et al.*, 2018). In the human body, the IM is significant when it comes to maintaining homeostasis, and the metabolic functionality of its taxa influences the health of the host (Davenport *et al.*, 2017; Selber-Hnatiw *et al.*, 2017; Das & Nair, 2019). An individual's IM composition changes throughout their lifetime, influenced by changes in health or diseases (Odamaki *et al.*, 2016; Chen *et al.*, 2018). Many factors influence the IM, such as diet, the use of medication or even seasonality (Schnorr *et al.*, 2014; Smits *et al.*, 2017; Kurilshikov *et al.*, 2021).

The microbes comprising our IM produce metabolites such as short chain fatty acids (SCFAs), which assist in regulating homeostasis in the human body (Feng *et al.*, 2018). SCFAs are a byproduct of the fermentation processes of food undergone by IM microbiota, these SCFAs are absorbed and used as energy sources (Torii *et al.*, 2010; Silva *et al.*, 2020; Jacobson *et al.*, 2021). IM-produced SCFAs has been a research focus in recent years (Sanna *et al.*, 2019; Jacobson *et al.*, 2021; Oliver *et al.*, 2021; Vitale *et al.*, 2021) and have been proposed to be manipulated as a form of treatment for diseases caused by IM dysbiosis (Dinan & Cryan, 2017). The main SCFAs are acetate, propionate, and butyrate (Silva *et al.*, 2020; Jacobson *et al.*, 2021). Several human IM metagenomes were analysed to assess the metabolic potential for SCFAs and the impact of lifestyle thereon (Jacobson *et al.*, 2021). In terms of their production of SCFA, non-industrial IMs were shown to have more resilient microbiota, along with an increased ratio of acetate:butyrate:proprionate seen in these populations (Jacobson *et al.*, 2021). A study which examined the impact of dietary fibre on the composition of the IM over 2 weeks, concluded that this had a significant effect on composition, but not on the abundance of SCFAs (Oliver *et al.*, 2021). The impact of diet, drinking water and other factors on the IM composition should have a cause and effect on SCFA levels in the human body over an extended time period.

The influence of diet on the IM has been the focus of many past studies, whereas the role of drinking water on our IM needs further exploration. Few studies explore the impact of the source of our drinking water on our IM (Vanhaecke *et al.*, 2022). Possible associations between the available tap water in the UK and the composition of the IM in UK residents was studied and it was suggested that the composition of the human IM may serve as mediators between tap water and human health (Bowyer *et al.*, 2020). They found that an increased consumption of sodium in drinking water led to a reduction in species richness, along with sulphate and chloride concentrations affecting IM composition. Research concerning lifestyle gradients effect on the IM in the Himalayas reported the source of drinking water was associated with the composition of the gut microbiota but could not efficiently explain the

functioning behind this (Jha *et al.*, 2018). The source of drinking water, bottled, tap, filtered or well water, had an effect on the alpha diversity of the host's IM in a study. This study concluded that water may have a significant effect on our IM and should be considered in future IM studies (Vanhaecke *et al.*, 2022).

This study aimed to explore the potential correlation between bacterial taxonomic abundance and measured volatile fatty acid (VFA) levels, as well as examine the probable association between the consumption of unfiltered borehole drinking water and observed variability in the Ju|'hoansi IM. In order to do so VFA levels were measured per participant during 2021 sampling and samples of the unfiltered borehole drinking water taken from each village, sent for sequencing, and subsequently analysed.

Results

Correlation analysis

Correlation analysis was done on the VFAs per participant, the chemical analysis of the different villages' borehole drinking water, and the general metadata from village to dietary preferences in winter using Spearman correlations. There were quite a few significant correlations between villages and the chemical analysis of the drinking water sources and there was a significant correlation between decanoic acid and villages. Along with significant correlations with the different villages, the chemical analysis saw significant correlations with the occurrence of tuberculosis in participants, the use of antibiotics as well as their dietary preferences in winter.

Looking at the correlation analysis of the VFAs, Decanoic acid was the only one found to have significant correlations with the different villages ($p = 8.95 \times 10^{-6}$) and with the participants' dietary preferences in winter ($p = 0.02$). The other eight fatty acids had no correlation to any of the factors present in the metadata. Also taking a look at the β-diversity metrices, nothing of the VFAs held any significance. Many of the chemicals from the analysis of the water samples had a correlation with the villages, with calcium ($p = 0$) and potassium ($p = 0$) proving most significant, copper, nickel and lead were subsequently significant ($p = 3.17 \times 10^{-8}$), along with other factors (SI Table 7.1). Tuberculosis had some statistically significant correlations with copper, nickel, and lead ($p = 4 \times 10^{-5}$) (SI Table 7.2). Dietary preferences in winter correlated significantly with decanoic acid ($p = 0.02$), chloride, ammonium, and the turbidity of the water source ($p = 0.034$) along with others (SI Table 7.3).

Water microbiome analysis

The total number of reads from our water sampling was 233,675, which resulted in an average number of 33,382.14 reads per sample. After quality control with DADA2 (Callahan *et al.*, 2016), we were left with 290 ASVs. No core microbiome could be determined for the water samples collected from the villages. In our water sampling, Bacteroidota (36.34%) and Firmicutes (35%) were the dominant phyla. Following Bacteroidota and Firmicutes, the top three subsequent dominant taxa comprised Proteobacteria (12.17%), Fusobacterota (9.66%), and Patescibacteria (3.93%) (SI Table 8). A total of 63 genera were identified in the water samples.

Table 1. Bray-Curtis PERMANOVA results looking at differences in the Ju|'hoansi IM with regards to 1) Village, 2) pH, 3) Electrical conductivity, 4) Total dissolved solids, 5) Alkalinity, 6) Chloride, 7) Sulphate, 8) Nitrate, 9) Ammonium, 10) Fluoride, 11) Calcium, 12) Magnesium, 13) Sodium, 14) Potassium, 15) Manganese, 16) Copper, 17) Nickel, 18) Lead, 19) Total coliform, 20) Turbidity, 21) Total hardness, 22) Total organic carbon, 23) Langelier saturation index and 24) Total oxidised nitrogen.

Group	PERMANOVA R-squared	PERMANOVA F-statistic	PERMANOVA p value	Significance
Village	0.09417	1.1436	0.0473	*
pH	0.02997	1.0813	0.248	-
Electrical conductivity	0.02727	0.9811	0.5124	-
Total dissolved solids	0.02715	0.9768	0.5267	-
Alkalinity	0.02756	0.9919	0.477	-
Chloride	0.02669	0.9598	0.5809	-
Sulphate	0.02848	1.0261	0.3754	-
Nitrate	0.03428	1.2424	0.0448	*
Ammonium	0.02904	1.0469	0.3181	-
Fluoride	0.02899	1.0449	0.3187	-
Calcium	0.02857	1.0293	0.3736	-
Magnesium	1.1575	1.1575	0.1187	-
Sodium	0.02686	0.9662	0.5591	-
Potassium	0.02667	0.9591	0.5829	-
Manganese	0.03393	1.2292	0.06	-
Copper	0.02668	0.9595	0.5802	-
Nickel	0.02668	0.9595	0.5802	-
Lead	0.02668	0.9595	0.5802	-
Total coliform	0.03393	1.2292	0.06	-
Turbidity	0.03111	1.1238	0.1645	-
Total hardness	0.03196	1.1557	0.1202	-
Total organic carbon	0.02956	1.0661	0.2813	-
Langelier saturation index	0.03727	1.3551	0.0152	*
Total oxidised nitrogen	0.03428	1.2424	0.0448	*

When looking at the Bray-Curtis dissimilarity for 2021 (Table 1), we observed a few statistically significant factors including, participants' village of residency ($p = 0.047$), nitrate levels in the water ($p = 0.045$), the Langelier saturation index (LSI) ($p = 0.015$) and the total oxidised nitrogen concentration of the water source ($p = 0.045$). Although these factors were not highly significant there is literature to back up these factors. The statistical significance of the villages is a reason behind this study as it could be reliant on the villages' water sources. Differential abundance testing was performed regarding the 3 statistically significant water chemical composition factors, nitrate concentration, the LSI and the total amount of oxidised nitrogen in the water (SI Table 9). Three genera were found to be differentially abundant in the Ju|'hoansi IM when looking at nitrate concentration and the total oxidised nitrogen in the borehole drinking water; *Streptococcus*, *Selenomonas*, and [*Ruminococcus*]_torques_group. When looking at the LSI, 12 different bacterial taxa were differentially abundant. One common OTU belonging to the genus *Selenomonas* was differentially abundant amongst all 3 components. Notably it had differentially decreased with nitrate concentration and total oxidised nitrogen, but it was differentially enriched when looking at the Langelier saturation index.

Discussion

With the aim of exploring the correlation between Ju|'hoansi IM bacterial taxonomic abundance and measured VFA levels, faecal samples were collected, subsequently sequenced from a total cohort of 37 participants in 2021 and their VFA levels were analysed. An additional objective was to examine the association between the consumption of unfiltered borehole drinking water and observed variability in the Ju|'hoansi IM, thus borehole water samples were taken from each village of residence of the participants. Based on the results of this analysis, we detected one statistically significant VFA correlation between decanoic acid and a couple of factors. Furthermore, the water chemical composition had a few statistically significant correlations with factors relating to the Ju|'hoansi IM.

Decanoic acid, also known as capric acid, is a VFA that regulates fatty acid biosynthesis by bacteria (Sado-Kamdem *et al.*, 2009; Shuai *et al.*, 2022). This VFA has been suggested as a biomarker for colorectal cancer (Crotti *et al.*, 2016). According to research, the presence of decanoic acid reflects the intake of fibre in the oral microbiome (Hansen *et al.*, 2018), this could account for its elevated presence in those Ju|'hoansi participants whose dietary preference in winter consists of bush food. None of the other VFAs analysed from the faecal samples showed statistical correlations with the identified factors. Along with no significant correlations, there was also no significance within the β-diversity metrices for any of the VFAs.

Several studies have examined the effects of drinking water on our IM (Sofi *et al.*, 2014; Ayeni *et al.*, 2018; Bowyer *et al.*, 2020; Lugli *et al.*, 2022; Vanhaecke *et al.*, 2022). According to a study of UK tap water consumption, the composition of human IM may be a mediator between tap water and health (Bowyer *et al.*, 2020). Calcium and potassium concentrations in the water had strong correlations ($p = 0$) with the origin of the water sources, the different villages. The calcium concentrations ranged from 9.69 to 46.30 mg/L and the potassium concentrations ranged from 2.47 to 75.40 mg/L. These large ranges could account for the strong correlation between these chemicals and the villages. For copper, nickel and lead, 3 of the 4 villages had the same readings, whereas Den/ui had significantly higher readings of 0.057, 0.038 and 0.017 mg/L respectively.

The LSI ranged in value from 0.43 to 1.18 LSI and presented with a different value within that range at each villages' borehole. The LSI measures the water's capacity to dissolve calcium carbonate and is frequently used to determine the corrosivity of the water sources (Larson *et al.*, 1942; Langelier, 1946). The LSI of the Ju|'hoansi villages fall within normal ranges and does not indicate any corrosivity. Increasing nitrate levels in drinking water has been a recent cause for concern in many areas around the world, due to fertilizers and agricultural run-offs (Ward *et al.*, 2018). In the United States, the regulation for nitrate levels in public drinking water is capped at 10mg/L, however the Ju|'hoansi drinking water

sits within the ranges of 0.21 to 3.6 mg/L and should not be concerning (Ward *et al.*, 2018). *Streptococcus* was a common differentially enriched genus amongst changes in nitrate concentration and total oxidised nitrogen of the water sources. This lactic acid producing bacteria has been shown to improve the health of patients with chronic kidney disease (Vitetta *et al.*, 2019). Research regarding the genus *Selenomonas* role in the IM is lacking, but *Selenomonas* is a nitrate-reducing bacteria which is consistent with the nitrate concentration and the total oxidised nitrogen causing the differential abundance of *Selenomonas* (Asanuma *et al.*, 1999).

In our study, one sample was taken from each borehole drinking water source. This data was duplicated for each participant belonging to the relevant village the samples were taken from. For future investigations into drinking water effects on the IM, multiple sampling should be considered.

Conclusion

Faecal samples were collected and sequenced from a cohort of 37 Ju|'hoansi participants in 2021, and along with this, borehole water from the different villages of residence was collected, sequenced, and analysed for chemical composition so as to examine the probable association between the consumption of unfiltered borehole drinking water and observed variability in the Ju|'hoansi IM. In addition, the faecal samples were analysed for VFA levels in order to explore the potential correlation between bacterial taxonomic abundance and measured VFAs levels. Background information was collected about each participant from their biological data, i.e., age, and sex, to history, i.e., travel, medical history, and the village of residency, with the aim of finding correlations to VFA levels or the water composition. As per our analysis, we only detected one statistically significant VFA correlation between decanoic acid and some of the participants' background information. In comparison to the VFAs, multiple water chemical compositional factors had statistically significant correlations with factors in association with the Ju|'hoansi IM. This research is a continuation of an ongoing study elucidating the Ju|'hoansi IM.

Materials and Methods

Data collection

Faecal samples collected during July 2021 were sent in to assess the VFAs present. Refer to Chapter 2 on sampling methods used for the faecal samples. The VFAs analysed were acetic acid, formic acid, propanoic acid, isobutyric acid, butyric acid, isovaleric acid, hexanoic acid, octanoic acid, and decanoic acid. The water samples consisted of 2L of tap water taken from each village, i.e., Duinpos (H20_TDP), Den/ui (H20_TDE), Mountain Pos (H20_TMP) and !Om!o!o (H20_TOM), as well as 2L of tap water from our lab (H20_Lab) and 2L of MilliQ from the lab (H20_MQ). Tap water collected at the villages was first filtered using the Pall Manifold and the Easy-Load Masterflex peristaltic pump. Each sample was subsequently filtered using Isopore™ filters with a 0,2µm PC membrane. DNA was extracted and sent for sequencing. The borehole water samples collected were also sent to Aquatico Laboratories (Pty). Ltd for full composition analysis with regards to; pH, electrical conductivity, total dissolved solids, alkalinity, chemical composition, *E. coli*, total coliform, turbidity, total hardness, total organic carbon, cyanide screening, LSI, and total oxidised nitrogen (SI Table 1.2).

Bioinformatic pipeline for water samples

Raw sequence reads from the water samples were imported into QIIME 2 v.2022.2 (Bolyen *et al.*, 2019) and demultiplexed using the q2-demux plugin. Denoising of the paired-end sequences was done by DADA2 (Callahan *et al.*, 2016) using the q2-alignment plugin with a truncation length of f250/r250 and trim lengths of f26 and r24. A feature table was produced using the generated table output and a metadata text file. Alignment of the sequences was done using mafft (Katoh *et al.*, 2002) and a phylogenetic tree was created using fasttree2 (Price *et al.*, 2010) with the q2-phylogeny plugin. Taxonomy was determined with an untrained taxonomic classifier using q2-feature-classifier (Bokulich *et al.*, 2018) classify-sklearn with SILVA database (Quast *et al.*, 2012). After taxonomic classification, the relative abundances were calculated for the data and the phylogenetic trees were rerun using IQ-Tree (Nguyen *et al.*, 2015) for increased accuracy.

Analysis in R

The water data was imported into R to create a phyloseq object (Mcmurdie & Holmes, 2013) and rarefied to 1160 reads, which was the minimum number of reads detected in the samples. With rarefaction, only 3.85% of OTUs were removed from the dataset. These sequencing reads were discarded if they had only occurred once in the dataset. Using the R packages microbiome (Lahti & Shetty, 2017), Hmisc (Harrell Jr & Harrell Jr, 2019), stats (R Core Team, 2021) and vegan (Dixon, 2003), correlations analysis was done on the VFAs per participant, the chemical analysis of the different villages' borehole drinking water, and the general participants' metadata using Spearman correlations.

Chapter 4: Conclusions

Summary

This is to serve as the conclusive chapter of this dissertation, providing a summary of the important findings presented in the previous chapters. The intention is to highlight the most significant outcomes and conclusions derived from the results obtained during my research. The aim of this dissertation was to detect temporal changes in the Ju|'hoansi intestinal microbiome (IM), while also exploring the impact of water sources on their IM, and the levels of the IM volatile fatty acids (VFAs). To accomplish this, faecal samples were collected and subsequently sequenced from a total cohort of 40 participants taken 2 years apart in 2019 and 2021. The aim and objectives of this project were achieved through the use various bioinformatic techniques. Notable differences were found in the taxonomic composition of the Ju|'hoansi IM after the 2 year hiatus. Presented below is a summary of the key findings from each research chapter.

Temporal change in the Namibian Ju|'hoansi intestinal microbiome

In an effort to detect temporal changes in the Ju|'hoansi IM between 2019 and 2021, faecal samples were collected and sequenced from a cohort of 40 participants. To accomplish this background information was collected about each participant from their biological data, i.e., age, and sex, to history, i.e., travel, medical history, and the village of residency. Based on our research, there were observed changes in the species evenness of the Ju|'hoansi IM, yet species richness did not display see any changes during the recorded time period (Figure 4). The village where participants lived proved statistically significant when looked at considered in conjunction with the composition of their IM from year to year (SI Table 5.3). The water sources differ per village which could account for village having a significant impact on the Ju|'hoansi IM composition. Additionally, it is worth highlighting the decrease in number of species in the Ju|'hoansi core microbiome IM from baseline to follow-up, many bacterial genera were differentially decreased (SI Table 3.1, SI Table 3.2).

The effect of unfiltered borehole drinking water on the Namibian Ju|'hoansi intestinal microbiome

In addition to 37 faecal samples collected and sequenced in 2021, borehole water from the different villages of residence was collected, sequenced, and analysed for chemical composition to investigate the potential link between the consumption of unfiltered borehole drinking water and any observed variability in the Ju|'hoansi IM. Additionally, we explored the VFA levels present in the faecal samples collected in 2021 with the aim of exploring a possible correlation to IM bacterial taxonomic abundances. We looked at the background information collected about each participant with the intention of correlating the variables to VFA levels or the chemical composition of the borehole water. Following our analysis, we only detected one statistically significant VFA, decanoic acid. Decanoic acid had some

significant correlations between village of primary residence for participant and dietary preferences in winter (SI Tables 7.1 and 7.3). Decanoic acid can reflect the consumption of fibre in the oral microbiome, which is consistent with the dietary habits of the Ju|'hoansi participants. In contrast to the lack of significant correlations to the VFAs, the water chemical composition proved statistically significant with multiple factors such as chloride or potassium associated with the Ju|'hoansi IM (SI Tables 7.1 – 7.4).

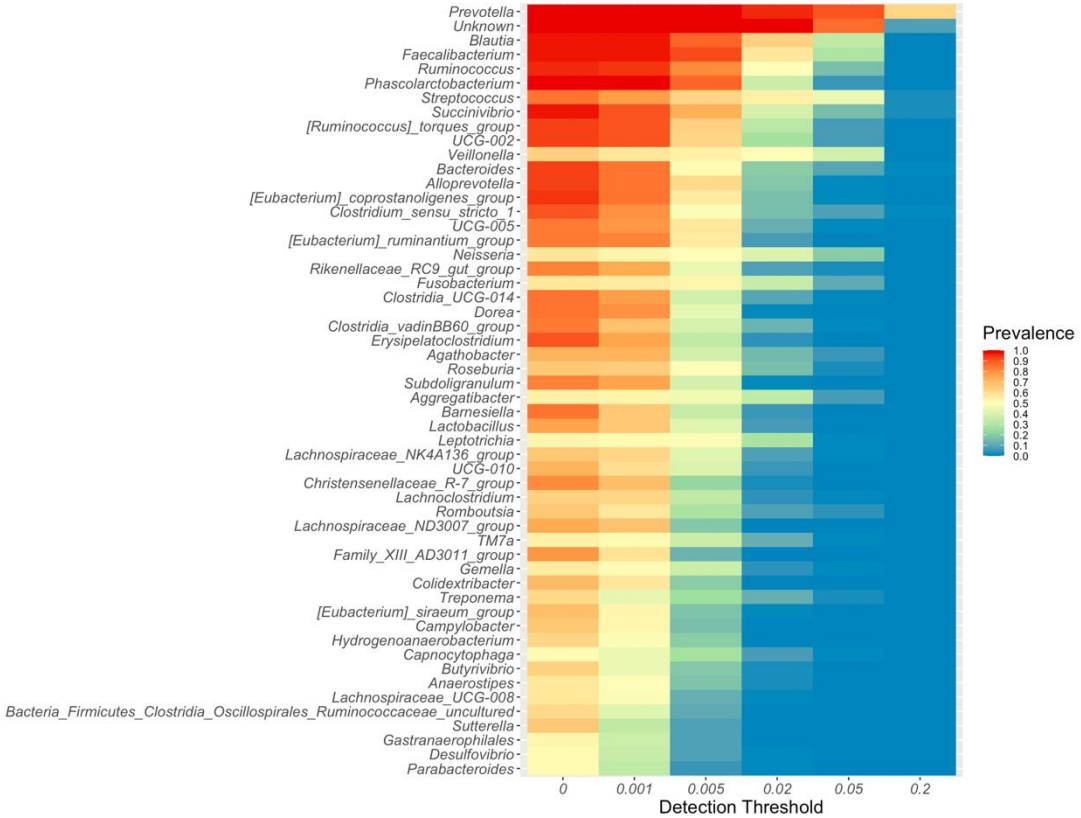
Future prospects

There are many future prospects for a study of this nature. Should we gain permits for further sampling, we could perform a longitudinal study of the Ju|'hoansi IM providing more data on an IM in transition from a traditional to a Western lifestyle. Should we still be interested in investigating VFAs linkage to the Ju|'hoansi IM, I would recommend doing a time series study with those VFAs important to the IM, along with having more in-depth questionnaires surrounding the dietary habits of the participants. This would provide some insight into any potential shift in VFAs that would accompany the contemporary hunter-gatherers transition to a more Western lifestyle. The impact of water composition, both chemically and taxonomically, on the composition of the human IM should be a topic of future studies, particularly in terms of traditional communities which are becoming increasingly sedentary and are depending more on water derived from boreholes. This data was duplicated for each participant belonging to the relevant village the samples were taken from. For future investigations into drinking water effects on the IM, multiple sampling should be considered.

Conclusion

The aim of this research to detect temporal changes in the Ju|'hoansi IM over a 2-year period, was successfully achieved through the use of bioinformatic analyses. This research is a continuation of an ongoing study elucidating the Ju|'hoansi IM. Studying contemporary hunter-gatherers prove useful in understanding how a traditional IM changes with Western influences over time. In summary, statistically significant changes in the taxonomic composition of the Ju|'hoansi IM were observed after the 2-year period. Further investigation is warranted in order to determine how fast the Ju|'hoansi IM may be transitioning towards a more Western IM.

Supplementary Information



SI Figure 1: Core Ju|'hoansi IM at a genus level, elucidated at 50% prevalence, and at a detection threshold of 0.001%.

SI Table 1.1: Metadata collected for this study.

Sample ID	Sampling date		Village		Gender	Age		Travel		Dietary Preferences in Winter		Intestinal Infections		Imodium, traditional, both or no medication for intestinal infection		Malaria		Malaria Medication		TB		Antibiotics		Lifestyle changes (2yrs)
	2019	2021	2019	2021		2019	2021	2019	2021	2019	2021	2019	2021	2019	2021	2019	2021	2019	2021	2019	2021	2019	2021	
TDP01	22/07/2019	25/07/2021	Duinpos	Duinpos	Male	35	37	Local	Regional	Shop food	Mix	No	No	None	None	No	No	No	No	Yes	No	Yes	No	No
TDP02	22/07/2019	25/07/2021	Duinpos	Duinpos	Female	49	51	Regional	International	Shop food	Bush food	No	No	None	None	No	No	No	No	Yes	No	Yes	No	No
TDP03	22/07/2019	25/07/2021	Duinpos	Duinpos	Male	69	71	Local	x	Shop food	Bush food	Yes	No	None	None	No	No	No	No	No	No	No	No	No
TDP04	22/07/2019	25/07/2021	Duinpos	Duinpos	Female	46	53	Local	x	Shop food	Bush food	No	Yes	Not sure	Traditional	No	No	Not sure	No	No	No	Not sure	No	No
TDP05	22/07/2019	25/07/2021	Duinpos	Duinpos	Male	35	37	Local	Local	Shop food	Bush food	No	No	None	None	No	No	No	No	Yes	No	Yes	No	No
TDP06	22/07/2019	25/07/2021	Duinpos	Duinpos	Female	51	53	Local	Regional	Shop food	Bush food	No	Not sure	Not sure	None	No	No	No	No	No	No	Not sure	No	Yes
TDP07	22/07/2019	25/07/2021	Duinpos	Duinpos	Male	25	27	Local	Regional	Shop food	Bush food	No	Yes	None	None	No	No	No	No	Yes	No	Yes	No	No
TDP08	22/07/2019	25/07/2021	Duinpos	Duinpos	Female	49	51	Local	Local	Shop food	Bush food	No	No	None	None	No	No	No	No	Yes	No	Yes	No	No
TDP09	22/07/2019	25/07/2021	Duinpos	Duinpos	Male	33	35	Local	Local	Shop food	Mix	No	No	None	None	No	No	No	No	Yes	No	Yes	No	No
TDP10	22/07/2019	25/07/2021	Duinpos	Duinpos	Female	41	43	Local	Regional	Shop food	Bush food	No	No	None	None	No	Yes	No	?	Yes	No	Yes	No	Yes
TDE01	23/07/2019	26/07/2021	Den/ui	Den/ui	Male	32	33	Local	Local	Shop food	Shop food	No	No	None	None	No	No	No	No	Yes	Yes	Yes	No	No
TDE02	23/07/2019	26/07/2021	Den/ui	Den/ui	Female	27	29	Local	International	Shop food	Bush food	Yes	No	Imodium	None	Yes	No	Yes	No	Yes	No	Yes	No	No
TDE03	23/07/2019	26/07/2021	Den/ui	Den/ui	Male	43	43	Local	Local	Shop food	Bush food	Not sure	No	Imodium	None	No	Yes	No	Yes	Yes	Yes	Yes	No	No
TDE04	23/07/2019	26/07/2021	Den/ui	Den/ui	Female	53	55	Local	Local	Bush food	Bush food	No	No	None	None	No	No	No	No	Yes	Yes	Yes	No	No
TDE05	23/07/2019	26/07/2021	Den/ui	Den/ui	Male	50	52	International	International	Bush food	Bush food	No	No	None	None	No	No	No	No	Yes	Yes	Yes	Not sure	Yes
TDE06	23/07/2019	26/07/2021	Den/ui	Den/ui	Female	22	24	Local	Local	Shop food	Bush food	Yes	Not sure	Imodium	Not sure	No	No	No	No	Yes	No	Yes	Not sure	No
TDE07	23/07/2019	26/07/2021	Den/ui	Den/ui	Male	44	46	Regional	Regional	Bush food	Bush food	No	No	None	None	No	No	No	No	Yes	Yes	Yes	No	No
TDE08	23/07/2019	26/07/2021	Den/ui	Den/ui	Male	58	60	Local	x	Shop food	Bush food	Not sure	Yes	None	Not sure	No	No	No	No	Yes	Yes	Yes	Not sure	No
TDE09	23/07/2019	26/07/2021	Den/ui	Den/ui	Female	19	21	Local	Local	Shop food	Bush food	Yes	No	Imodium	None	No	No	No	No	Yes	Yes	Yes	No	Yes
TDE10	23/07/2019	26/07/2021	Den/ui	Den/ui	Female	38	40	Local	Local	Shop food	Mix	Yes	No	Imodium	None	No	No	No	No	Yes	Yes	Yes	No	No
TMP01	24/07/2019	27/07/2021	Mountain Pos	Mountain Pos	Male	31	32	Local	Regional	Shop food	Bush food	No	No	Both	None	No	No	No	No	No	No	No	No	No
TMP02	24/07/2019	27/07/2021	Mountain Pos	Mountain Pos	Female	24	26	Local	Local	Shop food	Bush food	No	No	None	None	No	No	No	No	No	No	No	No	No
TMP03	24/07/2019	27/07/2021	Mountain Pos	Mountain Pos	Male	63	65	Regional	Regional	Shop food	Mix	No	Yes	None	None	No	No	No	No	Yes	Yes	Yes	No	No
TMP04	24/07/2019	27/07/2021	Mountain Pos	Mountain Pos	Female	60	62	Local	International	Shop food	Mix	No	No	None	None	No	No	No	No	Yes	Yes	Yes	No	Yes
TMP05	24/07/2019	27/07/2021	Mountain Pos	Mountain Pos	Male	35	37	Local	x	Shop food	Bush food	No	No	None	None	No	No	No	No	Yes	No	Yes	No	No
TMP06	24/07/2019	27/07/2021	Mountain Pos	Mountain Pos	Female	52	55	Local	Local	Shop food	Mix	No	No	None	None	No	No	No	No	No	No	No	No	No
TMP07	24/07/2019	27/07/2021	Mountain Pos	Mountain Pos	Male	36	37	Local	Local	Shop food	Mix	No	No	None	None	No	No	No	No	Yes	No	Yes	Yes	No
TMP08	24/07/2019	27/07/2021	Mountain Pos	Mountain Pos	Female	21	23	Local	Local	Shop food	Mix	Yes	No	Imodium	None	No	No	No	No	Yes	No	Yes	No	No
TMP09	24/07/2019	27/07/2021	Mountain Pos	Mountain Pos	Male	38	39	Local	Local	Shop food	Mix	No	No	None	None	No	No	No	No	No	No	No	No	No
TMP10	24/07/2019	27/07/2021	Mountain Pos	Mountain Pos	Female	29	31	Local	Local	Shop food	Mix	No	No	None	None	No	No	No	No	No	No	No	No	No
TOM01	25/07/2019	28/07/2021	!Om!o!o	!Om!o!o	Female	45	47	Local	Local	Bush food	Mix	Not sure	Not sure	Traditional	None	Yes	No	Yes	No	No	No	No	Not sure	Yes
TOM02	25/07/2019	28/07/2021	!Om!o!o	!Om!o!o	Male	25	27	Local	Regional	Bush food	Shop food	No	No	None	Not sure	Yes	Yes	Yes	No	No	No	No	Not sure	No
TOM03	25/07/2019		!Om!o!o		Female	57		International		Bush food		Yes		None		Yes		Yes		No		No		
TOM04	25/07/2019	28/07/2021	!Om!o!o	!Om!o!o	Male	31	33	Local	Regional	Bush food	Shop food	Yes	No	None	None	No	No	No	No	Yes	Yes	Yes	No	No
TOM05	25/07/2019	28/07/2021	!Om!o!o	!Om!o!o	Female	40	42	Local	x	Bush food	Mix	Yes	No	Imodium	None	No	No	No	No	No	No	No	No	No
TOM06	25/07/2019		!Om!o!o		Male	35		Local		Shop food		No		None		No		No		No		No		
TOM07	25/07/2019		!Om!o!o		Female	50		Local		Bush food		Yes		Imodium		No		No		No		No		
TOM08	25/07/2019	28/07/2021	!Om!o!o	!Om!o!o	Male	24	26	Regional	Local	Shop food	Shop food	Yes	Yes	None	Not sure	No	Yes	No	Yes	No	No	No	Not sure	Yes
TOM09	25/07/2019	28/07/2021	!Om!o!o	!Om!o!o	Female	28	30	Local	Regional	Shop food	Mix	Yes	Yes	Imodium	None	No	No	No	No	No	No	No	Yes	No
TOM10	25/07/2019	28/07/2021	!Om!o!o	!Om!o!o	Male	43	45	Local	Regional	Shop food	Bush food	Yes	Yes	Both	None	No	No	No	No	No	No	No	Not sure	No
CON-CTRL	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
KIT-CTRL	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

SI Table 1.2: Borehole drinking water metadata collected for Chapter 3 study.

Sample ID	Village	pH	Electrical Conductivity	Total Dissolved Solids	Alkalinity
H2O_TDE	Den/ui	9.09	191	1273	720
H2O_TDP	Duinpos	8.27	32.2	177	144
H2O_TMP	Mountain Pos	8.6	37	233	208
H2O_TOM	≠!Om!o!o	8.17	83.3	473	359
H2O_MQ	NA	0	0	0	0
H2O_Con	NA	0	0	0	0
H2O_Lab	NA	0	0	0	0

SI Table 1.2: Borehole drinking water metadata collected for Chapter 3 study (cont.).

Sample ID	Chloride	Sulphate	Nitrate	Ammonium	Flouride	Calcium	Magnesium
H2O_TDE	259	65.9	2.34	0.024	3.06	9.69	26.2
H2O_TDP	11.3	5.51	1.15	0.036	-0.263	23.7	9.94
H2O_TMP	6.76	8.27	0.209	0.045	0.516	27.9	15.4
H2O_TOM	36	41.7	3.6	0.025	2.05	46.3	38
H2O_MQ	0	0	0	0	0	0	0
H2O_Con	0	0	0	0	0	0	0
H2O_Lab	0	0	0	0	0	0	0

SI Table 1.2: Borehole drinking water metadata collected for Chapter 3 study (cont.).

Sample ID	Sodium	Potassium	Aluminium	Iron	Manganese	Chromium	Copper
H2O_TDE	384	75.4	-0.002	-0.004	-0.001	-0	0.057
H2O_TDP	26.2	7.65	-0.002	-0.004	-0.001	-0	-0.002
H2O_TMP	39.4	6.99	-0.002	-0.004	0.119	-0	-0.002
H2O_TOM	71.5	2.47	-0.002	-0.004	-0.001	-0	-0.002
H2O_MQ	0	0	0	0	0	0	0
H2O_Con	0	0	0	0	0	0	0
H2O_Lab	0	0	0	0	0	0	0

SI Table 1.2: Borehole drinking water metadata collected for Chapter 3 study (cont.).

Sample ID	Nickel	Zinc	Cadium	Lead	E. Coli	Total Coliform	Turbidity	Total Hardness
H2O_TDE	0.038	-0.002	-0.002	0.02	-1	-1	0.194	132
H2O_TDP	-0.002	-0.002	-0.002	-0	-1	-1	4.95	100
H2O_TMP	-0.002	-0.002	-0.002	-0	-1	7	15.7	133
H2O_TOM	-0.002	-0.002	-0.002	-0	-1	-1	3.78	272
H2O_MQ	0	0	0	0	0	0	0	0
H2O_Con	0	0	0	0	0	0	0	0
H2O_Lab	0	0	0	0	0	0	0	0

SI Table 1.2: Borehole drinking water metadata collected for Chapter 3 study (cont.).

Sample ID	Total Organic Carbon	Cyanide Screening	Langelier Saturation Index (LSI)	Total Oxidised Nitrogen
H2O_TDE	34.5	-1.1	0.93	2.34
H2O_TDP	1.27	-1.1	0.61	1.15
H2O_TMP	14.4	-1.1	1.18	0.209
H2O_TOM	1.57	-1.1	0.43	3.6
H2O_MQ	0	0	0	0
H2O_Con	0	0	0	0
H2O_Lab	0	0	0	0

SI Table 2.1.1: The Ju|'hoansi IM consisted of 14 identified bacterial phyla in 2019.

Phylum	**Relative Abundance**
Firmicutes	24.49351
Bacteroidota	10.67508
Proteobacteria	3.66831
Spirochaetota	0.37722
Actinobacteriota	0.32748
Desulfobacterota	0.11903
Verrucomicrobiota	0.08722
Cyanobacteria	0.03876
Fusobacteriota	0.03511
Campilobacterota	0.01828
Synergistota	0.01518
Patescibacteria	0.00914
Elusimicrobiota	0.00165
Fibrobacterota	0.00055

SI Table 2.1.2: The Ju|'hoansi IM consisted of 14 identified bacterial phyla in 2021.

Phylum	**Relative Abundance**
Firmicutes	15.05668
Bacteroidota	13.38599
Proteobacteria	4.53355
Fusobacteriota	2.07552
Patescibacteria	0.99726
Actinobacteriota	0.45767
Campilobacterota	0.15085025
Spirochaetota	0.12836
Cyanobacteria	0.08630
Verrucomicrobiota	0.02249
Desulfobacterota	0.01134
Synergistota	0.00457
Planctomycetota	0.00274
Acidobacteriota	0.00110

SI Table 2.2.1: The Ju|'hoansi IM consisted of 23 bacterial classes in 2019.

Class	Relative Abundance
Clostridia	21.07332
Bacteroidia	10.67508
Gammaproteobacteria	3.42311
Bacilli	1.9744
Negativicutes	1.39367
Spirochaetia	0.37703
Alphaproteobacteria	0.24575
Actinobacteria	0.19675
Coriobacteriia	0.13074
Desulfovibrionia	0.11903
Lentisphaeria	0.05979
Vampirivibrionia	0.03876
Desulfotomaculia	0.03748
Fusobacteriia	0.03511
Verrucomicrobiae	0.02743
Campylobacteria	0.01828
Synergistia	0.01518
Saccharimonadia	0.00914
Endomicrobia	0.00165
Fibrobacteria	0.00055
Brachyspirae	0.00018
Limnochordia	0.00018

SI Table 2.2.2: The Ju|'hoansi IM consisted of 22 bacterial classes in 2021.

Class	Relative Abundance
Bacteroidia	13.38599
Clostridia	7.10660
Gammaproteobacteria	4.50009
Bacilli	4.18998
Negativicutes	3.56409
Fusobacteriia	2.07552
Saccharimonadia	0.94661
Actinobacteria	0.37411
Campylobacteria	0.15085
Spirochaetia	0.12287
Coriobacteriia	0.08356
Vampirivibrionia	0.08045
Gracilibacteria	0.04900
Alphaproteobacteria	0.03346
Verrucomicrobiae	0.01810
Desulfovibrionia	0.01134
Desulfotomaculia	0.00622
Brachyspirae	0.00549
Synergistia	0.00457
Lentisphaeria	0.00439
Planctomycetes	0.00274
Acidobacteriae	0.00110

SI Table 2.3.1: 49 bacterial orders were identified in the Ju|'hoansi IM in 2019.

Order	**Relative Abundance**	**Order**	**Relative Abundance**
Bacteroidales	10.67069	Gastranaerophilales	0.03876
Lachnospirales	8.98354	Desulfotomaculales	0.03748
Oscillospirales	7.94313	Fusobacteriales	0.03511
B55-F-B-G02	3.26678	Staphylococcales	0.02962
Aeromonadales	2.22765	Verrucomicrobiales	0.02743
Lactobacillales	1.34522	Campylobacterales	0.01828
Enterobacterales	1.09545	Acholeplasmatales	0.0181
Acidaminococcales	0.97824	Rhodospirillales	0.0181
Clostridiales	0.93436	Synergistales	0.01518
Peptostreptococcales-Tissierellales	0.75864	Saccharimonadales	0.00914
Clostridia_vadinBB60_group	0.46645	Pasteurellales	0.00567
Christensenellales	0.44304	Bifidobacteriales	0.00329
Erysipelotrichales	0.4328	Flavobacteriales	0.00311
Veillonellales-Selenomonadales	0.41543	Eubacteriales	0.00183
Spirochaetales	0.37703	Endomicrobiales	0.00165
Clostridia_UCG-014	0.29055	Lentisphaerales	0.00146
Rhizobiales	0.21942	Corynebacteriales	0.00128
Coriobacteriales	0.13074	Halothiobacillales	0.000731
Desulfovibrionales	0.11903	Dongiales	0.00073
Micrococcales	0.11593	Streptomycetales	0.00073
Burkholderiales	0.09179	Fibrobacterales	0.00055
Oligosphaerales	0.05833	Sphingomonadales	0.00037
Actinomycetales	0.05083	Brachyspirales	0.00018
Clostridia	0.04827	Sphingobacteriales	0.00018
Monoglobales	0.04462		

SI Table 2.3.2: 56 bacterial orders were identified in the Ju|'hoansi IM in 2021.

Order	Relative Abundance	Order	Relative Abundance
Bacteroidales	12.83598	Actinomycetales	0.03328
Lactobacillales	3.60413	Verrucomicrobiales	0.01700
Lachnospirales	3.15304	Rhodospirillales	0.01682
Veillonellales-Selenomonadales	3.08466	Desulfovibrionales	0.01134
Oscillospirales	2.41488	Corynebacteriales	0.01097
Fusobacteriales	2.07552	Monoglobales	0.00695
Burkholderiales	2.04132	JGI_0000069-P22	0.00677
Pasteurellales	1.06363	Clostridia	0.00622
Saccharimonadales	0.94661	Desulfotomaculales	0.00622
Aeromonadales	0.82319	Brachyspirales	0.00549
Flavobacteriales	0.54014	Mycoplasmatales	0.00457
Acidaminococcales	0.47943	Synergistales	0.00457
Enterobacterales	0.41598	Acholeplasmatales	0.00457
Peptostreptococcales-Tissierellales	0.39367	Oligosphaerales	0.00439
Staphylococcales	0.37649	Sphingobacteriales	0.00421
Micrococcales	0.26440	Bifidobacteriales	0.00219
Clostridiales	0.24520	Rhizobiales	0.00201
Clostridia_UCG-014	0.20333	Eubacteriales	0.00183
Campylobacterales	0.15085	Chitinophagales	0.00183
Clostridia_vadinBB60_group	0.17133	Sphingomonadales	0.00165
Erysipelotrichales	0.14445	Planctomycetales	0.00165
Spirochaetales	0.12287	Pedosphaerales	0.00110
Xanthomonadales	0.09051	Pirellulales	0.00110
Coriobacteriales	0.08356	Thiomicrospirales	0.00073
Gastranaerophilales	0.08045	Bryobacterales	0.00073
Christensenellales	0.07661	Dongiales	0.00018
Pseudomonadales	0.05796	Acidobacteriales	0.00017
Gracilibacteria	0.04224		

SI Table 2.4.1: The Ju|'hoansi IM consisted of 81 bacterial families in 2019.

Family	**Family**
[Clostridium]_methylpentosum_group	*Halothiobacillaceae*
[Eubacterium]_coprostanoligenes_group	*Helicobacteraceae*
Acholeplasmataceae	*Hungateiclostridiaceae*
Acidaminococcaceae	*JGI_0000069-P22*
Actinomycetaceae	*Lachnospiraceae*
Aerococcaceae	*Lactobacillaceae*
Anaerovoracaceae	*Lenti-02*
Atopobiaceae	*Lentisphaeraceae*
B55-F-B-G02	*Leptotrichiaceae*
Bacteroidaceae	*Marinifilaceae*
Barnesiellaceae	*Micrococcaceae*
Bifidobacteriaceae	*Monoglobaceae*
Brachyspiraceae	*Neisseriaceae*
Butyricicoccaceae	*Oscillospiraceae*
Caloramatoraceae	*Oscillospirales*
Campylobacteraceae	*Pasteurellaceae*
Christensenellaceae	*Peptostreptococcaceae*
Clostridia_UCG-014	*Peptostreptococcales-Tissierellales*
Clostridia_vadinBB60_group	*Porphyromonadaceae*
Clostridiaceae	*Prevotellaceae*
Coriobacteriaceae	*Rhizobiaceae*
Coriobacteriales_Incertae_Sedis	*Rikenellaceae*
Corynebacteriaceae	*Ruminococcaceae*
Defluviitaleaceae	*Saccharimonadaceae*
Desulfotomaculales	*Saccharimonadales*
Desulfovibrionaceae	*Selenomonadaceae*
Dongiaceae	*Sphingobacteriaceae*
Dysgonomonadaceae	*Sphingomonadaceae*
Eggerthellaceae	*Spirochaetaceae*
Endomicrobiaceae	*Streptococcaceae*
Enterobacteriaceae	*Streptomycetaceae*
Enterococcaceae	*Succinivibrionaceae*
Erysipelatoclostridiaceae	*Sutterellaceae*
Erysipelotrichaceae	*Synergistaceae*
Eubacteriaceae	*Tannerellaceae*
Fibrobacteraceae	*Terasakiellaceae*
Flavobacteriaceae	*UCG-010*
Fusobacteriaceae	*Veillonellaceae*
Gastranaerophilales	*Verrucomicrobiaceae*
Gemellaceae	*Xanthobacteraceae*
Gracilibacteraceae	

SI Table 2.4.2: The Ju|'hoansi IM consisted of 88 bacterial families in 2021.

Family	**Family**
[Eubacterium]_coprostanoligenes_group	*Lactobacillaceae*
Acholeplasmataceae	*Lenti-02*
Acidaminococcaceae	*Lentimicrobiaceae*
Acidobacteriaceae_(Subgroup_1)	*Leptotrichiaceae*
Actinomycetaceae	*Marinifilaceae*
Aerococcaceae	*Micrococcaceae*
Alcaligenaceae	*Monoglobaceae*
Anaerovoracaceae	*Moraxellaceae*
Atopobiaceae	*Morganellaceae*
Bacteroidaceae	*Mycoplasmataceae*
Barnesiellaceae	*Neisseriaceae*
Bifidobacteriaceae	*Nocardiaceae*
Brachyspiraceae	*Oscillospiraceae*
Bryobacteraceae	*Oscillospirales*
Butyricicoccaceae	*Pasteurellaceae*
Campylobacteraceae	*Pedosphaeraceae*
Chitinophagaceae	*Peptostreptococcaceae*
Christensenellaceae	*Peptostreptococcales-Tissierellales*
Clostridia_UCG-014	*Pirellulaceae*
Clostridia_vadinBB60_group	*Porphyromonadaceae*
Clostridiaceae	*Prevotellaceae*
Comamonadaceae	*Pseudomonadaceae*
Coriobacteriales_Incertae_Sedis	*Rhizobiaceae*
Corynebacteriaceae	*Rikenellaceae*
Defluviitaleaceae	*Ruminococcaceae*
Desulfotomaculales	*Saccharimonadaceae*
Desulfovibrionaceae	*Saccharimonadales*
Dongiaceae	*Selenomonadaceae*
Dysgonomonadaceae	*Sphingomonadaceae*
Enterobacteriaceae	*Spirochaetaceae*
Enterococcaceae	*Streptococcaceae*
Erysipelatoclostridiaceae	*Succinivibrionaceae*
Erysipelotrichaceae	*Sutterellaceae*
Eubacteriaceae	*Synergistaceae*
Flavobacteriaceae	*Tannerellaceae*
Fusobacteriaceae	*Terasakiellaceae*
Gastranaerophilales	*Thioglobaceae*
Gemellaceae	*UCG-010*
Gimesiaceae	*Veillonellaceae*
Gracilibacteraceae	*Verrucomicrobiaceae*
Gracilibacteria	*Weeksellaceae*
Helicobacteraceae	*Xanthobacteraceae*
Hungateiclostridiaceae	*Xanthomonadaceae*
JGI_0000069-P22	*Lachnospiraceae*

SI Table 2.5.1: The Ju|'hoansi IM consisted of 141 bacterial genera in 2019.

Genus	**Genus**	**Genus**
[Clostridium]_methylpentosum_group	*Corynebacterium*	*Leptotrichia*
[Eubacterium]_brachy_group	*Defluviitaleaceae_UCG-011*	*Macellibacteroides*
[Eubacterium]_coprostanoligenes_group	*Desulfovibrio*	*Marvinbryantia*
[Eubacterium]_ruminantium_group	*Dialister*	*Mesorhizobium*
[Eubacterium]_siraeum_group	*DNF00809*	*Monoglobus*
[Eubacterium]_ventriosum_group	*Dongia*	*Mucilaginibacter*
[Eubacterium]_xylanophilum_group	*Dorea*	*Murdochiella*
[Ruminococcus]_gauvreauii_group	*Dysgonomonas*	*Neisseria*
[Ruminococcus]_gnavus_group	*Eggerthella*	*NK4A214_group*
[Ruminococcus]_torques_group	*Endomicrobium*	*Novosphingobium*
Abiotrophia	*Enterococcus*	*Odoribacter*
Acetoanaerobium	*Erysipelatoclostridium*	*Oribacterium*
Actinomyces	*Escherichia-Shigella*	*Oscillibacter*
Agathobacter	*Eubacterium*	*Paeniclostridium*
Aggregatibacter	*Ezakiella*	*Parabacteroides*
Alistipes	*Faecalibacterium*	*Parvimonas*
Alloprevotella	*Family_XIII_AD3011_group*	*Pelotomaculum*
Amnipila	*Fenollaria*	*Peptoniphilus*
Anaerococcus	*Fibrobacter*	*Peptostreptococcus*
Anaerostipes	*Finegoldia*	*Phascolarctobacterium*
Anaerotruncus	*Fretibacterium*	*Porphyromonas*
Anaerovorax	*Frisingicoccus*	*Prevotella*
Atopobium	*Fusobacterium*	*Pseudobutyrivibrio*
B55-F-B-G02	*Gallicola*	*Raoultibacter*
Bacteroides	*Gastranaerophilales*	*Rikenellaceae_RC9_gut_group*
Barnesiella	*Gemella*	*Romboutsia*
Bifidobacterium	*Gracilibacter*	*Roseburia*
Blautia	*Helicobacter*	*Ruminiclostridium*
Brachyspira	*Hungateiclostridium*	*Ruminococcus*
Bradyrhizobium	*Hungatella*	*Saccharimonadales*
Butyricicoccus	*Hydrogenoanaerobacterium*	*Selenomonas*
Butyrivibrio	*Incertae_Sedis*	*Senegalimassilia*
CAG-56	*Intestinibacter*	*Slackia*
Campylobacter	*JGI_0000069-P22*	*Sphaerochaeta*
Candidatus_Phytoplasma	*Klebsiella*	*Stomatobaculum*
Capnocytophaga	*Lachnoanaerobaculum*	*Streptococcus*
Catonella	*Lachnoclostridium*	*Subdoligranulum*
Christensenella	*Lachnospira*	*Succinivibrio*
Christensenellaceae_R-7_group	*Lachnospiraceae_FCS020_group*	*Sutterella*
Clostridia_UCG-014	*Lachnospiraceae_ND3007_group*	*Thiovirga*
Clostridia_vadinBB60_group	*Lachnospiraceae_NK4A136_group*	*TM7a*
Clostridium_sensu_stricto_1	*Lachnospiraceae_UCG-004*	*Treponema*
Clostridium_sensu_stricto_6	*Lachnospiraceae_UCG-008*	*Turicibacter*
Clostridium_sensu_stricto_7	*Lachnospiraceae_UCG-009*	*UCG-002*
Colidextribacter	*Lactobacillus*	*UCG-005*
Collinsella	*Lenti-02*	*UCG-010*
Coprobacter	*Lentisphaera*	*Veillonella*

SI Table 2.5.2: The Ju|'hoansi IM consisted of 142 bacterial genera in 2021.

Genus	**Genus**	**Genus**
[Eubacterium]_brachy_group	*Defluviitaleaceae_UCG-011*	*Mesorhizobium*
[Eubacterium]_coprostanoligenes_group	*Desulfovibrio*	*Monoglobus*
[Eubacterium]_ruminantium_group	*Dialister*	*Mycoplasma*
[Eubacterium]_siraeum_group	*Dongia*	*Neisseria*
[Eubacterium]_ventriosum_group	*Dorea*	*NK4A214_group*
[Eubacterium]_yurii_group	*Dysgonomonas*	*Odoribacter*
[Ruminococcus]_gauvreauii_group	*Enhydrobacter*	*Oribacterium*
[Ruminococcus]_gnavus_group	*Enterococcus*	*Oscillibacter*
[Ruminococcus]_torques_group	*Erysipelatoclostridium*	*Paenarthrobacter*
Abiotrophia	*Escherichia-Shigella*	*Parabacteroides*
Acetitomaculum	*Eubacterium*	*Parvimonas*
Acinetobacter	*Faecalibacterium*	*Pelotomaculum*
Actinomyces	*Family_XIII_AD3011_group*	*Peptoniphilus*
Agathobacter	*Fenollaria*	*Peptostreptococcaceae*
Aggregatibacter	*Finegoldia*	*Peptostreptococcus*
Alistipes	*Flavisolibacter*	*Phascolarctobacterium*
Alloprevotella	*Fretibacterium*	*Polaromonas*
Amnipila	*Frisingicoccus*	*Porphyromonas*
Anaerosporobacter	*Fusobacterium*	*Prevotella*
Anaerostipes	*Gastranaerophilales*	*Proteus*
Anaerovorax	*Gemella*	*Pseudomonas*
Atopobium	*Gracilibacter*	*Raoultibacter*
Bacteroides	*Gracilibacteria*	*Rhodococcus*
Barnesiella	*Helicobacter*	*Rikenellaceae_RC9_gut_group*
Bifidobacterium	*Hungateiclostridium*	*Romboutsia*
Blastopirellula	*Hungatella*	*Roseburia*
Blautia	*Hydrogenoanaerobacterium*	*Ruminiclostridium*
Bordetella	*Incertae_Sedis*	*Ruminococcus*
Brachyspira	*Intestinibacter*	*Saccharimonadales*
Bradyrhizobium	*JGI_0000069-P22*	*Selenomonas*
Bryobacter	*Kingella*	*Sphaerochaeta*
Butyricicoccus	*Klebsiella*	*Sphingomonas*
Butyrivibrio	*Lachnoanaerobaculum*	*Stenotrophomonas*
CAG-56	*Lachnoclostridium*	*Stomatobaculum*
Campylobacter	*Lachnospira*	*Streptococcus*
Candidatus_Phytoplasma	*Lachnospiraceae_FCS020_group*	*Subdoligranulum*
Candidatus_Saccharimonas	*Lachnospiraceae_ND3007_group*	*Succinivibrio*
Capnocytophaga	*Lachnospiraceae_NK4A136_group*	*SUP05_cluster*
Catonella	*Lachnospiraceae_UCG-004*	*Sutterella*
Christensenella	*Lachnospiraceae_UCG-008*	*Tannerella*
Christensenellaceae_R-7_group	*Lachnospiraceae_UCG-009*	*TM7a*
Cloacibacterium	*Lactobacillus*	*Treponema*
Clostridia_UCG-014	*Lenti-02*	*Turicibacter*
Clostridia_vadinBB60_group	*Lentimicrobium*	*UCG-002*
Clostridium_sensu_stricto_1	*Leptotrichia*	*UCG-005*
Colidextribacter	*Macellibacteroides*	*UCG-010*
Corynebacterium	*Marvinbryantia*	*Veillonella*

SI Table 2.6.1: 44 bacterial species were identified in the Ju|'hoansi IM in 2019.

Species	**Relative Abundance**	**Species**	**Relative Abundance**
Bacteroides_sp.	0.31999	*Leptotrichia_sp.*	0.00567
Bacteroidales_genomosp.	0.20845	uncultured_*Firmicutes*	0.00559
Bradyrhizobium_japonicum	0.17316	uncultured_*Ruminococcaceae*	0.00549
Christensenella_minuta	0.17115	uncultured_*Spirochaetes*	0.0053
Anaerostipes_hadrus	0.13494	*Streptococcus_sp.*	0.00494
Ruminococcus_flavefaciens	0.13421	uncultured_*Lactobacillus*	0.00439
Parabacteroides_gordonii	0.10825	uncultured_*Campylobacteraceae*	0.00421
uncultured_*Bacteroidetes*	0.05065	*Anaerococcus_prevotii*	0.00384
uncultured_*cyanobacterium*	0.03876	*Ruminococcus_albus*	0.0031
Prevotella_loescheii	0.02889	uncultured_delta	0.00219
Mycobacterium_tuberculosis	0.02779	*Bacteroidales_bacterium*	0.00201
Lactobacillus_salivarius	0.02743	uncultured_*Ruminococcus*	0.00146
Desulfovibrio_vulgaris	0.0267	*Eubacterium_sp.*	0.00091
Gemella_sp.	0.02505	uncultured_*Candidatus*	0.00091
uncultured_*Christensenella*	0.01755	uncultured_*Blautia*	0.00055
Bacillus_firmus	0.017	uncultured_*Bradyrhizobium*	0.00055
Clostridium_bornimense	0.01426	uncultured_*Clostridium*	0.00055
Campylobacter_cuniculorum	0.01298	uncultured_*Fibrobacteres*	0.00055
uncultured_*Gracilibacter*	0.01006	*Lactobacillus_sp.*	0.00054
Hebe_yellow	0.00823	*Bifidobacterium_longum*	0.00037
Klebsiella_pneumoniae	0.00768	*Blautia_sp.*	0.00018
uncultured_*Ruminiclostridium*	0.00695	uncultured_*Clostridia*	0.00018

SI Table 2.6.2: 45 bacterial species were identified in the Ju|'hoansi IM in 2021.

Species	**Relative Abundance**	**Species**	**Relative Abundance**
Gemella_sp.	0.37649	*Bifidobacterium_longum*	0.00165
Leptotrichia_sp.	0.3198	*Blastopirellula_marina*	0.0011
uncultured_*Candidatus*	0.23953	*Bacillus_firmus*	0.01024
Bacteroides_sp.	0.18815	*Ruminococcus_albus*	0.00896
Rhodococcus_erythropolis	0.12781	uncultured_*Ruminococcus*	0.00896
uncultured_*Bacteroidetes*	0.09874	*Paenarthrobacter_aurescens*	0.0064
uncultured_*Stenotrophomonas*	0.09051	*Proteus_mirabilis*	0.00512
uncultured_cyanobacterium	0.08045	*Chlamydia_abortus*	0.00457
Bacteroidales_genomosp.	0.07076	*Hebe_yellow*	0.00421
Anaerostipes_hadrus	0.0554	*Lactobacillus_salivarius*	0.00329
Ruminococcus_flavefaciens	0.04809	uncultured_*Ruminiclostridium*	0.00329
Parabacteroides_gordonii	0.04644	*Primula_veris*	0.00293
Clostridiales_bacterium	0.04242	*Klebsiella_pneumoniae*	0.00274
Gracilibacteria_bacterium	0.04224	uncultured_*Ruminococcaceae*	0.00201
Prevotella_loescheii	0.0384	*Eubacterium_sp.*	0.00183
uncultured_*Campylobacteraceae*	0.02816	*Sphingomonas_sp.*	0.00165
Christensenella_minuta	0.02761	*Escherichia_coli*	0.00091
Streptococcus_sp.	0.02487	*Prevotella_sp.*	0.00073
Pseudomonas_sp.	0.02468	*Corynebacterium_mastitidis*	0.00037
Mycobacterium_tuberculosis	0.02322	uncultured_*Blautia*	0.00018
uncultured_*Firmicutes*	0.01993	uncultured_*Bradyrhizobium*	0.00018
Acinetobacter_sp.	0.01426	uncultured_*Spirochaetes*	0.00018
Campylobacter_cuniculorum	0.01207		

SI Table 3.1: 2019 core microbiome.

Phylum	**Class**	**Order**	**Family**	**Genus**	**Species**
Proteobacteria	Gammaproteobacteria	Enterobacterales	*Enterobacteriaceae*	-	-
Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Prevotella*	uncultured_bacterium
Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Prevotella*	-
Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Prevotella*	-
Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Prevotella*	unidentified
Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Prevotella*	uncultured_bacterium
Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Prevotella*	unidentified
Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Prevotella*	unidentified
Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Prevotella*	unidentified
Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Prevotella*	unidentified
Firmicutes	Negativicutes	Acidaminococcales	*Acidaminococcaceae*	*Phascolarctobacterium*	uncultured_bacterium
Firmicutes	Clostridia	Oscillospirales	*Ruminococcaceae*	*Faecalibacterium*	uncultured_bacterium
Firmicutes	Clostridia	Oscillospirales	*Ruminococcaceae*	*Faecalibacterium*	uncultured_bacterium
Firmicutes	Clostridia	Oscillospirales	*Ruminococcaceae*	*Faecalibacterium*	uncultured_bacterium
Firmicutes	Clostridia	Oscillospirales	*Ruminococcaceae*	*Faecalibacterium*	uncultured_bacterium
Firmicutes	Clostridia	Oscillospirales	*Ruminococcaceae*	*Ruminococcus*	uncultured_bacterium
Firmicutes	Clostridia	Oscillospirales	*Oscillospiraceae*	*UCG-002*	uncultured_bacterium
Firmicutes	Clostridia	Lachnospirales	*Lachnospiraceae*	*Blautia*	uncultured_bacterium
Firmicutes	Clostridia	Lachnospirales	*Lachnospiraceae*	*Blautia*	uncultured_bacterium
Firmicutes	Clostridia	Lachnospirales	*Lachnospiraceae*	*Blautia*	uncultured_bacterium
Firmicutes	Clostridia	Lachnospirales	*Lachnospiraceae*	*[Ruminococcus]_torques_group*	uncultured_bacterium
Firmicutes	Clostridia	Lachnospirales	*Lachnospiraceae*	*[Ruminococcus]_torques_group*	uncultured_bacterium
Firmicutes	Clostridia	Lachnospirales	*Lachnospiraceae*	*Lachnospiraceae_ND3007_group*	uncultured_bacterium
Firmicutes	Clostridia	Lachnospirales	*Lachnospiraceae*	*Dorea*	uncultured_bacterium
Firmicutes	Clostridia	Lachnospirales	*Lachnospiraceae*	*Dorea*	uncultured_bacterium

SI Table 3.2: 2021 core microbiome.

Phylum	**Class**	**Order**	**Family**	**Genus**	**Species**
Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Prevotella*	uncultured_bacterium
Bacteroidota	Bacteroidia	Bacteroidales	*Prevotellaceae*	*Prevotella*	uncultured_bacterium

SI Table 5.1: Bray-Curtis PERMANOVA results looking at 1) Year, 2) Subject, 3) Village, 4) Gender, 5) Age, 6) Travel, 7) Dietary Preferences in Winter, 8) Intestinal Infections, 9) Treatment of Intestinal Infections, 10) Malaria, 11) Malaria Medication, 12) Tuberculosis and 13) Antibiotics.

Group	**PERMANOVA R-squared**	**PERMANOVA F-statistic**	**PERMANOVA p value**	**Significance**
Year	0.1271	10.484	1×10^{-4}	***
Subject	0.43695	0.7976	1	-
Village	0.04575	1.1188	0.1564	-
Gender	0.01198	0.8728	0.7302	-
Age	0.5191	0.941	0.9377	-
Travel	0.05787	1.4333	0.0032	**
Dietary Preferences in Winter	0.0918	3.5885	1×10^{-4}	***
Intestinal Infections	0.03045	1.115	0.1748	-
Treatment of Intestinal Infections	0.07782	1.4557	4×10^{-4}	***
Malaria	0.01427	1.042	0.297	-
Use of Malaria Medication	0.02586	0.9424	0.6073	-
Tuberculosis	0.03091	2.2962	0.0014	**
Antibiotics	0.07842	3.0209	1×10^{-4}	***

SI Table 5.2: Bray-Curtis PERMANOVA results looking at 3) Village, 4) Gender, 5) Age, 6) Travel, 7) Dietary Preferences in Winter, 8) Intestinal Infections, 9) Treatment of Intestinal Infections, 10) Malaria, 11) Malaria Medication, 12) Tuberculosis and 13) Antibiotics.

Group	**PERMANOVA R-squared**	**PERMANOVA F-statistic**	**PERMANOVA p value**	**Significance**
Village	0.11226	1.5174	8×10^{-4}	***
Gender	0.02505	0.9763	0.493	-
Age	0.02874	1.1246	0.2167	-
Travel	0.04153	0.8017	0.9555	-
Dietary Preferences in Winter	0.03697	1.4587	0.0306	*
Intestinal Infections	0.06464	1.2786	0.0423	*
Treatment of Intestinal Infections	0.12403	1.2389	0.033	*
Malaria	0.02856	1.1171	0.2084	-
Use of Malaria Medication	0.0515	1.0044	0.4404	-
Tuberculosis	0.0352	1.3863	0.0441	*
Antibiotics	0.05922	1.1645	0.1261	-

SI Table 5.3: Bray-Curtis PERMANOVA results looking at 3) Village, 4) Gender, 5) Age, 6) Travel, 7) Dietary Preferences in Winter, 8) Intestinal Infections, 9) Treatment of Intestinal Infections, 10) Malaria, 11) Malaria Medication, 12) Tuberculosis and 13) Antibiotics.

Group	**PERMANOVA R-squared**	**PERMANOVA F-statistic**	**PERMANOVA p value**	**Significance**
Village	0.09417	1.1436	0.0473	*
Gender	0.02623	0.9428	0.6391	-
Age	0.02851	1.0272	0.3731	-
Travel	0.09158	1.109	0.0863	.
Dietary Preferences in Winter	0.05422	0.9746	0.5747	-
Intestinal Infections	0.05316	0.9545	0.682	-
Treatment of Intestinal Infections	0.05544	0.9977	0.4885	-
Malaria	0.02823	1.0167	0.4023	-
Use of Malaria Medication	0.05857	1.0576	0.2245	-
Tuberculosis	0.03502	1.27	0.0387	*
Antibiotics	0.0624	1.1314	0.074	.

SI Table 6.1: Differentially abundant bacterial genera between samples taken from different years.

Genus	Adj. p-value	log2Fold Change	Genus	Adj. p-value	log2Fold Change
Prevotella	0.03773	5.22419	*Ruminococcus*	0.03928	-5.3954
Prevotella	0.01416	6.21069	*Dorea*	2.16×10^{-6}	-5.94769
Prevotella	0.0423	5.86766	*Lachnospiraceae_UCG-008*	7.27×10^{-5}	-6.30754
Prevotella	0.00933	-6.84486	[*Eubacterium*]_ruminantium_group	4.05×10^{-7}	-7.33214
Prevotella	0.02981	-6.65325	[*Eubacterium*]_ruminantium_group	5.63×10^{-6}	-6.91623
Prevotella	0.01931	5.94424	*Lachnospiraceae_ND3007_group*	9.12×10^{-7}	-6.7212
Prevotella	1.49×10^{-17}	21.58131	*Anaerostipes*	0.0035	-6.74698
Prevotella	2.92×10^{-20}	23.22289	*Butyrivibrio*	0.01531	-5.62313
Prevotella	0.00815	5.52694	*Lactobacillus*	0.00896	5.44341
Prevotella	1.87×10^{-11}	-8.26282	*CAG-56*	3.44×10^{-23}	-24.93929
Prevotella	0.00289	-6.87543	*Agathobacter*	0.00051	-6.41502
Prevotella	1.12×10^{-6}	7.33392	*Roseburia*	0.03773	-7.19601
Prevotella	1.17×10^{-20}	-10.01799	*Agathobacter*	0.03346	-5.99767
Prevotella	0.00383	6.04187	*Roseburia*	1.19×10^{-28}	-24.61481
Prevotella	5.2×10^{-15}	-8.40054	[*Ruminococcus*]_torques_group	1.02×10^{-8}	-7.58637
Prevotella	0.00831	6.26148	[*Ruminococcus*]_torques_group	2.3×10^{-17}	-9.29543
Prevotella	3.77×10^{-9}	-8.53229	[*Ruminococcus*]_torques_group	0.02981	5.91147
Prevotella	2.65×10^{-14}	-9.84964	*Blautia*	3.4×10^{-26}	-22.28677
Prevotella	0.00933	6.15742	*Blautia*	0.00025	6.84525
Prevotella	0.01836	-6.25186	*Blautia*	2.59×10^{-23}	-10.0179
Prevotella	1.12×10^{-15}	-9.75364	*Blautia*	0.00063	-6.59358
Prevotella	6.88×10^{-6}	6.67837	*Blautia*	6.14×10^{-20}	22.78046
Prevotella	0.03738	6.05547	*Blautia*	5.01×10^{-25}	-9.15443
Prevotella	0.00121	-7.13465	*Blautia*	1.46×10^{-13}	-7.58963
Prevotella	1.57×10^{-23}	-9.39263	*Blautia*	0.00711	-5.80826
Prevotella	0.0005	-6.3018	[*Ruminococcus*]_gauvreauii_group	0.00933	-5.94133
Prevotella	1.39×10^{-8}	-8.10613	*Dorea*	1.12×10^{-6}	-6.21563
Prevotella	0.00014	6.73542	*Lachnoclostridium*	0.00379	-6.23591
Prevotella	0.0061	5.71372	*Lachnospiraceae_NK4A136_group*	1.63×10^{-43}	-25.46811
Prevotella	0.01969	-6.14385	*Leptotrichia*	1.95×10^{-17}	22.61215
Prevotella	0.00334	6.15713	*Fusobacterium*	8.07×10^{-44}	24.96149
Prevotella	0.0011	5.72251	*Fusobacterium*	1.13×10^{-20}	23.92865
Prevotella	1.1×10^{-7}	-7.13766	*Fusobacterium*	2.67×10^{-18}	23.64013
Prevotella	6.16E-05	-7.30839	*Streptococcus*	9.07×10^{-8}	11.11306
Prevotella	0.01931	-5.07699	*Streptococcus*	1.91×10^{-5}	9.25263
Prevotella	0.02885	-5.74407	*Streptococcus*	6.29×10^{-27}	24.9995
Prevotella	4.95×10^{-20}	22.91718	*Christensenella*	0.0031	-6.84402
Prevotella	0.04786	5.72137	*Clostridia_vadinBB60_group*	0.00022	-6.02535
Prevotella	1.1×10^{-33}	25.39412	*Phascolarctobacterium*	0.00498	-6.47828
Prevotella	4.68×10^{-29}	24.76965	*Phascolarctobacterium*	0.03621	-6.42001
Prevotella	7.05×10^{-41}	25.45355	*Phascolarctobacterium*	0.04032	-5.75023
Prevotella	0.00971	-5.52885	*Phascolarctobacterium*	0.00082	-6.80153
Prevotella	0.00172	6.21625	*Phascolarctobacterium*	0.00757	-6.10782
Alloprevotella	0.02175	5.83169	*Phascolarctobacterium*	2.64×10^{-15}	-8.4854
Rikenellaceae_RC9_gut_group	0.04457	-5.25964	*Veillonella*	0.04705	5.06055
Bacteroides	0.0198	-6.58929	*Veillonella*	0.00033	7.36729
Colidextribacter	0.00648	-5.8106	*Veillonella*	0.00229	6.39056
UCG-002	3.55×10^{-9}	-7.27057	*Veillonella*	4.77×10^{-23}	23.66295
UCG-002	0.00014	-6.93628	*Veillonella*	4.29×10^{-23}	23.69307
UCG-002	0.00017	-7.23989	*Veillonella*	0.01932	5.51896
UCG-002	0.03738	-5.08995	*Veillonella*	0.03134	5.82809
Ruminococcus	1.32×10^{-10}	-9.06343	*Veillonella*	2.01×10^{-20}	23.37894
Ruminococcus	4.53×10^{-18}	-24.72039	*Veillonella*	6.35×10^{-23}	23.52667
Ruminococcus	0.03773	-6.28091	*Selenomonas*	0.00017	-7.2982
UCG-005	0.04556	-5.6108	*Erysipelatoclostridium*	0.00187	-6.74855

SI Table 6.1: Differentially abundant bacterial genera between samples taken from different years. (cont.)

Genus	**Adj. p-value**	**log2Fold Change**	**Genus**	**Adj. p-value**	**log2Fold Change**
UCG-005	5.46×10^{-5}	-5.8067	*Gemella*	0.03928	5.15682
UCG-010	0.03737	-4.72310	*Gemella*	0.00092	6.61123
Faecalibacterium	0.02734	5.56676	*Treponema*	0.01748	-6.32138
Faecalibacterium	0.03928	5.17451	*Clostridium_sensu_stricto_1*	1.28×10^{-29}	-25.99599
Faecalibacterium	2.59×10^{-23}	-8.67702	*Clostridium_sensu_stricto_1*	4.8×10^{-32}	-25.18382
Faecalibacterium	0.01609	-5.15175	*Clostridium_sensu_stricto_1*	0.00253	-6.94437
Faecalibacterium	5.07×10^{-29}	-24.71267	*Romboutsia*	3.34×10^{-29}	-25.39288
Faecalibacterium	0.02614	-6.29964	*Succinivibrio*	4.52×10^{-10}	-10.69132
Faecalibacterium	0.03287	5.78741	*Succinivibrio*	1.42×10^{-24}	22.24016
Faecalibacterium	3.85×10^{-21}	-9.92727	*Succinivibrio*	1.98×10^{-33}	-26.77907
Faecalibacterium	4.2×10^{-10}	-7.93464	*Aggregatibacter*	6.52×10^{-29}	23.86661
Faecalibacterium	1.74×10^{-10}	-7.41566	*Aggregatibacter*	1.17×10^{-20}	23.69122
Subdoligranulum	0.00603	-5.92325	*Aggregatibacter*	7.8×10^{-26}	23.83291
Subdoligranulum	0.00019	-6.30731	*Neisseria*	7.26×10^{-26}	23.97587
uncultured	0.0319	-6.06601	*Neisseria*	7.02×10^{-36}	24.52749
[*Eubacterium*]_coprostanoligenes_group	4.89×10^{-5}	-8.71177	*Neisseria*	1.17×10^{-20}	23.69853
Hydrogenoanaerobacterium	0.02195	-6.00521	*Desulfovibrio*	0.03737	-5.25837
[*Eubacterium*]_siraeum_group	0.00136	-6.18886	*TM7a*	1.78×10^{-22}	20.89626
Ruminococcus	0.00222	-7.12185	*TM7a*	0.00148	6.70238

SI Table 6.2: Differentially abundant bacterial genera between samples from different villages.

Genus	**Adj. p-value**	**log2FoldChange**
[*Eubacterium*]_coprostanoligenes_group	0.00035	-19.95688
Anaerostipes	5.72×10^{-9}	-23.73010
Blautia	0.00028	22.40824
Streptococcus	0.00146	20.62368
Veillonella	0.00143	20.76665
Succinivibrio	0.00185	-20.39012
Succinivibrio	0.00012	-23.33072
Succinivibrio	2.24×10^{-7}	-25.50905

SI Table 6.4: Differentially abundant bacterial genera for travel.

Genus	Adj. p-value	log2FoldChange
Prevotella	0.02785	18.83369
Prevotella	0.0086	-23.07694
Prevotella	0.00023	18.92323
Prevotella	0.00122	19.47719
Prevotella	0.0086	-22.97591
Prevotella	0.0086	-23.03116
Alloprevotella	0.00836	-22.53046
Bacteroides	0.0086	-23.44739
Odoribacter	0.0086	-23.70165
Porphyromonas	0.0086	-23.92036
Bacteroides	0.0086	-23.44625
Bacteroides	0.00091	-29.86095
Bacteroides	0.0086	-23.49146
Ruminococcus	0.0086	-24.15377
Faecalibacterium	0.0086	-24.67312
[Ruminococcus]_gnavus_group	0.00836	-24.79641
Butyrivibrio	0.00997	-22.78945
[Ruminococcus]_torques_group	2.6×10^{-8}	21.63728
Fusobacterium	0.0086	-23.25792
Streptococcus	0.00099	-29.64967
Veillonella	0.0086	-23.06536
Veillonella	0.01477	-19.75658
Veillonella	0.00893	-22.85565
Veillonella	0.00341	-27.33429
Gemella	0.0086	-23.15001
Succinivibrio	0.0086	-21.17809
Neisseria	0.00836	-25.09964
Neisseria	0.00588	-24.51665
Neisseria	0.0086	-23.53556
Neisseria	0.03678	19.95175
Neisseria	0.00836	-24.79193
Neisseria	0.0086	-22.99055
TM7a	0.00588	19.92122

SI Table 6.5: Differentially abundant bacterial genera for dietary preferences in winter.

Genus	Adj. p-value	log2FoldChange
Prevotella	3.28×10^{-13}	-23.0741
Prevotella	0.04104	5.01596
Prevotella	0.02053	5.25073
Prevotella	1.03×10^{-25}	-25.9999
UCG-002	0.00288	6.9567
Ruminococcus	0.02988	7.14216
Roseburia	3.89×10^{-12}	23.79687
[*Ruminococcus*]_torques_group	0.00206	5.21471
Leptotrichia	8×10^{-14}	-23.57935
Fusobacterium	0.01712	-6.91469
Fusobacterium	1.2×10^{-11}	-23.24761
Streptococcus	0.00206	-8.87278
Veillonella	5.86×10^{-18}	-24.48354
Romboutsia	6.69×10^{-20}	23.74619
Succinivibrio	3.31×10^{-19}	-23.96796
Aggregatibacter	1.78×10^{-11}	-23.04783
Aggregatibacter	1.04×10^{-17}	-24.39057
Neisseria	4.34×10^{-16}	-24.48633

SI Table 6.6: Differentially abundant bacterial genera for intestinal infections.

Genus	Adj. p-value	log2FoldChange
Prevotella	1.34×10^{-5}	-19.3874
Prevotella	3.18×10^{-5}	-18.80803
Prevotella	3.62×10^{-15}	-26.06069
Bacteroides	0.00098	-16.34065
Faecalibacterium	1.58×10^{-13}	-22.76802
Blautia	0.00016	-17.74123
Leptotrichia	1.39×10^{-6}	-20.9421
Phascolarctobacterium	0.00021	-17.49481
Veillonella	0.00001	-19.99688
Treponema	0.00062	-16.71658
Treponema	0.01831	-14.09325
Succinivibrio	1.03×10^{-5}	-19.62276

SI Table 6.7: Differentially abundant bacterial genera for malaria.

Genus	Adj. p-value	log2FoldChange
Prevotella	4.93×10^{-6}	-23.74691
Blautia	5.43×10^{-7}	-22.01647
Streptococcus	1.3×10^{-6}	-24.13589
Veillonella	5.43×10^{-7}	-22.89358
Succinivibrio	5.19×10^{-11}	-26.66188

SI Table 6.8: Differentially abundant bacterial families for tuberculosis.

Family	Adj. p-value	log2FoldChange
Tannerellaceae	4.12×10^{-18}	-24.09395
Tannerellaceae	9.67×10^{-14}	-23.85277

SI Table 6.9: Differentially abundant bacterial genera for antibiotics.

Genus	Adj. p-value	log2FoldChange
Prevotella	4.47×10^{-16}	-23.58663
Prevotella	0.0467	-5.91688
Prevotella	2.34×10^{-20}	-23.94643
Faecalibacterium	0.04358	4.22568
[*Ruminococcus*]_torques_group	9.96×10^{-15}	-21.77356
Blautia	0.0172	-6.60015
Fusobacterium	7.52×10^{-16}	-24.1424
Streptococcus	2.8×10^{-24}	-25.52229
Veillonella	5.72×10^{-16}	-23.57705
Veillonella	3.41×10^{-14}	-23.17216
Gemella	1.11×10^{-25}	-23.47158
Succinivibrio	2.19×10^{-19}	-23.82889
Aggregatibacter	8.53×10^{-20}	-23.71288
Aggregatibacter	3.41×10^{-14}	-23.35397
Aggregatibacter	6.53×10^{-16}	-23.60441
Neisseria	5.47×10^{-18}	-23.91733
Neisseria	1.86×10^{-22}	-23.5964
TM7a	1.94×10^{-19}	-23.50703
Candidatus_Saccharimonas	2.17×10^{-9}	-20.03244

SI Table 7.1: Correlations for villages in 2021.

Factor	p-value
Total organic carbon	0.036
Tuberculosis	0.017
Ammonium	0.0041
Chloride	0.0041
Turbidity	0.0041
Decanoic acid	8.94954×10^{-6}
pH	7.13825×10^{-7}
Total hardness	7.13825×10^{-7}
Copper	3.17269×10^{-8}
Lead	3.17269×10^{-8}
Nickel	3.17269×10^{-8}
Calcium	0
Potassium	0

SI Table 7.2: Correlations for tuberculosis in 2021.

Factor	p-value
Calcium	0.01781
Potassium	0.01781
Village	0.01781
Ammonium	0.00816
Chloride	0.00816
Turbidity	0.00816
pH	0.00101
Alkalinity	0.00036
Electrical conductivity	0.00036
Fluoride	0.00036
Sodium	0.00036
Sulphate	0.00036
Total dissolved solids	0.00036
Total organic carbon	0.00017
Copper	0.00004
Nickel	0.00004
Lead	0.00004

SI Table 7.3: Correlations for dietary preferences in winter in 2021.

Factor	p-value
Ammonium	0.03376
Chloride	0.03376
Turbidity	0.03376
Manganese	0.02112
Total coliform	0.02112
Decanoic acid	0.01989

SI Table 7.4: Correlations for antibiotics in 2021.

Factor	p-value
Nitrate	0.03861
Total oxidised nitrogen	0.03861
Total hardness	0.02381
Magnesium	0.00413

SI Table 8: The water microbiome comprised 8 identified bacterial phyla.

Phylum	**Relative Abundance**
Bacteroidia	2.54397
Firmicutes	2.45
Proteobacteria	0.85172
Fusobacteriota	0.67586
Patescibacteria	0.275
Actinobacteriota	0.18448
Campilobacterota	0.01638
Nitrospirota	0.00259

SI Table 9: Differentially abundant bacterial genera for significant water chemical composition factors.

Chemical component	**Genus**	**Adj. p-value**	**log2FoldChange**
Nitrate	[*Ruminococcus*]_torques_group	0.00495	-5.449745
Nitrate	*Streptococcus*	3.28×10^{-13}	9.85679
Nitrate	*Selenomonas*	1.54×10^{-83}	-23.85534
Langelier saturation index	*Prevotella*	0.01193	20.59427
Langelier saturation index	*Phascolarctobacterium*	0.00606	21.49698
Langelier saturation index	*Blautia*	0.00429	21.99429
Langelier saturation index	*Prevotella*	0.00059	-24.307
Langelier saturation index	*Veillonella*	0.00018	-25.62286
Langelier saturation index	*Bacteroides*	2.38×10^{-5}	-27.70176
Langelier saturation index	*Prevotella*	2.14×10^{-5}	27.95192
Langelier saturation index	*Ruminococcus*	2.14×10^{-5}	28.07344
Langelier saturation index	*Veillonella*	2.14×10^{-5}	-28.00157
Langelier saturation index	*Prevotella*	2.03×10^{-5}	28.60027
Langelier saturation index	*Selenomonas*	2.03×10^{-5}	28.50218
Langelier saturation index	*Veillonella*	2.03×10^{-5}	-28.34034
Total oxidised nitrogen	[*Ruminococcus*]_torques_group	0.00495	-5.449745
Total oxidised nitrogen	*Streptococcus*	3.28×10^{-13}	9.85679
Total oxidised nitrogen	*Selenomonas*	1.54×10^{-83}	-23.85534

Supplementary R Code

Load the R packages

```
library(DECIPHER)
library(ape)
library(DESeq2)
library(ggplot2)
library(phyloseq)
library(plotly)
library(vegan)
library(philr)
library(tidyverse)
library(adespatial)
library(devtools
library(qiime2R)
library(MicrobeR)
library(microbiome)
library(microbiomeSeq)
library("pander")
library(ranacapa)
library(grid)
library(gridExtra)
library(knitr)
library(png)
library("ggdendro")
library(ggpubr)
library(RColorBrewer)
library(microbiomeutilities)
library(rstatix)
library(pheatmap)
library(ggrepel)
library(Hmisc)
```

Load the data

```
t <- read.table("merged_table.tsv", header=T, sep="\t")
t <- t[,-c(44)]
t <- column_to_rownames(t, var="OTU.ID")
t <- as.matrix(t)
OTU = otu_table(t, taxa_are_rows = TRUE)

taxonomy <- read.table ("merged_taxonomy.tsv",header=T,sep="\t")
taxonomy <- taxonomy %>% separate(Taxon, c("Domain", "Phylum", "Class", "Order", "Family",
"Genus", "Species"), sep = ";")
taxonomy <- column_to_rownames(taxonomy, var="Feature.ID")
taxonomy$Domain <- gsub("d__", "",taxonomy$Domain)
taxonomy$Phylum <- gsub(" p__", "",taxonomy$Phylum)
taxonomy$Class <- gsub(" c__", "",taxonomy$Class)
taxonomy$Order <- gsub(" o__", "",taxonomy$Order)
taxonomy$Family <- gsub(" f__", "",taxonomy$Family)
taxonomy$Genus <- gsub(" g__", "",taxonomy$Genus)
taxonomy$Species <- gsub(" s__", "",taxonomy$Species)
taxonomy <- as.matrix(taxonomy)
TAX = tax_table(taxonomy)
TAXA = tax_table(taxonomy)[,-8]

metadata <- import_qiime_sample_data("merged_metadata.txt")
metadata = metadata[-1,]
tree <- read_tree("merged_tree.nwk")

physeq <- phyloseq(OTU, TAX, metadata, tree)
phy <- phyloseq(OTU, TAXA, metadata, tree)
```

Delete unwanted taxa

```
physeq <- physeq %>%
  subset_taxa(Domain != "Archaea",
              Family  != "Mitochondria" &
              Order  != "Chloroplast")

phy <- phy %>%
  subset_taxa(Domain != "Archaea",
              Family  != "Mitochondria" &
              Order  != "Chloroplast")
```

Remove withdrawn participants

```
physeq <- subset_samples(physeq, sample_names(physeq) != "TOM03")
physeq <- subset_samples(physeq, sample_names(physeq) != "TOM06")
physeq <- subset_samples(physeq, sample_names(physeq) != "TOM07")

phy <- subset_samples(phy, sample_names(phy) != "TOM03")
phy <- subset_samples(phy, sample_names(phy) != "TOM06")
phy <- subset_samples(phy, sample_names(phy) != "TOM07")
```

Remove singletons

```
physeq <-prune_taxa(taxa_sums(physeq) > 1, physeq)
physeq

phy <-prune_taxa(taxa_sums(phy) > 1, phy)
phy
```

Check features of data

```
summarize_phyloseq(physeq)
print_ps(physeq)
summary(sample_sums(physeq))
```

Rarefy the data

```
physeq_rarefy <- rarefy_even_depth(physeq, rngseed=1, sample.size=5469, replace=FALSE)
phy_rarefy <- rarefy_even_depth(phy, rngseed=1, sample.size=5469, replace=FALSE)
```

Composition plots

```
physeq_phy <- microbiome::aggregate_rare(phy_rarefy, level = "Phylum", detection = 0.1/100, prevalence = 1/100, include.lowest = TRUE)

physeq.phy.rel <- microbiome::transform(physeq_phy, "compositional")

physeq.phy.rel <- physeq %>%
  aggregate_rare(level = "Phylum", detection = 0.1/100, prevalence = 1/100, include.lowest = TRUE) %>%
  microbiome::transform(transform = "compositional")

plot_composition(physeq.phy.rel,sample.sort = "year", x.label = "SampleID") + theme(legend.position = "bottom") + scale_fill_brewer("Phylum", palette = "Spectral") + theme_bw() + theme(axis.text.x = element_text(angle = 90)) + ggtitle("Relative abundance") + theme(legend.title = element_text(size = 18))
```

Heatmaps

```
p <- plot_taxa_heatmap(phy_rarefy,
                       subset.top = 25,
                       VariableA = ("year"),
                       transformation = "log10",
                       cluster_rows = T,
                       cluster_cols = F,
                       show_colnames = F,
                       heatcolors = colorRampPalette(rev(brewer.pal(n = 7, name = "RdYlBu")))(100),)

h.map <- plot_heatmap(physeq.phy.rel, method="PCoA", distance="bray", taxa.label = "Phylum", sample.order = unique(sample_names(physeq))) + facet_grid(~year, scales = "free_x", drop = TRUE) + theme_bw() + theme(axis.text.x = element_text(face = "bold", angle = 45, hjust = 1)) + theme(legend.key = element_blank(),strip.background = element_rect(colour="black", fill="white"))
h.map <- h.map + theme(text = element_text(size = 15))
h.map <- h.map + theme(axis.text.y = element_text(colour = 'black', size = 15, face = 'italic'))
h.map <- h.map + scale_fill_distiller("Abundance", palette = "RdYlBu")
h.map <- h.map + rremove("x.text")
print(h.map)
```

```
mycols <- c("red", "blue", "red3", "sienna2")
comps <- make_pairs(sample_data(phy.f)$year)
```

Determine dominant taxa

```
physeq.gen <- aggregate_taxa(phy_rarefy,"Genus")

dom.tax <- dominant_taxa(phy_rarefy,level = "Genus", group="year")
head(dom.tax$dominant_overview)
```

Calculate relative abundance

```
phy.rarefy.rel.abund <- transform_sample_counts(phy_rarefy, function(x) x / sum(x) )
rel.abund <- phy.rarefy.rel.abund %>% psmelt()
write.csv(rel.abund, file = "rel.abund.csv")
```

Alpha diversities

```
plot_richness(phy_rarefy, measures="Simpson")
plot_richness(phy_rarefy, measures=c("Simpson", "Shannon"))+theme(text=element_text(size=12))
```

Figure 4: Plot observed, Shannon and Simpson richness

```
a.div <- plot_richness(phy_rarefy, x="year", measures=c("Shannon", "simpson", "Observed"), color =
"year") + geom_boxplot() + theme_bw() + theme(text=element_text(size=15)) +
scale_colour_brewer(palette="Set1")
```

Adding statistical support

```
a.div + stat_compare_means(
  comparisons = comps,
  label = "p.signif",
  tip.length = 0.05,
  symnum.args = list(
    cutpoints = c(0, 0.0001, 0.001, 0.01, 0.05, 1),
    symbols = c("xxxx", "***", "**", "*", "ns")
  ),
  method = "wilcox.test")

richness <- estimate_richness(phy_rarefy,measures=c("Shannon", "simpson", "Observed"))
write.csv(richness, file = "alpha_div.csv")
```

Shannon diversity

```
sample_data(phy_rarefy)$year <- factor(sample_data(phy_rarefy)$year, levels = c("2019", "2021"))
diver <-estimate_richness(phy_rarefy, measures=c("Shannon"))
data.tests <- cbind(sample_data(phy_rarefy), diver)
data.tests1 <-gather(data.tests, key=diversity, value=measure, Shannon)
```

Figure 5.A: Shannon plot for year

```
ggplot(data.tests1, aes(x=year, y=measure, color=year)) +
  geom_boxplot() +
  facet_grid(~village) +
  geom_point(size=1, alpha=0.7) +
  theme_bw(base_size = 14) +
  theme(axis.text.x= element_text(angle = 45, hjust = 1)) +
  xlab("year") +
  ylab("shannon diversity index") +
  scale_color_manual(values = c("#48A11C", "#5CE8FF")) +
  theme(panel.background = element_blank(), panel.grid.major = element_blank(), panel.grid.minor =
element_blank(),
        panel.border = element_rect(fill = NA), axis.text.x=element_blank(),
axis.ticks.x=element_blank(), axis.title.x = element_blank(),
        strip.background = element_blank()) + ggtitle("Bacteria shannon diversity per village")
```

Simpson diversity

```
sample_data(phy_rarefy)$year <- factor(sample_data(phy_rarefy)$year, levels = c("2019", "2021"))
diver <-estimate_richness(phy_rarefy, measures=c("Simpson"))
data.tests <- cbind(sample_data(phy_rarefy), diver)
data.tests1 <-gather(data.tests, key=diversity, value=measure, Simpson)
```

Figure 5.B: Simpson plot for year

```
ggplot(data.tests1, aes(x=year, y=measure, color=year)) +
  geom_boxplot() +
  facet_grid(~village) +
  geom_point(size=1, alpha=0.7) +
  theme_bw(base_size = 14) +
  theme(axis.text.x= element_text(angle = 45, hjust = 1)) +
  xlab("year") +
  ylab("simpson diversity index") +
  scale_color_manual(values = c("#48A11C", "#5CE8FF")) +
  theme(panel.background = element_blank(), panel.grid.major = element_blank(), panel.grid.minor =
element_blank(),
        panel.border = element_rect(fill = NA), axis.text.x=element_blank(),
axis.ticks.x=element_blank(), axis.title.x = element_blank(),
        strip.background = element_blank())+
  ggtitle("Bacteria simpson diversity per village")
```

Statistical alpha diversity calculations

```
alphaObserved = estimate_richness(phy_rarefy, measures="Observed")
alphaShannon = estimate_richness(phy_rarefy, measures="Shannon")
alphaSimpson = estimate_richness(phy_rarefy, measures="simpson")
alpha.stats <- cbind(alphaObserved, sample_data(phy_rarefy))
alpha.stats2 <- cbind(alpha.stats, alphaShannon)
alpha.stats3 <- cbind(alpha.stats2, alphaSimpson)
```

For two factors

```
wilcox.test(Simpson~year, data = alpha.stats3)
```

For more than 2 factors

```
pairwise.wilcox.test(alpha.stats3$Shannon, alpha.stats3$village, p.adjust.method="none")
```

Beta diversity

Figure 6: Bray-Curtis PCoA

```
physeq.ord <- ordinate(phy_rarefy, "PCoA", "bray")
b.div.bray <- plot_ordination(physeq_rarefy, physeq.ord, type= "samples", color= "year") +
geom_point(size=3)
b.div.bray <- b.div.bray + stat_ellipse() + ggtitle("Bray-Curtis Dissimilarity for Year")  +
theme_classic() + scale_color_brewer("year", palette = "Set1")
print(b.div.bray)

physeq_rel <- microbiome::transform(phy_rarefy, "compositional")
```

Permanova

```
otu <- abundances(physeq_rel)
meta <- meta(physeq_rel)
```

SI Table 5.1. - 5.3: Statistical beta diversity calculations

```
set.seed(1)
permanova <- adonis2(t(otu) ~ year, data = meta, permutations=9999, method = "bray")
permanova
```

Core microbiota

```
physeq.tmp.rel <- microbiome::transform(physeq_rarefy, "compositional")
physeq.tmp.rel2 <- prune_taxa(taxa_sums(physeq.tmp.rel) > 0, physeq.tmp.rel)
core.taxa.standard <- core_members(physeq.tmp.rel2, detection = 0.001, prevalence = 50/100)
print(core.taxa.standard)
```

Visualize the taxonomy table

```
taxonomy_core <- as.data.frame(tax_table(physeq.tmp.rel2))
core_taxa_id <- subset(taxonomy_core, rownames(taxonomy_core) %in% core.taxa.standard)
DT::datatable(core_taxa_id)
core.abundance <- sample_sums(core(physeq.tmp.rel2, detection = 0.001, prevalence = 50/100))
DT::datatable(as.data.frame(core.abundance))
```

Core heatmaps

```
pseq.unfiltered.rel <- microbiome::transform(physeq_rarefy, "compositional")
pseq.unfiltered.rel@tax_table <-gsub("g__","",pseq.unfiltered.rel@tax_table)
df <- data.frame(pseq.unfiltered.rel@tax_table)
df$Genus <- gsub("unidentified",NA, df$Genus)
tax_table(pseq.unfiltered.rel) <- as.matrix(df)
```

Determine detection threshold and prevalence

```
det <- c(0, 0.1, 0.5, 2, 5, 20)/100
prevalences <- c(10,20,30,40,50,60,70,80,90,100)
p.16<-plot_core(pseq.unfiltered.rel, prevalences = prevalences,
                detections = det, plot.type = "lineplot") +
  xlab("Detection threshold (%)") +
  theme_bw()
```

Figure 3.A. - B. and SI Figure 1:Plot core genera

```
pseq.core.taxa <- aggregate_taxa(pseq.unfiltered.rel,"Genus")
pseq.core.taxa <- core(pseq.core.taxa, detection = 0.01/100, prevalence = .5)
tax <- as.matrix(pseq.core.taxa@tax_table)
tax<- tax[-c(32,46),]
tax_table(pseq.core.taxa) <- tax

p <- plot_core(pseq.core.taxa, plot.type = "heatmap",
               prevalences = prevalences,
               detections = det,
               colours = rev(brewer.pal(5, "Spectral")),
               min.prevalence = .5, horizontal = F)

p + theme(axis.text = element_text(size = 12, face="italic"),
          axis.title = element_text(size=15),
          legend.title = element_text(size=15))+
  coord_fixed(ratio=.2)

prevalences <- seq(.05, 1, .05)
detections <- 10^seq(log10(1e-3), log10(.2), length = 10)
```

```
gray <- gray(seq(0,1,length=5))
p.core <- plot_core(pseq.core.taxa,
                    plot.type = "heatmap",
                    colours = rev(brewer.pal(5, "Spectral")),
                    prevalences = prevalences,
                    detections = det,
                    min.prevalence = .5) +
  xlab("Detection Threshold (Relative Abundance (%))")
print(p.core)

physeq.tmp.rel2.f <- microbiomeutilities::format_to_besthit(physeq.tmp.rel2)

p.core <- plot_core(physeq.tmp.rel2.f,
                    plot.type = "heatmap",
                    colours = rev(brewer.pal(5, "Spectral")),
                    prevalences = prevalences,
                    detections = detections,
                    min.prevalence = .5) +
  xlab("Detection Threshold (Relative Abundance (%))")

p.core + theme(axis.text.y = element_text(face="italic"))
```

Differential abundance testing

Convert phyloseq object to DeSeq

```
bsdds <- phyloseq_to_deseq2(physeq_rarefy, ~ year)
gm_mean <- function(x, na.rm=TRUE){
  exp(sum(log(x[x > 0]), na.rm=na.rm) / length(x))}
geoMeans <- apply(counts(bsdds), 1, gm_mean)
bsdds <- estimateSizeFactors(bsdds, geoMeans = geoMeans)
bsdds <- DESeq(bsdds, test="Wald", fitType="parametric")
```

Create table of results

```
res <- results(bsdds, cooksCutoff = FALSE)
alpha <- 0.05
sigtab <- res[which(res$padj < alpha), ]
sigtab <- cbind(as(sigtab, "data.frame"), as(tax_table(physeq_rarefy)[rownames(sigtab), ], "matrix"))
head(sigtab)
posigtab <- sigtab[sigtab[, "log2FoldChange"] > 0, ]
posigtab <- posigtab[, c("baseMean", "log2FoldChange", "lfcSE", "padj", "Phylum", "Class",
"Family", "Genus")]
```

Write to .csv file

```
sigtabgen <- subset(sigtab, !is.na(Genus))
write.csv(sigtabgen, file = "difabutesgen.csv")
```

Figure 7: Bar plot showing the log2-fold-change with Genus and Phylum

```
x <- tapply(sigtabgen$log2FoldChange, sigtabgen$Phylum, function(x) max(x))
x <- sort(x, TRUE)
sigtabgen$Phylum = factor(as.character(sigtabgen$Phylum), levels=names(x))
x <- tapply(sigtabgen$log2FoldChange, sigtabgen$Genus, function(x) max(x))
x <- sort(x, TRUE)
sigtabgen$Genus = factor(as.character(sigtabgen$Genus), levels=names(x))
ggplot(sigtabgen, aes(y=Genus, x=log2FoldChange, color=Phylum)) +
  geom_vline(xintercept = 0.0, color = "gray", size = 0.5) +
  geom_point(size=2) +
  theme(axis.text.x = element_text(angle = -90, hjust = 0, vjust=0.5)) +
  scale_colour_brewer(palette="Set1")
```

Spearman correlations

```
meta = read.table(file.choose(), header = TRUE, row.names = 1, sep = ",")
meta[, 21:43] = decostand(meta[, 21:43], method = "standardize") # z-score standardize water variables because they are measured in different units
cormatrix = rcorr(as.matrix(meta), type='spearman')
cormatrix$P = p.adjust(cormatrix$P, method = "fdr") # FDR automatically adjust to best value
cormat = signif(cormatrix$r, 2)
pmat = signif(cormatrix$P, 2)

flattenCorrMatrix <- function(cormat, pmat) {
  ut <- upper.tri(cormat)
  data.frame(
    row = rownames(cormat)[row(cormat)[ut]],
    column = rownames(cormat)[col(cormat)[ut]],
    cor  =(cormat)[ut],
    p = pmat[ut]
  )}

cormatrix_table = flattenCorrMatrix(cormatrix$r, cormatrix$P)
write.csv(cormatrix_table, file = "Spearman_correlations.csv")
```

References

Angelakis, E., Bachar, D., Yasir, M., Musso, D., Djossou, F., Gaborit, B., Brah, S., Diallo, A., Ndombe, G. & Mediannikov, O. 2019. Treponema species enrich the gut microbiota of traditional rural populations but are absent from urban individuals. *New microbes and new infections*, 27:14-21.

Asanuma, N., Iwamoto, M. & Hino, T. 1999. Effect of the addition of fumarate on methane production by ruminal microorganisms in vitro. *Journal of Dairy Science*, 82(4):780-787.

Ayeni, F.A., Biagi, E., Rampelli, S., Fiori, J., Soverini, M., Audu, H.J., Cristino, S., Caporali, L., Schnorr, S.L., Carelli, V., Brigidi, P., Candela, M. & Turroni, S. 2018. Infant and Adult Gut Microbiome and Metabolome in Rural Bassa and Urban Settlers from Nigeria. *Cell Reports*, 23(10):3056-3067.

Bäckhed, F., Fraser, C.M., Ringel, Y., Sanders, M.E., Sartor, R.B., Sherman, P.M., Versalovic, J., Young, V. & Finlay, B.B. 2012. Defining a healthy human gut microbiome: current concepts, future directions, and clinical applications. *Cell Host & Microbe*, 12(5):611-622.

Berer, K., Mues, M., Koutrolos, M., Rasbi, Z.A., Boziki, M., Johner, C., Wekerle, H. & Krishnamoorthy, G. 2011. Commensal microbiota and myelin autoantigen cooperate to trigger autoimmune demyelination. *Nature*, 479(7374):538-541.

Berry, D. 2016. The emerging view of Firmicutes as key fibre degraders in the human gut. *Environmental Microbiology*, 18(7):2081-2083.

Bikel, S., López-Leal, G., Cornejo-Granados, F., Gallardo-Becerra, L., García-López, R., Sánchez, F., Equihua-Medina, E., Ochoa-Romo, J.P., López-Contreras, B.E. & Canizales-Quinteros, S. 2021. Gut dsDNA virome shows diversity and richness alterations associated with childhood obesity and metabolic syndrome. *Iscience*, 24(8):102900.

Blaser, M.J. & Falkow, S. 2009. What are the consequences of the disappearing human microbiota? *Nature Reviews Microbiology*, 7(12):887-894.

Bokulich, N.A., Kaehler, B.D., Rideout, J.R., Dillon, M., Bolyen, E., Knight, R., Huttley, G.A. & Gregory Caporaso, J. 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome*, 6(1):1-17.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M. & Asnicar, F. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature biotechnology*, 37(8):852-857.

Bowyer, R.C.E., Schillereff, D.N., Jackson, M.A., Le Roy, C., Wells, P.M., Spector, T.D. & Steves, C.J. 2020. Associations between UK tap water and gut microbiota composition suggest the gut microbiome as a potential mediator of health differences linked to water quality. *Science of the Total Environment*, 739:139697.

Budden, K.F., Gellatly, S.L., Wood, D.L.A., Cooper, M.A., Morrison, M., Hugenholtz, P. & Hansbro, P.M. 2017. Emerging pathogenic links between microbiota and the gut–lung axis. *Nature Reviews Microbiology,* 15(1):55-63.

Callahan, B.J., Mcmurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J. & Holmes, S.P. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods,* 13(7):581-583.

Chaima, D., Pickering, H., Hart, J.D., Burr, S.E., Maleta, K.M., Kalua, K., Bailey, R.L. & Holland, M.J. 2021. Fecal biomarkers of environmental enteric dysfunction and the gut microbiota of rural Malawian children: an observational study. *Heliyon,* 7(10):e08194.

Chen, H., Mozzicafreddo, M., Pierella, E., Carletti, V., Piersanti, A., Ali, S.M., Ame, S.M., Wang, C. & Miceli, C. 2021a. Dissection of the gut microbiota in mothers and children with chronic Trichuris trichiura infection in Pemba Island, Tanzania. *Parasites & Vectors,* 14(1):1-13.

Chen, L., Garmaeva, S., Zhernakova, A., Fu, J. & Wijmenga, C. 2018. A system biology perspective on environment-host-microbe interactions. *Human Molecular Genetics,* 27(R2):R187-R194.

Chen, L., Wang, D., Garmaeva, S., Kurilshikov, A., Vich Vila, A., Gacesa, R., Sinha, T., Lifelines Cohort, S., Segal, E., Weersma, R.K., Wijmenga, C., Zhernakova, A. & Fu, J. 2021b. The long-term genetic stability and individual specificity of the human gut microbiome. *Cell,* 184(9):2302-2315 e2312.

Crotti, S., Agnoletto, E., Cancemi, G., Di Marco, V., Traldi, P., Pucciarelli, S., Nitti, D. & Agostini, M. 2016. Altered plasma levels of decanoic acid in colorectal cancer as a new diagnostic biomarker. *Analytical and bioanalytical chemistry,* 408:6321-6328.

Das, B. & Nair, G.B. 2019. Homeostasis and dysbiosis of the gut microbiome in health and disease. *Journal of Biosciences,* 44(5).

Davenport, E.R., Sanders, J.G., Song, S.J., Amato, K.R., Clark, A.G. & Knight, R. 2017. The human microbiome in evolution. *BMC Biology,* 15(1):127.

Davis, N.M., Proctor, D.M., Holmes, S.P., Relman, D.A. & Callahan, B.J. 2017. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. bioRxiv.

De Bandt, J.-P., Waligora-Dupriet, A.-J. & Butel, M.-J. 2011. Intestinal microbiota in inflammation and insulin resistance: relevance to humans. *Current Opinion in Clinical Nutrition & Metabolic Care,* 14(4):334-340.

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

Devkota, S., Wang, Y., Musch, M.W., Leone, V., Fehlner-Peach, H., Nadimpalli, A., Antonopoulos, D.A., Jabri, B. & Chang, E.B. 2012. Dietary-fat-induced taurocholic acid promotes pathobiont expansion and colitis in Il10−/− mice. *Nature,* 487(7405):104-108.

Dinan, T.G. & Cryan, J.F. 2017. Microbes, Immunity, and Behavior: Psychoneuroimmunology Meets the Microbiome. *Neuropsychopharmacology,* 42(1):178-192.

Ding, R.X., Goh, W.R., Wu, R.N., Yue, X.Q., Luo, X., Khine, W.W.T., Wu, J.R. & Lee, Y.K. 2019. Revisit gut microbiota and its impact on human health and disease. *Journal of Food and Drug Analysis,* 27(3):623-631.

Dixon, P. 2003. VEGAN, a package of R functions for community ecology. *Journal of vegetation science.*

Elvers, K.T., Wilson, V.J., Hammond, A., Duncan, L., Huntley, A.L., Hay, A.D. & Van Der Werf, E.T. 2020. Antibiotic-induced changes in the human gut microbiota for the most commonly prescribed antibiotics in primary care in the UK: a systematic review. *BMJ open,* 10(9):e035677.

Faith, J.J., Guruge, J.L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A.L., Clemente, J.C., Knight, R., Heath, A.C., Leibel, R.L., Rosenbaum, M. & Gordon, J.I. 2013. The long-term stability of the human gut microbiota. *Science,* 341(6141):1237439.

Feng, W., Ao, H. & Peng, C. 2018. Gut Microbiota, Short-Chain Fatty Acids, and Herbal Medicines. *Front Pharmacol,* 9:1354.

Frost, I., Van Boeckel, T.P., Pires, J., Craig, J. & Laxminarayan, R. 2019. Global geographic trends in antimicrobial resistance: the role of international travel. *Journal of Travel Medicine,* 26(8).

Gilbert, J.A., Blaser, M.J., Caporaso, J.G., Jansson, J.K., Lynch, S.V. & Knight, R. 2018. Current understanding of the human microbiome. *Nat Med,* 24(4):392-400.

Gill, S.R., Pop, M., Deboy, R.T., Eckburg, P.B., Turnbaugh, P.J., Samuel, B.S., Gordon, J.I., Relman, D.A., Fraser-Liggett, C.M. & Nelson, K.E. 2006. Metagenomic analysis of the human distal gut microbiome. *Science,* 312(5778):1355-1359.

Gillevet, P.M., Sikaroodi, M. & Torzilli, A.P. 2009. Analyzing salt-marsh fungal diversity: comparing ARISA fingerprinting with clone sequencing and pyrosequencing. *Fungal Ecology,* 2(4):160-167.

Gillings, M.R. & Paulsen, I.T. 2014. Microbiology of the Anthropocene. *Anthropocene,* 5:1-8.

Gomaa, E.Z. 2020. Human gut microbiota/microbiome in health and diseases: a review. *Antonie Van Leeuwenhoek,* 113(12):2019-2040.

Gomez, A., Petrzelkova, K.J., Burns, M.B., Yeoman, C.J., Amato, K.R., Vlckova, K., Modry, D., Todd, A., Jost Robinson, C.A., Remis, M.J., Torralba, M.G., Morton, E., Umana, J.D., Carbonero, F., Gaskins, H.R., Nelson, K.E., Wilson, B.A., Stumpf, R.M., White, B.A., Leigh, S.R. & Blekhman, R. 2016. Gut Microbiome of Coexisting BaAka Pygmies and Bantu Reflects Gradients of Traditional Subsistence Patterns. *Cell Reports,* 14(9):2142-2153.

Hansen, T.H., Kern, T., Bak, E.G., Kashani, A., Allin, K.H., Nielsen, T., Hansen, T. & Pedersen, O. 2018. Impact of a vegan diet on the human salivary microbiota. *Scientific Reports,* 8(1):5847.

Harrell Jr, F.E. & Harrell Jr, M.F.E. 2019. Package 'hmisc'. *CRAN2018,* 2019:235-236.

Hu, Y., Feng, Y., Wu, J., Liu, F., Zhang, Z., Hao, Y., Liang, S., Li, B., Li, J. & Lv, N. 2019. The gut microbiome signatures discriminate healthy from pulmonary tuberculosis patients. *Frontiers in cellular and infection microbiology,* 9:90.

Iliev, I.D., Funari, V.A., Taylor, K.D., Nguyen, Q., Reyes, C.N., Strom, S.P., Brown, J., Becker, C.A., Fleshner, P.R. & Dubinsky, M. 2012. Interactions between commensal fungi and the C-type lectin receptor Dectin-1 influence colitis. *Science,* 336(6086):1314-1317.

Jacobson, D.K., Honap, T.P., Ozga, A.T., Meda, N., Kagone, T.S., Carabin, H., Spicer, P., Tito, R.Y., Obregon-Tito, A.J., Reyes, L.M., Troncoso-Corzo, L., Guija-Poma, E., Sankaranarayanan, K. & Lewis, C.M., Jr. 2021. Analysis of global human gut metagenomes shows that metabolic resilience potential for short-chain fatty acid production is strongly influenced by lifestyle. *Scientific Reports,* 11(1):1724.

Jha, A.R., Davenport, E.R., Gautam, Y., Bhandari, D., Tandukar, S., Ng, K.M., Fragiadakis, G.K., Holmes, S., Gautam, G.P., Leach, J., Sherchand, J.B., Bustamante, C.D. & Sonnenburg, J.L. 2018. Gut microbiome transition across a lifestyle gradient in Himalaya. *PLoS Biology,* 16(11):e2005396.

Jiang, Z., Li, L., Chen, J., Wei, G., Ji, Y., Chen, X., Liu, J. & Huo, J. 2021. Human gut-microbiome interplay: Analysis of clinical studies for the emerging roles of diagnostic microbiology in inflammation, oncogenesis and cancer management. *Infection, Genetics and Evolution,* 93:104946.

Kassambara, A. & Kassambara, M.A. 2020. Package 'ggpubr'. *R package version 0.1.*

Katoh, K., Misawa, K., Kuma, K.I. & Miyata, T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic acids research,* 30(14):3059-3066.

Kolde, R. 2019. pheatmap: Pretty Heatmaps.

Koot, S. 2019. The Limits of Economic Benefits: Adding Social Affordances to the Analysis of Trophy Hunting of the Khwe and Ju/'hoansi in Namibian Community-Based Natural Resource Management. *Society & Natural Resources,* 32(4):417-433.

Krajmalnik-Brown, R., Ilhan, Z.E., Kang, D.W. & Dibaise, J.K. 2012. Effects of gut microbes on nutrient absorption and energy regulation. *Nutrition in clinical practice,* 27(2):201-214.

Kurilshikov, A., Medina-Gomez, C., Bacigalupe, R., Radjabzadeh, D., Wang, J., Demirkan, A., Le Roy, C.I., Raygoza Garay, J.A., Finnicum, C.T., Liu, X., Zhernakova, D.V., Bonder, M.J., Hansen, T.H., Frost, F., Ruhlemann, M.C., Turpin, W., Moon, J.Y., Kim, H.N., Lull, K., Barkan, E., Shah, S.A., Fornage, M., Szopinska-Tokov, J., Wallen, Z.D., Borisevich, D., Agreus, L., Andreasson, A., Bang, C., Bedrani, L., Bell, J.T., Bisgaard, H., Boehnke, M., Boomsma, D.I., Burk, R.D.,

Claringbould, A., Croitoru, K., Davies, G.E., Van Duijn, C.M., Duijts, L., Falony, G., Fu, J., Van Der Graaf, A., Hansen, T., Homuth, G., Hughes, D.A., Ijzerman, R.G., Jackson, M.A., Jaddoe, V.W.V., Joossens, M., Jorgensen, T., Keszthelyi, D., Knight, R., Laakso, M., Laudes, M., Launer, L.J., Lieb, W., Lusis, A.J., Masclee, A.a.M., Moll, H.A., Mujagic, Z., Qibin, Q., Rothschild, D., Shin, H., Sorensen, S.J., Steves, C.J., Thorsen, J., Timpson, N.J., Tito, R.Y., Vieira-Silva, S., Volker, U., Volzke, H., Vosa, U., Wade, K.H., Walter, S., Watanabe, K., Weiss, S., Weiss, F.U., Weissbrod, O., Westra, H.J., Willemsen, G., Payami, H., Jonkers, D., Arias Vasquez, A., De Geus, E.J.C., Meyer, K.A., Stokholm, J., Segal, E., Org, E., Wijmenga, C., Kim, H.L., Kaplan, R.C., Spector, T.D., Uitterlinden, A.G., Rivadeneira, F., Franke, A., Lerch, M.M., Franke, L., Sanna, S., D'amato, M., Pedersen, O., Paterson, A.D., Kraaij, R., Raes, J. & Zhernakova, A. 2021. Large-scale association analyses identify host factors influencing human gut microbiome composition. *Nature genetics,* 53(2):156-165.

Lahti, L. & Shetty, S. 2017. Microbiome r package: Tools for microbiome analysis in r.

Lahti, L., Shetty, S., Blake, T. & Salojarvi, J. 2017. Tools for microbiome analysis in R Version 2.1. 26.

Langelier, W.F. 1946. Chemical equilibria in water treatment. *Journal (American Water Works Association)*:169-178.

Larson, T.E., Buswell, A.M., Ludwig, H.F. & Langelier, W. 1942. Calcium carbonate saturation index and alkalinity interpretations [with discussion]. *Journal (American Water Works Association)*:1667-1684.

Lee, R.B. 1968. The hunters: scarce resources in the Kalahari. *Conformity and Conflict: Readings In Cultural Anthropology*.

Liang, G. & Bushman, F.D. 2021. The human virome: assembly, composition and host interactions. *Nature Reviews Microbiology,* 19(8):514-527.

Lloyd-Price, J., Abu-Ali, G. & Huttenhower, C. 2016. The healthy human microbiome. *Genome Medicine,* 8(1).

Love, M.I., Huber, W. & Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology.*

Lugli, G.A., Longhi, G., Mancabelli, L., Alessandri, G., Tarracchini, C., Fontana, F., Turroni, F., Milani, C., Van Sinderen, D. & Ventura, M. 2022. Tap water as a natural vehicle for microorganisms shaping the human gut microbiome. *Environ Microbiol,* 24(9):3912-3923.

Mancabelli, L., Milani, C., Lugli, G.A., Turroni, F., Ferrario, C., Van Sinderen, D. & Ventura, M. 2017. Meta-analysis of the human gut microbiome from urbanized and pre-agricultural populations. *Environ Microbiol,* 19(4):1379-1390.

Martínez, 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.

Mcmurdie, P.J. & Holmes, S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One.* Public Library of Science.

Mehta, R.S., Abu-Ali, G.S., Drew, D.A., Lloyd-Price, J., Subramanian, A., Lochhead, P., Joshi, A.D., Ivey, K.L., Khalili, H., Brown, G.T., Dulong, C., Song, M., Nguyen, L.H., Mallick, H., Rimm, E.B., Izard, J., Huttenhower, C. & Chan, A.T. 2018. Stability of the human faecal microbiome in a cohort of adult men. *Nature microbiology*, 3(3):347-355.

Mukhopadhya, I., Segal, J.P., Carding, S.R., Hart, A.L. & Hold, G.L. 2019. The gut virome: the 'missing link' between gut bacteria and host immunity? *Therapeutic Advances in Gastroenterology*, 12:1756284819836620.

Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A. & Minh, B.Q. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution,* 32(1):268-274.

Ninkova, V. 2020. Perpetuating the Myth of the "Wild Bushman": Inclusive Multicultural Education for the Omaheke Ju|'hoansi in Namibia. *Comparative Education Review,* 64(2):159-178.

Norman, J.M., Handley, S.A., Baldridge, M.T., Droit, L., Liu, C.Y., Keller, B.C., Kambal, A., Monaco, C.L., Zhao, G. & Fleshner, P. 2015. Disease-specific alterations in the enteric virome in inflammatory bowel disease. *Cell,* 160(3):447-460.

Ochoa-Reparaz, J., Mielcarz, D.W., Haque-Begum, S. & Kasper, L.H. 2010. Induction of a regulatory B cell population in experimental allergic encephalomyelitis by alteration of the gut commensal microflora. *Gut Microbes,* 1(2):103-108.

Odamaki, T., Kato, K., Sugahara, H., Hashikura, N., Takahashi, S., Xiao, J.Z., Abe, F. & Osawa, R. 2016. Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study. *BMC Microbiology,* 16:90.

Olejniczak-Staruch, I., Ciążyńska, M., Sobolewska-Sztychny, D., Narbutt, J., Skibińska, M. & Lesiak, A. 2021. Alterations of the skin and gut microbiome in psoriasis and psoriatic arthritis. *International journal of molecular sciences,* 22(8):3998.

Oliver, A., Chase, A.B., Weihe, C., Orchanian, S.B., Riedel, S.F., Hendrickson, C.L., Lay, M., Sewall, J.M., Martiny, J.B.H. & Whiteson, K. 2021. High-Fiber, Whole-Food Dietary Intervention Alters the Human Gut Microbiome but Not Fecal Short-Chain Fatty Acids. *mSystems,* 6(2).

Pennycook, J.H. & Scanlan, P.D. 2021. Ecological and evolutionary responses to antibiotic treatment in the human gut microbiota. *FEMS Microbiology Reviews*, 45(5):fuab018.

Pérez, J.C. 2021. Fungi of the human gut microbiota: Roles and significance. *International Journal of Medical Microbiology,* 311(3):151490.

Piquer-Esteban, S., Ruiz-Ruiz, S., Arnau, V., Diaz, W. & Moya, A. 2022. Exploring the universal healthy human gut microbiota around the World. *Computational and Structural Biotechnology Journal*, 20:421-433.

Price, M.N., Dehal, P.S. & Arkin, A.P. 2010. FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS One,* 5(3):e9490.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D.R., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J.M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H.B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Li, S., Jian, M., Zhou, Y., Li, Y., Zhang, X., Li, S., Qin, N., Yang, H., Wang, J., Brunak, S., Dore, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., Meta, H.I.T.C., Bork, P., Ehrlich, S.D. & Wang, J. 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature,* 464(7285):59-65.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. & Glöckner, F.O. 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research,* 41(D1):D590-D596.

Quercia, S., Candela, M., Giuliani, C., Turroni, S., Luiselli, D., Rampelli, S., Brigidi, P., Franceschi, C., Bacalini, M.G., Garagnani, P. & Pirazzini, C. 2014. From lifetime to evolution: timescales of human gut microbiota adaptation. *Front Microbiol,* 5:587.

R Core Team, R. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rampelli, S., Schnorr, S.L., Consolandi, C., Turroni, S., Severgnini, M., Peano, C., Brigidi, P., Crittenden, A.N., Henry, A.G. & Candela, M. 2015. Metagenome Sequencing of the Hadza Hunter-Gatherer Gut Microbiota. *Current Biology,* 25(13):1682-1693.

Rifkin, R.F., Vikram, S., Ramond, J.B., Rey-Iglesia, A., Brand, T.B., Porraz, G., Val, A., Hall, G., Woodborne, S., Le Bailly, M., Potgieter, M., Underdown, S.J., Koopman, J.E., Cowan, D.A., Van De Peer, Y., Willerslev, E. & Hansen, A.J. 2020. Multi-proxy analyses of a mid-15th century Middle Iron Age Bantu-speaker palaeo-faecal specimen elucidates the configuration of the 'ancestral' sub-Saharan African intestinal microbiome. *Microbiome,* 8(1):62.

Rook, G.A., Raison, C.L. & Lowry, C.A. 2014. Microbial 'old friends', immunoregulation and socioeconomic status. *Clin Exp Immunol,* 177(1):1-12.

Rosas-Plaza, S., Hernandez-Teran, A., Navarro-Diaz, M., Escalante, A.E., Morales-Espinosa, R. & Cerritos, R. 2022. Human Gut Microbiome Across Different Lifestyles: From Hunter-Gatherers to Urban Populations. *Front Microbiol,* 13:843170.

Rstudio Team, R. 2022. RStudio: integrated development environment for R. RStudio, PBC, Boston, MA.

Russell, A.B., Wexler, A.G., Harding, B.N., Whitney, J.C., Bohn, A.J., Goo, Y.A., Tran, B.Q., Barry, N.A., Zheng, H., Peterson, S.B., Chou, S., Gonen, T., Goodlett, D.R., Goodman, A.L. &

Mougous, J.D. 2014. A type VI secretion-related pathway in Bacteroidetes mediates interbacterial antagonism. *Cell Host & Microbe,* 16(2):227-236.

Sado-Kamdem, S.L., Vannini, L. & Guerzoni, M.E. 2009. Effect of α-linolenic, capric and lauric acid on the fatty acid biosynthesis in Staphylococcus aureus. *International journal of food microbiology,* 129(3):288-294.

Sanna, S., Van Zuydam, N.R., Mahajan, A., Kurilshikov, A., Vich Vila, A., Vosa, U., Mujagic, Z., Masclee, A.a.M., Jonkers, D., Oosting, M., Joosten, L.a.B., Netea, M.G., Franke, L., Zhernakova, A., Fu, J., Wijmenga, C. & Mccarthy, M.I. 2019. Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. *Nature genetics,* 51(4):600-605.

Schnorr, S.L., Candela, M., Rampelli, S., Centanni, M., Consolandi, C., Basaglia, G., Turroni, S., Biagi, E., Peano, C., Severgnini, M., Fiori, J., Gotti, R., De Bellis, G., Luiselli, D., Brigidi, P., Mabulla, A., Marlowe, F., Henry, A.G. & Crittenden, A.N. 2014. Gut microbiome of the Hadza hunter-gatherers. *Nature Communications,* 5:3654.

Schnorr, S.L., Sankaranarayanan, K., Lewis, C.M., Jr. & Warinner, C. 2016. Insights into human evolution from ancient and contemporary microbiome studies. *Curr Opin Genet Dev,* 41:14-26.

Segata, N. 2015. Gut Microbiome: Westernization and the Disappearance of Intestinal Diversity. *Curr Biol,* 25(14):R611-613.

Selber-Hnatiw, S., Rukundo, B., Ahmadi, M., Akoubi, H., Al-Bizri, H., Aliu, A.F., Ambeaghen, T.U., Avetisyan, L., Bahar, I., Baird, A., Begum, F., Ben Soussan, H., Blondeau-Ethier, V., Bordaries, R., Bramwell, H., Briggs, A., Bui, R., Carnevale, M., Chancharoen, M., Chevassus, T., Choi, J.H., Coulombe, K., Couvrette, F., D'abreau, S., Davies, M., Desbiens, M.P., Di Maulo, T., Di Paolo, S.A., Do Ponte, S., Dos Santos Ribeiro, P., Dubuc-Kanary, L.A., Duncan, P.K., Dupuis, F., El-Nounou, S., Eyangos, C.N., Ferguson, N.K., Flores-Chinchilla, N.R., Fotakis, T., Gado Oumarou, H.D.M., Georgiev, M., Ghiassy, S., Glibetic, N., Gregoire Bouchard, J., Hassan, T., Huseen, I., Ibuna Quilatan, M.F., Iozzo, T., Islam, S., Jaunky, D.B., Jeyasegaram, A., Johnston, M.A., Kahler, M.R., Kaler, K., Kamani, C., Karimian Rad, H., Konidis, E., Konieczny, F., Kurianowicz, S., Lamothe, P., Legros, K., Leroux, S., Li, J., Lozano Rodriguez, M.E., Luponio-Yoffe, S., Maalouf, Y., Mantha, J., Mccormick, M., Mondragon, P., Narayana, T., Neretin, E., Nguyen, T.T.T., Niu, I., Nkemazem, R.B., O'donovan, M., Oueis, M., Paquette, S., Patel, N., Pecsi, E., Peters, J., Pettorelli, A., Poirier, C., Pompa, V.R., Rajen, H., Ralph, R.O., Rosales-Vasquez, J., Rubinshtein, D., Sakr, S., Sebai, M.S., Serravalle, L., Sidibe, F., Sinnathurai, A., Soho, D., Sundarakrishnan, A., Svistkova, V., Ugbeye, T.E., Vasconcelos, M.S., Vincelli, M., Voitovich, O., Vrabel, P., Wang, L., Wasfi, M., Zha, C.Y. & Gamberi, C. 2017. Human Gut Microbiota: Toward an Ecology of Disease. *Front Microbiol,* 8:1265.

Shanahan, F., Ghosh, T.S. & O'toole, P.W. 2021. The Healthy Microbiome-What Is the Definition of a Healthy Gut Microbiome? *Gastroenterology,* 160(2):483-494.

Shareefdeen, H. & Hill, C. 2022. The gut virome in health and disease: new insights and associations. *Curr Opin Gastroenterol,* 38(6):549-554.

Shetty, S.A., Lahti, L., De Vos, W.M. & Smidt, H. 2018. microbiomeutilities: An R package for utilities to guide in-depth marker gene amplicon data analysis. *Ecophysiological insights into the human intestinal microbiota: from single strains to defined consortia.*

Shuai, M., Fu, Y., Zhong, H.-L., Gou, W., Jiang, Z., Liang, Y., Miao, Z., Xu, J.-J., Huynh, T. & Wahlqvist, M.L. 2022. Mapping the human gut mycobiome in middle-aged and elderly adults: multiomics insights and implications for host metabolic health. *Gut,* 71(9):1812-1820.

Silva, Y.P., Bernardi, A. & Frozza, R.L. 2020. The Role of Short-Chain Fatty Acids From Gut Microbiota in Gut-Brain Communication. *Frontiers in Endocrinology,* 11:25.

Smith, A.B. 2001. Ethnohistory and Archaeology of the Ju/'hoansi Bushmen. *African study monographs,* 26:15-25.

Smits, S.A., Leach, J., Sonnenburg, E.D., Gonzalez, C.G., Lichtman, J.S., Reid, G., Knight, R., Manjurano, A., Changalucha, J. & Elias, J.E. 2017. Seasonal cycling in the gut microbiome of the Hadza hunter-gatherers of Tanzania. *Science,* 357(6353):802-806.

Sofi, M.H., Gudi, R., Karumuthil-Melethil, S., Perez, N., Johnson, B.M. & Vasu, C. 2014. pH of drinking water influences the composition of gut microbiome and type 1 diabetes incidence. *Diabetes,* 63(2):632-644.

Thorburn, A.N., Macia, L. & Mackay, C.R. 2014. Diet, metabolites, and "western-lifestyle" inflammatory diseases. *Immunity,* 40(6):833-842.

Torii, T., Kanemitsu, K., Wada, T., Itoh, S., Kinugawa, K. & Hagiwara, A. 2010. Measurement of short-chain fatty acids in human faeces using high-performance liquid chromatography: specimen stability. *Annals of Clinical Biochemistry,* 47(Pt 5):447-452.

Trompette, A., Gollwitzer, E.S., Yadava, K., Sichelstiel, A.K., Sprenger, N., Ngom-Bru, C., Blanchard, C., Junt, T., Nicod, L.P., Harris, N.L. & Marsland, B.J. 2014. Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nature medicine,* 20(2):159-166.

Vangay, P., Johnson, A.J., Ward, T.L., Al-Ghalith, G.A., Shields-Cutler, R.R., Hillmann, B.M., Lucas, S.K., Beura, L.K., Thompson, E.A., Till, L.M., Batres, R., Paw, B., Pergament, S.L., Saenyakul, P., Xiong, M., Kim, A.D., Kim, G., Masopust, D., Martens, E.C., Angkurawaranon, C., Mcgready, R., Kashyap, P.C., Culhane-Pera, K.A. & Knights, D. 2018. US Immigration Westernizes the Human Gut Microbiome. *Cell,* 175(4):962-972 e910.

Vanhaecke, T., Bretin, O., Poirel, M. & Tap, J. 2022. Drinking Water Source and Intake Are Associated with Distinct Gut Microbiota Signatures in US and UK Populations. *The Journal of Nutrition,* 152(1):171-182.

Vitale, M., Giacco, R., Laiola, M., Della Pepa, G., Luongo, D., Mangione, A., Salamone, D., Vitaglione, P., Ercolini, D. & Rivellese, A.A. 2021. Acute and chronic improvement in postprandial glucose metabolism by a diet resembling the traditional Mediterranean dietary pattern: Can SCFAs play a role? *Clinical Nutrition,* 40(2):428-437.

Vitetta, L., Llewellyn, H. & Oldfield, D. 2019. Gut dysbiosis and the intestinal microbiome: Streptococcus thermophilus a key probiotic for reducing uremia. *Microorganisms,* 7(8):228.

Walter, J. & Ley, R. 2011. The human gut microbiome: ecology and recent evolutionary changes. *Annual Review of Microbiology,* 65:411-429.

Ward, M.H., Jones, R.R., Brender, J.D., De Kok, T.M., Weyer, P.J., Nolan, B.T., Villanueva, C.M. & Van Breda, S.G. 2018. Drinking water nitrate and human health: an updated review. *International journal of environmental research and public health,* 15(7):1557.

Wheeler, M.L., Limon, J.J., Bar, A.S., Leal, C.A., Gargus, M., Tang, J., Brown, J., Funari, V.A., Wang, H.L., Crother, T.R., Arditi, M., Underhill, D.M. & Iliev, I.D. 2016. Immunological Consequences of Intestinal Fungal Dysbiosis. *Cell Host & Microbe,* 19(6):865-873.

Wicham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer: New York, NY, USA.

Wiessner, P.W. 2014. Embers of society: Firelight talk among the Ju/'hoansi Bushmen. *Proceedings of the National Academy of Sciences,* 111(39):14027-14035.