MONASH University

Microbiomes and Health of the Orang Asli

Yeo Li-Fang

PgDip in Biomedical Science
BSc Medical Bioscience

A thesis submitted for the degree of Doctor of Philosophy at
Monash University in 2021
Jeffrey Cheah School of Medicine and Health Sciences

Copyright notice

Abstract

The decrease in next-generation sequencing costs increased accessibility to microbiome research. However, global microbiome studies have demonstrated that the human microbiome is more complex, dynamic, and variable than initially anticipated. Researchers are at the early stages of untangling the web of host-microbe and microbe-microbe interactions. In recent years, several large, international projects have been launched. Yet, they are mainly conducted in developed countries on mostly modern, urbanised populations. There are still many populations in Southeast Asia, especially the indigenous populations where little to almost nothing is known about their human microbiomes. Amongst them are the indigenous people of Peninsular Malaysia, the Orang Asli (OA).

This thesis aimed to address two substantial gaps in OA health and biology. Firstly this thesis explored the saliva and gut microbiomes of OA ranging from urban to semi-nomadic hunter-gatherer communities living in the rainforest using the V4 region of the 16S rRNA gene. The saliva microbiomes of semi-nomadic, hunter-gatherer Jehai and urban Temuan had higher alpha diversity, whereas the rural Temiar exhibited the lowest alpha diversity. The urban-like OA saliva microbiomes of Temuan and TemiarGM were elevated in pathogenic oral bacteria such as *Corynebacterium, Prevotella* and *Mogibacterium.* On the other hand, the alpha diversity of OA gut microbiomes revealed a transition from most rural to urban. Semi-nomadic Jehai exhibited the highest alpha diversity and urban Temuan the lowest. *Sutterella, Odoribacter, Blautia, Lachnoclostridium, Parabacteroides, Bacteroides* and Ruminococcaceae UCG-013 were found depleted in Jehai but enriched in urban Temuan gut microbiomes. There were 73 core species identified in the Jehai gut metagenome using the UHGG reference database. *Prevotella copri* was identified at high abundance. A total of 28 putative, novel species were discovered.

Secondly, the urban Temuan had poorer cardio-metabolic health while rural and semi-nomadic OA were underdoing epidemiological transition. The overall prevalence of metabolic syndrome among OA was 44.63%. The urban Temuan suffered higher prevalence of general and abdominal obesity. Interestingly, the Jehai had double the prevalence rate of dyslipidaemia and marginally higher prevalence of hypertension despite being the leanest. Type 2 diabetes and insulin resistance using a surrogate marker, TG/HDL, were generally low among the Jehai and Temiar. Framingham Risk Score was relatively low among the OA at 4.35%.

Intestinal carriage of antibiotic-resistant bacteria, VRE and ESBL, was found to be 23% and 33%, respectively, among the Jehai which was rather unexpected. Half of the Jehai (n=6) harboured <15 antibiotic resistant (ABR) genes in their gut resistome whereas the other half had 43-71 ABR genes. ABR genes found in >80% of the Jehai included *rpoB, CfxA6, dfrE* and *tetM, Q, W.* Lastly, *Helicobacter pylori* infection was the highest among the Jehai at 78% prevalence, followed by TemiarGM at 16% and Temuan at 6%.

These findings provide essential information about indigenous OA microbiomes, aspects of antibiotic resistance, novel elements and cast light on related aspects of their health. In addition to contributing positively to the UHGG database, they pave the way for more detailed microbiome investigations to be undertaken in underrepresented communities in Southeast Asia.

Declaration

This thesis is an original work of my research and contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Print Name: Yeo Li-Fang

Date: 9th July 2021

Publication during candidature (Appendix III)

YEO, L.-F., AGHAKHANIAN, F. F., TAN, J. S., GAN, H. M. & PHIPPS, M. E. 2019. Health and saliva microbiomes of a semi-urbanized indigenous tribe in Peninsular Malaysia. *F1000Research,* 8.

AGHAKHANIAN, F., WONG, C., TAN, J., YEO, L., RAMADAS, A., EDO, J., HOH, B., KHALID, B. & PHIPPS, M. 2019. Metabolic syndrome and cardiometabolic risk factors among indigenous Malaysians. *Public health,* 176**,** 106-113.

Acknowledgement

My sincerest gratitude goes to my supervisor, Professor Maude Phipps, who has taught me things above and beyond what would appear in my thesis. She offered a sense of security that she has got your back. She gave me the confidence and freedom to explore (on the premise that I have finished my work). In a competitive setting where one is under pressure to publish or perish, I realize that curiosity and the freedom to explore is a luxury that I am extremely grateful to have been afforded. Her involvement in the project and openness for discussions provided valuable insights which guided both the direction of the project and formation of my thought processes. The most valuable thing she trained me in was the ability to think, coupled with the confidence and skills to support those thoughts, no matter how unconventional.

I would also like to thank my co-supervisors, Dr Uma Palanisamy and Dr Amreeta Dhanoa. Their seemingly unshakeable faith in me has gotten me through many challenges. It is always a delight to work with them. I learnt to rely on them when I was lost and exhausted, especially during the last part of my thesis write-up.

I am greatly indebted to my brave teammates, Lee Sze Mei (for also figuring out how to merge the V3-V4 and V4 dataset), Heerman Kumar, fieldwork assistants and friends Goh Qi En, Alicia Yee, Eleena Chong and medical doctors from CSJB who happily participated and contributed to our numerous field trips. Special shoutout to Prof Khalid Kadir, Kak Syikin for without whom fieldwork would have been unsurmountable to me.

I want to thank every Orang Asli that I have met throughout this journey, who have been so kind and patient with me and for participating and trusting me. My close confidante Lee Yee Qian, Yong Foong Chen, Selina Gan, Athena Ng for watching over me; Mom and Dad for loving and supporting me throughout this journey and my siblings.

Belum Rainforest

- ♣ park rangers- Mr Ridzuan and Mr Ismail;
- ♣ Jehai friends from Kg Sg Tiang – Pak Ngah, Saadiah, Encik Eli, Ismael (for the insights into Jehai culture and unwarranted ghost stories)

Temiar, Gua Musang

- ♣ Tok Batin of Kg Bodek, Tok Damak bin Anggah and Kak Halima who gave me a ride to the Tok Batin's house

Temiar, Pos Piah

- ♣ Prof Yvonne Lim and Dr Lee Soo Ching

Temuan, Bukit Lanjan

- ♣ Encik Azman, Kak Gelang

Technical advisor

- ♣ Dr Chong Chun Wie who gave much technical assistance in microbiome analysis.

Grant

- ♣ JCMSHS – Graduate Research Merit Scholarship and supporting the sequencing of Jehai samples when uBiome ceased operations.
- ♣ Tropical Medicine and Biology (TMB) platform – for grant support of fieldwork, NGS sequencing and *H. pylori* pilot study
- ♣ Monash University Malaysia Genomics Facility (MUMGF) – Prof Qasim and lovely Ms Shu Yong

Table of Contents

List of Figures

List of Tables

Supplementary material

Thesis outline

My doctoral work and experiences over several years were, in many ways, unconventional and a departure from the usual biomedical laboratory-based project. Therefore, what I have written may be a bit different from a conventional, traditional thesis. For example, I have added a section in Chapter Two that addresses some fieldwork and sample size challenges. There is also an additional part called Researcher's Reflexivity, which is not typically included in many other 'biomedical' thesis types as there would be no need for it. However, my research journey and experiences were a mixture of field anthropology, biomedical and bioinformatics. I have mapped out the sections here to ease navigation through the various sections (Figure 0.1). To this end, I have also included QR codes throughout the thesis that can be scanned (or clicked) for relevant information.

Figure 0.1 An overview of thesis chapters

Chapter 1 Introduction

1.1 Human microbiome

Microbiomes are the collection of genes of all the microbes that inhabit a particular area, for example, human skin, gut, soil, and water. The upsurge in microbiome studies within the past five to ten years had resulted from interests in microbial determinants of human biology and health, decreasing costs and increasing accuracy of next-generation sequencing (NGS). In addition, accessible bioinformatics tools and interpretive technologies have allowed biologists, even those without an informatics background, to study microbiomes.

Microbes help carry out specific tasks that the human host cannot do independently, such as digest dietary fibre. By-products from microbe-mediated fermentation in the human gastrointestinal tract, such as short-chain fatty acids, can lead to physiological changes, which directly or indirectly affects the whole body. The commercialisation potential of microbiome manipulation to generate better outcomes and products have garnered much interest in the healthcare, cosmetics industry, and food industries. The idea of an organic therapeutic option mediated by microbes is equally fascinating to the general public as it is to researchers.

Microbiome research is a relatively young, albeit rapidly growing field. However, its concept may well be traced back to 4th century China with written records of a medical cure known as “yellow soup”. The faeces of a healthy person was made into a soup to treat patients with recurring diarrhoea. The practice was quickly banned but was rumoured to be secretly used among the poor. In modern times, the first reported clinical use of faecal transplant was by Eiseman and colleagues who treated suspected *Clostridium difficile* enterocolitis in 1958 (Eiseman et al., 1958). The term ‘microflora’ has been used in publications involving humans as early as the 1940s. For example, there was a study on infant intestinal microflora by Cataldi and Montagna (1944). Another study was published on dental plaque microflora of immune-compromised patients (Green et al., 1957). Several years later, scientists established human faecal bacteria in germ-free mice (Gibbons et al., 1964), studied human microflora of healthy people (Mata et al., 1969), the effect of different diets (Borriello et al., 1978, Drasar et al., 1986) and even made an urban-rural faecal microflora comparison (Benno et al., 1986).

This is not a comprehensive, chronological list of historical microbiome work. Nonetheless, it illustrates some similarities in concepts and research directions on microbiomes from the 20th to the 21st century. Despite being a relatively old concept, microbiome research has reached new heights in the 21st century with NGS technology that obviates the need for growing bacteria in culture. By using new methods and molecular taxonomic insights, we can study what our predecessors could not.

In *modern* microbiome research, most of the knowledge has been derived from populations in Western countries, such as North America, UK, and Europe. Most participants are Caucasians, who primarily consume ‘western diets’ and mostly live in urbanised areas. The human microbiome is largely affected by diet (David et al., 2014, Singh et al., 2017, De Angelis et al., 2020), especially dietary fibres (Zhao et al., 2018); environment (De Filippo et al., 2017, Rothschild et al., 2018), medical interventions (Jakobsson et al., 2010), and other stochastic events (Dill-McFarland et al., 2019). Researchers realised that human microbiomes are highly dynamic and varied. Microbiomes differed even among monozygotic twins who were living apart (Finnicum et al., 2019). A healthy microbiome is hard to define. Yet, most research seems to agree that a healthy person’s microbiome is rich in diversity with a high redundancy of bacterial species. High redundancy of species who occupy a similar niche and

function ensures the robustness of the host system and resist the overgrowth of pathogenic species. One species can quickly occupy the niche and take over its functionality should an event cause a decrease in number or a complete wipe-out. On the other hand, a diseased microbiome appears less diverse and often dominated by few bacterial species (low evenness) where the dysbiosis state seems to favour their growth.

1.1.1 Oral microbiome

The oral cavity is inhabited by a wide range of microbes where a unique subset of microbes colonises different surfaces within the oral cavity. The oral microbiome may include samples from the gingival sulcus, tongue, cheek, hard and soft palates, throat, saliva, and teeth (Dewhirst et al., 2010). The oral microbiome is relatively well-characterized due to the ease of sample collection (Deo and Deshmukh, 2019). Oral microbiomes have been explored from the early days to identify biomarkers for oral diseases, including lichen planus, oral cancer, psoriasis (Martina et al., 2020). However, oral microbiome studies have received less attention than the gut (Nath et al., 2021). In the year 2020, for every 100,000 citations in PubMed[1], there were approximately 500 gut microbiome citations (Figure 1.1), compared to 100 and 10 citations for oral and saliva microbiomes, respectively (Sperr, 2016).

Since the oral cavity is the gateway to the gut, studies have suggested the occurrence of oral-gut microbe cross talk (Benahmed et al., 2020) and that dysbiosis in the gut may be reflected in the oral microbiomes (Bajaj et al., 2015, Drago et al., 2019, Iwauchi et al., 2019). Shifts in the oral microbiome have been indicative of cardiovascular diseases, diabetes, and adverse pregnancy outcomes (Sampaio-Maia et al., 2016). Changes in saliva microbiomes have also been indicated in cardiovascular diseases such as rheumatic heart disease (Shi et al., 2021), atherosclerosis (Hoshiga et al., 2019), age-induced deterioration in cardiovascular health and cognitive function (Vanhatalo et al., 2021). The saliva microbiome can also be used as biomarkers for diseases such as periodontitis, caries and oropharyngeal cancer (Chattopadhyay and Panda, 2019). The development of saliva into a diagnostic tool is highly desirable because sample collection is quick, straightforward, and, most importantly, non-invasive (Martina et al., 2020). Unfortunately, there are some limitations in current technology for processing and measuring salivary elements (Chattopadhyay and Panda, 2019). We are also in the early days of understanding the saliva microbiome and the underlying mechanisms in oral health and disease.

Studies have denoted that modern, urbanized oral microbiomes were markedly less diverse than hunter-gatherers (Lassalle et al., 2018) and ancient humans (Adler et al., 2013). Calcified dental plaque on ancient human teeth from early European skeletons suggested that the transition from hunter-gatherer to farming had caused a shift in the oral microbiome to a disease-associated community (Adler et al., 2013). Unfortunately, the oral microbiomes of non-industrialised populations have mostly been left out of large cohort studies (Nath et al., 2021). Furthermore, the presence of pathogenic species in the oral cavity may not necessarily be linked to poor oral health (Lassalle et al., 2018, Fellows Yates et al., 2021). This indicates a gap in knowledge regarding the oral microbial constituents and host interaction. Increasing research among under-represented populations can improve oral health among people who suffer from a higher burden of oral diseases and provide new insights into the mechanisms that cause shifts in oral microbiomes (Nath et al., 2021).

[1] Results per 100,000 is calculated by $\frac{No.of\ search\ results\ that\ year}{Total\ no.of\ PubMed\ results\ that\ year}\ x\ 100{,}000$

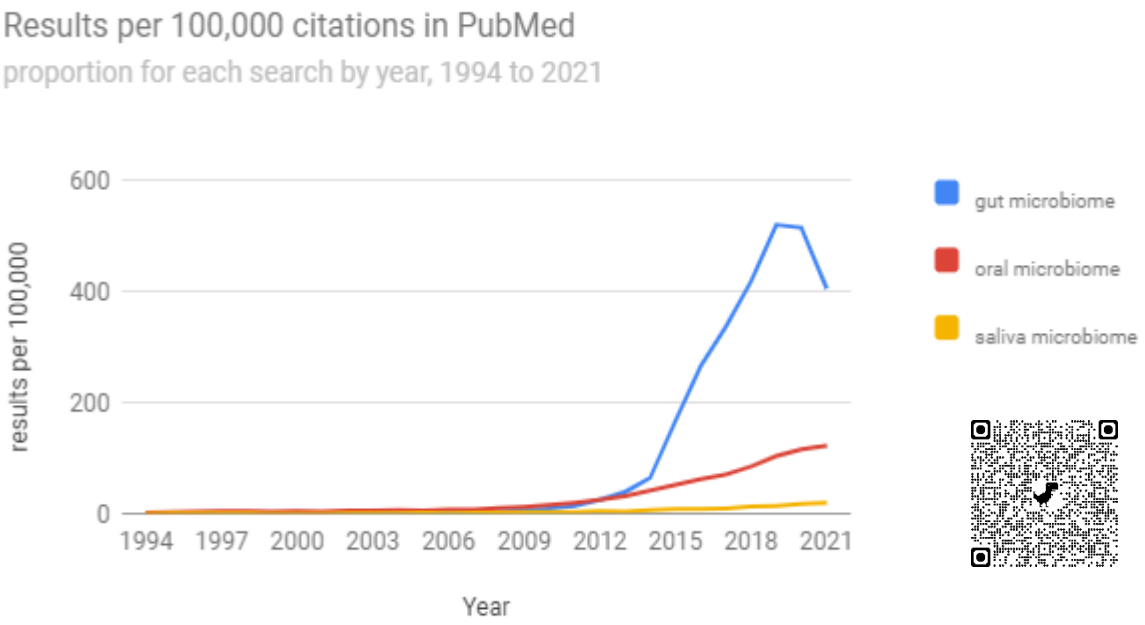


Figure 1.1 Publication trend of gut, oral and saliva microbiome studies in PubMed database. Made with PubMed by Year: http://esperr.github.io/pubmed-by-year

1.1.2 Gut microbiome

New studies and deeper sequencing data have facilitated the gradual transition of gut microbiome studies from correlational to establishing causal relationships. However, the shift is gradual and mired by the transferability of findings from existing animal models to human studies (Wade and Hall, 2019). Nonetheless, it is an exciting time to be doing microbiome research. For example, individuals with type 2 diabetes (T2D) tend to have decreased butyrate production (Jia et al., 2017). A recent multi-omics study demonstrated that host-genetic could influence the microbiome to increase butyrate production and improve insulin response (Sanna et al., 2019). In contrast, irregularity in propionate production was causally related to an increased risk of T2D (Sanna et al., 2019).

Similarly, there are studies that aim to establish causal inference on the gut microbiome, including non-alcoholic fatty liver disease (Wu et al., 2020), host genetic traits, anthropometrics and microbiome composition (Hughes et al., 2020), effect of environmental exposures on the gut microbiome (Sommer et al., 2021), and effect of *Bifidobacterium* on obesity (Pedret et al., 2019). These studies successfully reported some causal inferences and yet still err on the side of caution because these are preliminary studies, with a lot of work still needed to be done (Walter et al., 2020). For instance, epidemiological observations of associations between the microbiome and disease have been contradictory to the results investigated using causal methods (Lv et al., 2021). Secondly, there are challenges in isolating and culturing gut bacteria under current experimental conditions (Lagier et al., 2018, Lv et al., 2021).

We are in the early days of microbiome research. Despite much progress over the decade, data in this area is still lacking. Many of the research mentioned here utilises metagenomics data, human genomic data and anthropometrics obtained from large cohorts studies. This sort of data is currently only available in a few and predominantly Western cohorts. Therefore, foundational work has to be completed, especially for under-studied populations such as Southeast Asia, in the eventual development of personalised medicine via manipulation of the human microbiomes.

1.1.2.1 Shotgun metagenomics in improving strain-level diversity

Two of the most widely-used methods for microbiome studies are targeted 16S rRNA amplicon sequencing and shotgun metagenomics. Targeted 16S rRNA sequencing is a relatively cheap and quick method to provide a snapshot of the composition of a microbiome – the abundance and diversity of bacteria species present at a sampled site ("Which and how many microbes are there?"). On the other hand, shotgun metagenomics can address questions such as functional annotations, metabolic pathways that contribute to health and diseases, host-microbe, and microbe-microbe interactions ("What are the microbes doing?").

However, progress in the microbiome field may be limited by existing reference databases. Recent efforts to increase the diversity of bacteria genome reference databases have revealed a shocking number of putative *novel species[2] by reconstructing metagenome-assemble genomes (MAGs) (Parks et al., 2017, Tully et al., 2018, Almeida et al., 2019, Nayfach et al., 2019). MAGs are bacteria genomes pieced together using bioinformatics tools instead of the traditional method of first culturing and then extracting and sequencing the bacteria genome. The process of ascribing MAGs as *novel species requires sufficient sequencing depth and the most stringent quality filtering because these bacteria currently cannot be cultured in the laboratory. In other words, MAGs may only exist as genomic sequences until they are cultured in vitro (Cross et al., 2019). It would be potentially disastrous to base research and future medical interventions on 'erroneous' MAGs.

Almeida et al. (2019) and Nayfach et al. (2019) independently identified close to 2000 *novel bacteria species using MAGs from publicly available human gut metagenome datasets. Their invaluable contribution was a significant increase in the diversity of existing bacteria genome reference databases. In addition, most of these *novel species were discovered in gut microbiomes of non-western and rural populations. Once again, this finding illustrates how skewed and biased our microbiome and genomics knowledge is towards people of European descent from Western countries and rural communities with a long partnership with prominent academic institutes (Rogers et al., 2019). Asia is roughly 60% of the world's population as of 2021 (United Nations, 2019). Yet, we have a dearth of microbiome information from Asians, except for recent efforts from China (Rogers et al., 2019). There is an even more severe shortage of studies from Southeast Asian populations, particularly among indigenous populations.

Aside from demonstrating the urgent need for inclusivity of other populations, Nayfach et al. (2019) revealed that the newly identified bacteria species helped improve predictive disease models. As much as 40% of the *novel bacteria were associated with diseases including T2D, atherosclerotic cardiovascular disease, colorectal cancer and ankylosing spondylitis (Nayfach et al., 2019). Almeida et al. (2019) revealed that read classification of African and South American microbiomes, which mapped rather poorly to existing bacteria reference databases, improved by 200% following the discovery of *novel species. Similarly, reads mapping of the South American and African samples were improved by 50-87% with the addition of bacteria MAGs to the reference database (Nayfach et al., 2019, Pasolli et al., 2019).

In addition, shotgun metagenomics allows for strain-level analysis, another limitation of 16S rRNA studies. For example, *Fusobacterium nucleatum* is a symbiont and opportunist of the human mouth

[2] Putative bacterial species that have been identified through metagenomics and have yet to be characterized through culture will henceforth be denoted as *novel.

and gut that has been linked to colorectal cancer (Brennan and Garrett, 2019). *F. nucleatum* has also been implicated in periodontal health and disease by forming plaque biofilms with other oral microbiota (Brennan and Garrett, 2019). Moreover, *F. nucleatum* seems ubiquitous in the placenta of healthy, full-term pregnancies. Yet, it has also been found to induce premature births in mice (Han et al., 2004). *F. nucleatum* is just one example of the vast amount of knowledge to be discovered about specific bacteria and their interactions with other microbiota and host cells. It is essential to conduct more strain-level analysis. Significant genetic and functional differences can exist even between closely related strains, where their role under different environments and in relation to other commensals are largely unknown (Yan et al., 2020). Therefore, researchers have to be cautious about making overly simplistic claims, especially when presenting microbiome data from 16S rRNA sequencing, which only allows taxonomic resolution to the genus level.

1.1.2.2 Importance of 16S rRNA studies and inclusivity of indigenous populations

Nonetheless, many research questions have their foundation built using targeted 16S rRNA amplicon data. Studies using 16S rRNA data have revealed associations between perturbation of the human gut microbiome and diseases such as inflammatory bowel disease (IBD) (Halfvarson et al., 2017), obesity (Ley et al., 2006, Turnbaugh et al., 2009), allergies (Maslowski and Mackay, 2010), type-2 diabetes (Qin et al., 2012) and even psychiatric disorders like schizophrenia (Dinan et al., 2014), depression (Jiang et al., 2015) as well as sleeping disorders (Durgan et al., 2016).

In the absence of any previous knowledge about a cohort, the large amount of short-read data afforded by targeted 16S rRNA sequencing provides a quick snapshot of the microbiome compositions. 16S rRNA studies can also serve as a pilot before expanding to more in-depth studies that could incorporate different omics methods.

Advancement in technology has made available a variety of tools. Targeted 16S rRNA is used for preliminary studies, shotgun metagenomics for functional and strain-level analysis, and multi-omics method for eventual causal inferences. However, researchers are still at the tip of the iceberg that is human microbiomes in health and disease. Research involving minority indigenous populations poses a somewhat different challenge than urban populations (Malhi and Bader, 2015, Handsley-Davis et al., 2021). Perhaps due to inaccessibility to the settlements, potential communication barrier, extra time and effort required to ensure appropriate involvement of all stakeholders, cultural sensitivity training for investigators, and more resources needed to recruit a sufficient number of participants, may have resulted in indigenous populations often being left out of major cohort studies (Nath et al., 2021). Yet, indigenous people with their distinctive genetic makeup, unique lifestyle and culture may hold the key to unlocking intricacies of the human microbiome (Lewis Jr et al., 2012).

This is where opportunities lie for researchers to contribute to the global understanding of the human microbiome by taking the path less trodden (literally). Below is a rudimentary map of gut microbiome studies conducted on rural and indigenous populations (Figure 1.2). Several notable human microbiome studies involving indigenous communities across the world include the native Americans (Sankaranarayanan et al., 2015), Innuits of Canada (Girard et al., 2017), the Amerindians from the Amazon jungles (Clemente et al., 2015), rural communities in Papua New Guinea (Martínez et al., 2015), several African populations (Schnorr et al., 2014, Gomez et al., 2016), indigenous people from India (Dehingia et al., 2015), Nepal and several minority tribes in China (Li et al., 2016). The ability to

quickly map out all available gut microbiome studies published on indigenous people adequately demonstrates the scarcity of studies involving minority cohorts.

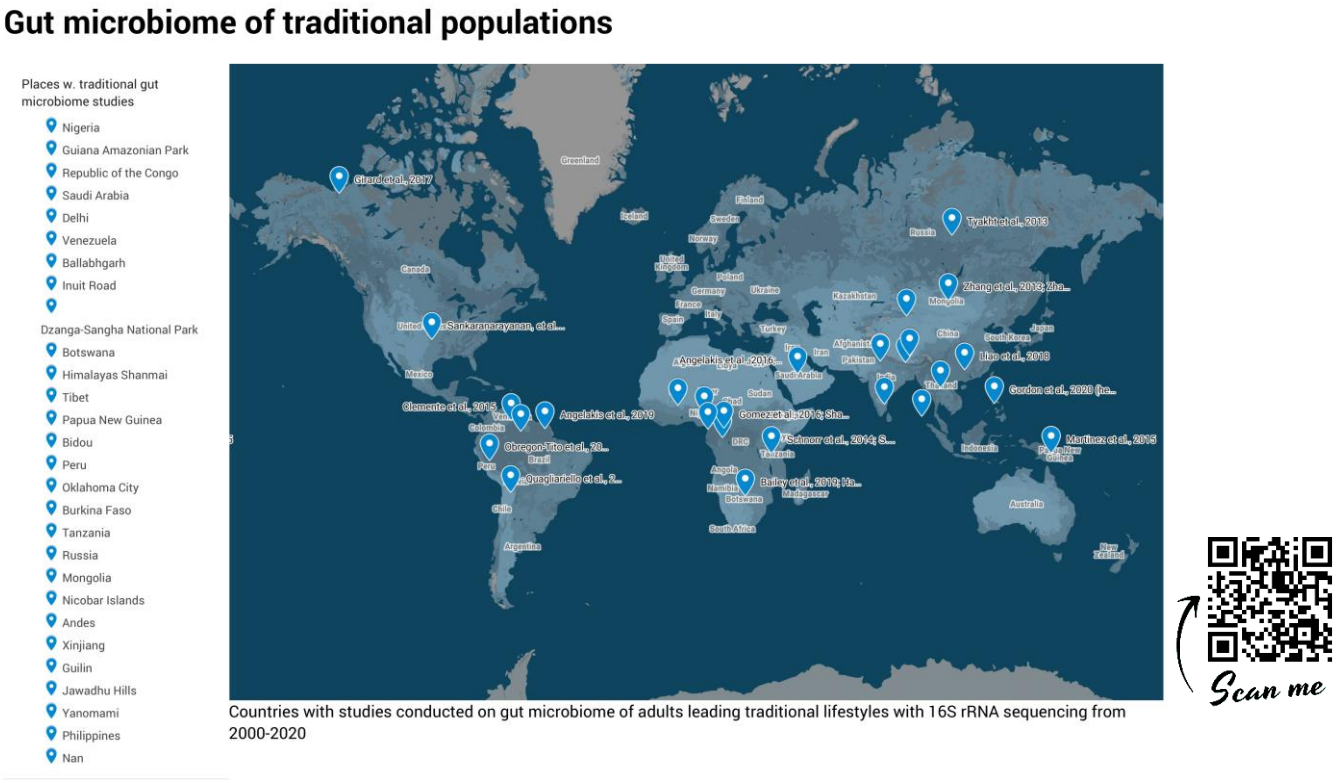


Figure 1.2 An overview of gut microbiome studies conducted on rural, indigenous populations across the world utilising pyrosequencing or targeted 16S rRNA from the year 2000-2020. (Attribution: Map data ©2021 Google, INEGI)

1.2 The Orang Asli of Malaysia

Malaysia is situated at the heart of Southeast Asia and is home to 32.7 million people. The Malay Peninsular played a vital role in the course of human migration and evolution. Some of the first humans who left Africa approximately 70,000 years ago headed eastwards and arrived on the Malay Peninsular, then part of a continental shelf known as Sundaland. Human genomic and anthropological evidence suggests that small groups of them settled on the Malay Peninsular permanently (Hill et al., 2006). These early settlers became the ancestors of the Orang Asli (OA) of Malaysia. The OAs are indigenous to Malaysia but remain relatively unknown to most Peninsular Malaysians of Malay, Chinese, and Indian descent.

The OA comprise various tribes or groups scattered throughout the Peninsular. They consist of approximately 0.7% of Malaysia's total population (IWGIA, 2020). There are three main groups, Proto-Malay, Senoi and Negritos; Each group is further divided into six sub-groups or tribes based on physical appearance, language, and culture. Each sub-group speak their language under the umbrella term "Aslian", which is of either Austroasiatic (mainland) or Austronesian (islands) origins.

OA are diverse. Those living near large cities have been resettled and are fully urbanised under government development schemes and partnerships. Some live in rural resettlement villages, usually next to rubber estates and palm oil plantations, where they sometimes obtain short-term employment. Small populations, usually the Negritos, still lead semi-nomadic lifestyles in the interiors of Malaysia's

rainforests, hunting, gathering, and trading forest produces in exchange for daily necessities. This thesis focuses on three sub-groups, namely Temuan, Temiar and Jehai, each representing different degrees of urbanization. The following descriptions of the OA are gathered through observation during fieldwork, interviews and published work.

1.2.1 Temuan

The Temuan, a sub-group of Proto-Malays, have mostly been resettled into urban areas in the central region of Peninsular Malaysia. Temuan are Austronesian speakers whose language closely resembles Malay and neighbouring Indonesian islands (Eberhard, 2021). Most Temuan were traditionally rice planters and lived close to Malay villages. They planted cash crops and collected forest produce like rattan, bamboo which were traded for rice, flour, fishing nets and other necessities.

1.2.2 Temiar

The Temiar, a sub-group of Senoi, are one of the largest OA populations at approximately 20,000 people. Temiar language is part of the Austroasiatic language family that closely resembles the Mon-Khmer languages spoken throughout mainland Southeast Asia. Temiar communities are mostly scattered around the central-north region of the Peninsular. Some communities have been resettled nearer to townships, such as the Temiar from Gua Musang, Kelantan. In contrast, others live in more rural areas, such as communities in Pos Piah, Perak who live in resettlement villages bordering the rainforest and palm oil plantations. Traditionally, Temiar practised swiddening (slash-and-burn farming) while those who lived by the fringe of rainforests also subsisted on fishing, occasional hunting, and trading forest produce.

1.2.3 Jehai

The Jehai, a sub-group of the Negrito, comprise the smallest OA communities. Actual numbers are difficult to estimate as the majority are still semi-nomadic (IWGIA, 2020). They mainly live in the north in the Royal Belum Rainforest, Perak (near the Temenggor Lake), and some parts of Kelantan. Physically, they have dark skin and tight, afro-like curls and short stature. Hence the name Negrito which means “Little African”. The Jehai language is part of the Austroasiatic language family. An interesting feature of their language is their extensive vocabulary describing odour (Majid and Burenhult, 2014). The Jehai are by large still hunter-gatherers who specialise in poison darts and blowpipe hunting. Nowadays, with most of the Jehai living more sedentary lifestyles, meat obtained through hunting appears to be mainly for self-consumption rather than trade (Loke et al., 2020). Younger Jehai may prefer to seek short-term employment in nearby towns while the state park has employed a few as local guides and park rangers.

1.2.4 An epidemiological transition

Epidemiological transition is defined as the shift in patterns of mortality, life expectancy, fertility, population age distribution and causes of death (McKeown, 2009). Now is a critical period for the OA who are undergoing epidemiological transition whereby the disease burden among the OA have gradually shifted from communicable diseases such as malaria, tuberculosis and leprosy to non-communicable diseases such as cardio-vascular diseases, type 2 diabetes, and obesity (Phipps et al., 2015; Wong et al., 2019b).

Urbanisation and the lack of well-planned, environmentally sensitive development projects are rapidly encroaching on the OA communities. As a result, OA lifestyles, health and well-being are being affected, bringing about many adverse effects. Non communicable diseases are affecting even the most remote communities (Phipps et al., 2015).

Deteriorating health among the Orang Asli

Other than the occasional dispute over land rights that makes it to the local newspapers, the general public knows little about Orang Asli. They remain scarcely mentioned in national history textbooks to this day. Orang Asli have rich cultures and traditions deeply rooted in the natural world and phenomenon, each unique to the 18 sub-groups. Their lives and well-being have been very much connected to the land over the millennia. However, in recent times, substantial realms of forest and thousands of biological species have been cleared for single species, monoculture palm oil plantations, and urban development projects (Shaq Koyok., 2020). As contemporary Temuan artist Shaq Koyok commented[3] on the de-gazettement of the Kuala Langat North Forest Reserve (KLFNR), where his people have been custodians of the land for many generations. "You cannot just recreate a peat-swamp forest…Lost is lost forever."

Scan me
Shaq Koyok's podcast

Deforestation is linked to deteriorations in the health and well-being of people whose livelihood, history and culture are so deeply entwined with the land they live on (Ellwanger et al., 2020), as the OA are. The OA were reportedly at the highest risk for cardiometabolic diseases compared to a cohort comprising Malaysia's major ethnic groups (Pell et al., 2016). The misconception about OA being "healthier" because they lived nearer to forests and thus have higher energy expenditure needs to be corrected. This theory may have held currency in the past nutrition-wise (Bolton, 1972), especially when their lands were left intact and un-appropriated (Kuchikura, 1988, Devaraj, 1999). However, many OA have now adopted more sedentary lifestyles due to resettlement schemes, and subsequent loss of their customary lands have made way for urban development projects (Nicholas, 2000, Endicott, 2012). Easy access to fast food that is relatively cheaper, high in calories and refined sugar but low in dietary fibre and other nutrients put OA health in dire need to be addressed urgently. A simple visit to most OA villages reveals the most obvious effect of urbanisation, notably the high prevalence of obesity in the community. The notion of lean OAs hunting with blowpipes and snares represents a tiny percentage of the OA (Ching and Leong, 2011, Loke et al., 2020, Bartholomew et al., 2021).

Gender disparity in Orang Asli health

There have been strong associations where gender inequality may directly or indirectly impact health and well-being, especially in lower-income countries (Wells et al., 2012). There is some evidence of gender differences among OA in cardio-metabolic risk factors. One most prominent example is the epidemiology of obesity that is being increasingly reported among many OA groups, where women appear to be more at risk of being overweight/obese. All the urban OA women who participated in the study were reportedly overweight or obese (Mohamud and Suraiami, 2010). This phenomenon was also reported in a rural Temiar community where almost 70% of the women suffered from general and abdominal obesity (Yeo et al., 2019). Poor maternal health (Lim and Chee, 1998, Khor and Zalilah, 2008), low food security (Nurfahilin and Norhasmah, 2015, Pei et al., 2018, Law et al., 2018) and malnutrition (Saibul et al., 2009) are among the few factors that seemingly impact OA women and children more severely. On top of that, the burden of infectious diseases affecting OA such as malaria,

[3] Ctrl + click the QR code for Shaq Koyok's podcast episode: "A Future Rooted in Our Forests".

tuberculosis, sexually transmitted diseases and helminth infections (Lim et al., 2009) puts OA women at overall poorer health.

The gender disparity among OA women was highlighted by Adela Baer and many other scholars who worked with various OA communities from the 1960s to the early 2000s (Baer, 1999, Baer, 2006). Their field observations suggested that women's subordination to men was introduced to the OA communities from increased contact with patriarchal societies during the British occupation and other major ethnic groups (Nicholas and Baer, 2007). As the OA lost more of their customary land, women's role in OA society as healers, midwives, shamans, and food gathering shrank (Endicott, 2012). They became more dependent on men for food and survival (Baer, 2006, Endicott, 2012). Few OA literature distinguishes health reports by gender. However, from those that do, gender inequality, as indicated by earlier researchers, was evident and may well have taken a toll on women's health and well-being.

1.2.5 Emergence of antibiotic-resistant bacteria and *Helicobacter pylori*

Government development schemes aimed at increasing OA income and food stability have failed to consider OA culture and communicate with them to understand their needs. Inequality, low literacy rates and exploitations have left OA groups vulnerable and at a higher risk for a myriad of infectious diseases. Among the multitude of infectious diseases plaguing OA communities include Lyme's disease (Khor et al., 2019), tuberculosis (Wong et al., 2019a), helminthic infections (Lim et al., 2009, Anuar et al., 2014) among others. This phenomenon is not unique to OA. Indigenous populations worldwide are at higher risk for emerging infectious diseases, including zoonotic diseases, infection from antibiotic-resistant organisms, and *Helicobacter pylori* infection (Butler, 2001). This thesis will focus on two aspects, namely antibiotic-resistant bacteria and *H. pylori* infections.

Antibiotic-resistant (ABR) bacteria

Antibiotic resistance is one of the biggest threats to global health and food security. ABR bacteria can be found in healthy people, farm animals and food along the food production chain to environmental samples such as ponds, sewage and soil (Boonyasiri et al., 2014). ABR bacteria hinder disease treatment, leading to extended hospital stays, increased medical costs and higher mortality rates (Davies and Davies, 2010). Farm animals are also indiscriminately fed antibiotics meant for human consumption, contributing to the rise of ABR bacteria (Marquardt and Li, 2018), making it harder to prevent and control multidrug-resistant strains.

There is a fair amount of information regarding ABR in the food chain (Alhaj et al., 2007, Letchumanan et al., 2015) and some in health care facilities in Malaysia (Hanafiah et al., 2019, Rashid et al., 2019). In addition, ABR is pervasive and increasingly seen in the community setting (Bezabih et al., 2020). However, information on ABR prevalence in Malaysia, such as ESBL (Extended-Spectrum Beta-Lactamase) and VRE (Vancomycin Resistant Enterococci) in the community setting, is somewhat limited and usually done independently by large research institutes and research groups. Information of ABR spread amongst the indigenous population is even more scarce.

Extended-spectrum Beta-Lactamase (ESBL) Enterobacteriaceae

Beta-lactam antibiotics such as penicillin, cephalosporin, monobactams and carbapenems have beta-lactam rings found in their chemical structures. An improved class of antibiotics called extended-spectrum beta-lactams was developed to overcome beta-lactam resistance. Soon after, Germany reported a clinical isolate of *Klebsiella ozaenae* with "transferable resistance…by plasmid pBP60" to broad-spectrum cephalosporins, an extended-spectrum beta-lactam antibiotics (Kliebe et al., 1985).

One of the earliest reports of ABR in Malaysia was data collected in 1991-1992 from six hospitals which showed 60% *E.coli*, 93.2% of *Klebsiella spp*. and 82% *Enterobacter spp*. resistant to ampicillin (Cheong et al., 1994). Biedenbach et al. (1999) reported the emergence and uprise surge of ESBL producers, especially *Klebsiella spp.* and *E.coli* in Malaysia and Singapore. Following that, Ariffin et al. (2000) reported that from 1996-1997, more than half the clinical isolates of *Klebsiella pneumoniae* in paediatric patients were resistant to ceftazidime. A committee was thus set up in Malaysia in the mid-1990s to monitor and manage the surge of ESBL-Enterobacteriaceae (ESBL-E) in hospitals (Ministry of Health Malaysia, 2001).

The worldwide prevalence of ESBL-E in the community setting among healthy individuals has shown a steady increase from 1978-2020 (Melzer and Petersen, 2007, Woerther et al., 2013, Bezabih et al., 2020). Beta-lactam antibiotics, namely penicillin and cephalosporins, were among the most prescribed antibiotics in primary-care clinics in Malaysia, accounting for 30.7% and 23.6%, respectively (Ab Rahman et al., 2016). A study further indicated that antibiotic prescription in Malaysia was higher among private clinics (30.8%) compared to public clinics (6.8%) and that upper respiratory tract infections accounted for 49.2% of antibiotic prescriptions (Ab Rahman et al., 2016). Inappropriate prescription of antibiotics and lack of guideline compliance, especially among private clinics, have been persistent in Malaysia for many years (Lim et al., 1994, Teng et al., 2011, Ab Rahman et al., 2016). Such action may further exacerbate the spread of ABR strains. Sadly, Malaysia is among many other Southeast Asian countries that lack proper management of antibiotics dispensing practices, especially in private clinics (Holloway et al., 2017, Nepal and Bhatta, 2018). The annual progression of human intestinal carriage of ESBL-E was highest in Southeast Asia from 2002-2011 (Woerther et al., 2013) and 2003-2018 (Bezabih et al., 2020). In contrast, the lowest carriage rate was consistently reported in Europe.

In the healthcare centres, ESBLs are commonly associated with urinary tract infection, ventilator-associated pneumonia, and bacteraemia. The problem posed by high human intestinal carriage of ESBL-E is that carriers are usually asymptomatic, whereby these asymptomatic carriers can transmit these bacteria via human-to-human, food chain, and non-environmental sources (Mughini-Gras et al., 2019). Therefore, when asymptomatic carriers of ESBL-E are infected by common infections, for example, urinary tract infections, they cannot be effectively treated with oral-administered antibiotics. Instead, they may require hospitalization for administration of intravenous antibiotics such as carbapenems, last-resort antibiotics to tackle ESBL-E. This hospitalization increases their exposure and risk of being infected by other nosocomial infections, including other ABR bacteria (Pana and Zaoutis, 2018). Therefore, it is crucial to monitor the community spread of ESBL, especially among the rural, indigenous communities, who besides being at a higher risk for infectious diseases (Butler, 2001), have limited access to healthcare facilities.

Vancomycin Resistant Enterococci (VRE)

Enterococci are commensals of human and animal gut microbiomes, of which *Enterococcus faecalis* and *E. faecium* account for most human enterococcal infections (Ahmed, 2018). Vancomycin is a glycopeptide antibody used to treat bloodstream infections, MRSA infections, *Clostridium difficile* associated diarrhoea and enterocolitis. Vancomycin is one of the last few drugs still effective against MRSA. Enterococci from clinical outbreaks are different from commensal enterococci. Clinical variants usually carry the *vanA* gene cluster that encodes for vancomycin resistance (Hammerum, 2012). The rise of vancomycin-resistant enterococci leaves us with dangerously limited treatment options.

Therefore, most hospitals closely monitor vancomycin administration to ensure that other alternatives have been considered prior.

Avoparcin and vancomycin are both glycopeptide antibiotics. Avoparcin was allowed in animal feed as a growth promoter since the 1950s. Meanwhile, vancomycin was mainly reserved for humans, especially those diagnosed with unknown bacterial sepsis. VRE was first reported in hospitalized patients in London and France in the 1980s (Uttley et al., 1988, Leclercq et al., 1988) and a few other cases sparingly. However, the association between VRE and avoparcin was not established until much later, when a study in Germany showed that VRE was more frequently isolated from farms that used avoparcin (Klare et al., 1995). Avoparcin was subsequently banned in all EU countries, but VRE had already become a widespread problem by then.

The earliest report of VRE in Malaysia was in the 1990s when a patient admitted for a bone marrow transplant was found to harbour *Enterococcus faecium* with high-level resistance to vancomycin (Riley et al., 1996). VRE was also isolated in farm animals and their animal handlers (Toosa et al., 2001, Shah-Majid et al., 2004, Getachew et al., 2013, Tan et al., 2018). *vanA* appears to be the predominant genotype isolated from VRE in Malaysia (Mohamed et al., 2015, Wada et al., 2019). Yet, its presence may not necessarily encode for vancomycin resistance as reported in environmental isolates (Daniel et al., 2017). Reservoirs for VRE in Malaysia from non-hospital settings appear to primarily originate from antibiotic use in livestock (Tan et al., 2006, Daniel et al., 2015, Wada et al., 2019), environmental sources, especially rivers and sewages polluted by run-offs from farms (Daniel et al., 2017), and human-to-human transmission possibly from asymptomatic carriers (Ibrahim et al., 2013, Ngoi et al., 2021). The intestinal carriage of VRE among rural OA communities in Malaysia is currently unknown.

Antibiotic resistance is ancient and part of the microbiome.

Selective pressure due to the overuse of antibiotics in humans and animal husbandry encourages the emergence of ABR bacteria. However, antibiotic resistance is, in fact, a natural phenomenon and ubiquitous that precedes the use of the first sulfonamide antibiotics in humans in the 1930s. Some of the oldest antimicrobial-resistant genes were found in 30,000year-old sediment DNA from Beringian permafrost (D'Costa et al., 2011) and 1,000year-old mummies from the Inca Empire (Santiago-Rodriguez et al., 2018). For millions of years, bacteria have been at war with each other and other species. They have evolved to produce antimicrobial products for defence and signalling. Competing bacteria will naturally develop resistant mechanisms against these products. DNA, often within plasmids, encoding resistance towards beta-lactams, tetracycline and glycopeptide antibiotics were found in both archaeological sources (D'Costa et al., 2011, Santiago-Rodriguez et al., 2018).

ABR genes identified may not necessarily play a role in only conferring antibiotic resistance. Instead, they may also play a regulatory role in their host. For example, the *vanGCd* gene cluster is functional and often found in *C. difficile* but does not confer vancomycin resistance (Ammam et al., 2013). The original bacterial host of ABR genes cannot be readily determined because of the promiscuity of genetic elements widely spread among various microbes (van Schaik, 2015).

Resident members of the microbiome may be naturally carrying antibiotic-resistant genes, collectively known as the resistome (Penders et al., 2013). The gut microbiome is thought to be a natural reservoir for ABR genes. ABR genes were identified in healthy infants and children in the absence of selective pressure from antibiotics (Moore et al., 2013). This study suggested that ABR genes exist in the gut reservoir through other routes such as mother's breast milk and intrapartum antibiotics prophylaxis

(Pärnänen et al., 2018). The environment, such as soil, can also contribute to the human resistome (Riesenfeld et al., 2004).

ABR genes identified in the human gut, which mostly originated from anaerobic gut commensals, were evolutionary distinct from ABR genes found in major pathogens (Sommer et al., 2009). This study suggested that ABR genes from the gut reservoir may be infrequently exchanged or perhaps inaccessible by human pathogens (Sommer et al., 2009). As always, there are exceptions. For example, macrolide resistance genes *ermB, ermF* and *ermG* tetracycline resistance genes *tetM* and *tetQ* seem to spread freely among Gram-negative and Gram-positive bacteria species (Courvalin, 1994).

One study showed that ABR genes were more abundant in countries like France and Spain than the US and Japan (Forslund et al., 2013). Another study reported ABR genes to be highest in Chinese samples, followed by Danish and Spanish (Hu et al., 2013). Despite the global threat of multi-drug resistant bacteria, there have not been published reports on the prevalence of ABR among OA communities. The isolation of ABR bacteria, namely ESBL and VRE from stool samples, and the use of metagenomics to study the gut resistome of OA, that I have undertaken may well be the first. This study will shed light on the intestinal carriage of ABR among the OA who have limited exposure and access to antibiotics.

Helicobacter pylori infection

Helicobacter pylori is a gram-negative spirochete that infects more than half the world's population (Hooi et al., 2017). *H. pylori* have causal links to chronic/atrophic gastritis, duodenal ulcers, gastric mucosa-associated lymphoid tissue (MALT) lymphoma and adenocarcinoma (Israel et al., 2001). Untreated *H. pylori* infections may lead to peptic ulcers and stomach cancer. Approximately 78% of the worldwide gastric cancer cases can be attributed to chronic *H. pylori* infections (IARC, 2014). *H. pylori* infections have a high prevalence among East Asians (China, Korea, Japan, Taiwan), Latin America, Eastern Europe (IARC, 2014) and Africa (Hooi et al., 2017). Eradication of *H. pylori* is feasible with mass screening and antibiotics therapy, a practice widely adopted by most developed countries.

The prevalence of *H. pylori* colonization in a Malaysian cohort was reported to range from 24.3% to 49% (Gunaletchumy et al., 2014). The risk factors for *H. pylori* colonization included older age, Chinese and Indian ethnicities, low educational levels and poverty (Gunaletchumy et al., 2014). Ethnic Malays in northern Peninsular Malaysia reported an exceptionally low infection rate of *H. pylori (Mahendra Raj et al., 2008, Maran et al., 2013)*. A similar pattern was also observed in Yogyakarta, Indonesia (Tokudome et al., 2005). However, there is scarcity of data on the prevalence of *H. pylori* infections among OA communities in Malaysia.

One study reported the highest prevalence of *H. pylori* infections among Negritos (65.7%) (Thevakumar et al., 2016). Tap water source and relocation from forest to resettlement villages were associated with decreased *H. pylori* infection rates among OA communities in Kelantan (Rahim et al., 2010). Since *H. pylori i*nfections are usually asymptomatic, it is vital to detect and investigate the prevalence of infection for communities with limited access to healthcare.

1.3 Problem statement and rationale for this study

1.3.1 Microbiomes

Recent years have seen an increase in international human microbiome projects with large sample sizes and usually involving multi-ethnic groups. However, Southeast Asian populations have largely

been left out, especially the indigenous people. Studying the microbiome of indigenous people, some of whom still live in forested areas, and practice traditional lifestyles may offer unique insights into the human microbiome from a time before communities were urbanised and 'overly sanitised'.

There are a few microbiome studies conducted in OA, including gut microbiome of OA children (Chong et al., 2015), saliva microbiomes (Yeo et al., 2019), gastric microbiomes (Chua et al., 2019), helminth infection and gut microbiomes (Lee et al., 2014, Lee et al., 2019) and nasal microbiomes (Cleary et al., 2021). Given the diversity of OA communities across Malaysia, OA microbiome studies are evidently still lacking.

Hence, I aimed to explore the saliva and gut microbiome of four distinct Orang Asli communities using targeted 16S rRNA sequencing. These communities range from urban, rural to forest-fringe OA communities (objective 1). Given much interest in semi-nomadic, hunter-gatherer communities and limited resources, I prioritised the Jehai group for further investigations. The study delved deeper into the core microbiome and identified *novel bacterial species in the Jehai gut using shotgun metagenomics (objective 2).

1.3.2 Health

Urbanisation can have negative impacts on the cardio-metabolic health of populations. Minority indigenous people such as the OA are especially vulnerable to poorer health. Studies have shown that they are experiencing epidemiological transition (Phipps et al., 2015, Pell et al., 2016, Aghakhanian et al., 2019). We have no information on their current health situation. **Hence, I aimed to assess a few current cardio-metabolic trends and metabolic syndrome among distinct OA communities (objective 3).**

The majority of antibiotic resistance investigations and surveillance in Malaysia have been focused primarily on patients admitted to hospitals, with little data from community studies. This is despite regional and global problems of antibiotic resistance that have been on the rise. The scarcity of research on antibiotic resistance amongst the Orang Asli prompted a preliminary exploration, especially the Jehai community who live in Belum Rainforest with limited access to modern healthcare and presumably fewer antibiotics exposures. **As part of a discovery-based study, I gathered evidence on the intestinal carriage of ABR bacteria among the Jehai community by culturing ESBL and VRE, one of the more common ABR found in Malaysia (objective 4). I furthered my search by identifying the gut resistome (collection of ABR genes) in Jehai to gain an overall perspective of ABR presence using metagenomics (objective 5).**

A preliminary investigation was undertaken to gather data about *H. pylori* prevalence in the demographically different OA communities. Aside from the few studies conducted (Rahim et al., 2010, Musa et al., 2014, Thevakumar et al., 2016, Chua et al., 2019), little is known about the spread of this medically important microbe in our indigenous groups. Hence, it was important to seek evidence for the feasibility of focused research projects in OA and help design future, larger-scale studies. **Therefore, I determined the prevalence of *H.pylori* infections (objective 6).**

1.4 Aims of this study

There were some notable gaps in the knowledge on microbiomes and health of indigenous populations in Southeast Asia. Therefore, I aimed to characterise Orang Asli microbiomes, aspects of cardio-metabolic health, antibiotic-resistant bacteria, and *H.pylori* infections in distinct Orang Asli communities.

1.4.1 Objectives

(i) Characterise the saliva and gut microbiomes of OA using targeted 16S rRNA sequencing and investigate any health associations.

(ii) Identify the core microbiome and *novel bacterial species in semi-nomadic Jehai gut with shotgun metagenomics.

(iii) Assess a few cardio-metabolic health indicators and metabolic syndrome in distinct OA communities living in urban, rural, and forest-fringe areas.

(iv) Determine the intestinal carriage of clinically important antibiotic-resistant bacteria, especially ESBL and VRE, and explore the Jehai gut resistomes.

(v) Investigate the prevalence of *H. pylori* infection among the OA using stool antigen kit

1.4.2 Flow diagram of research approaches

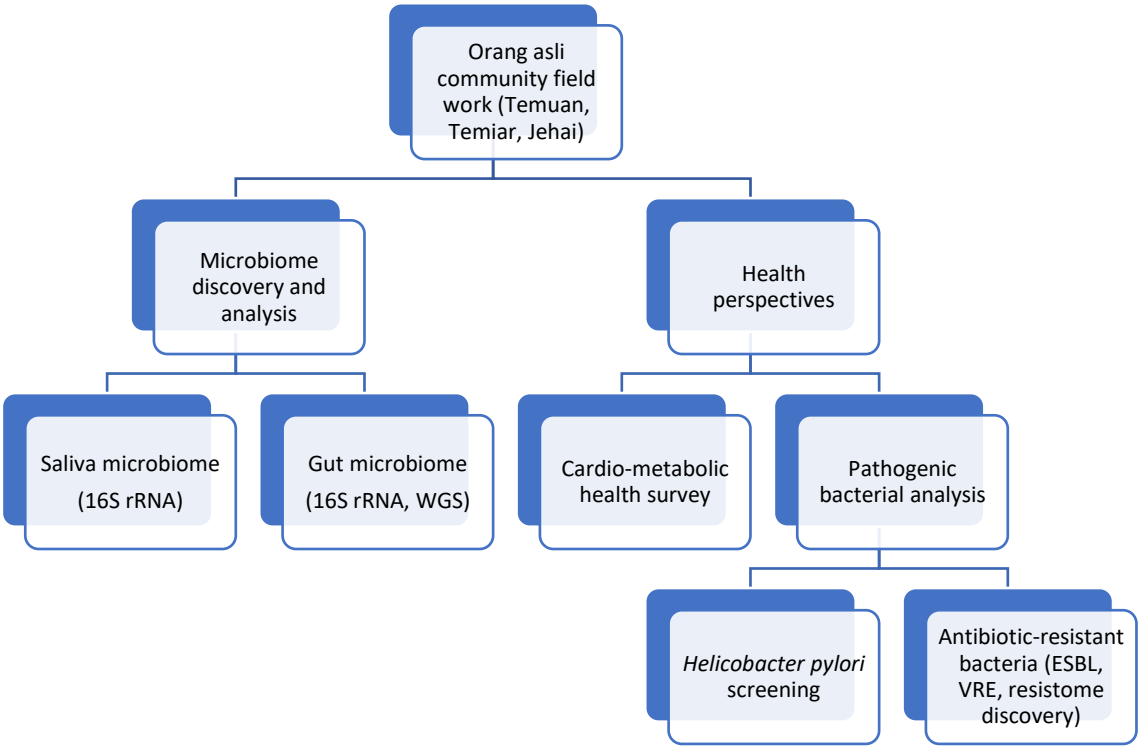


Figure 1.3 Flow diagram indicating approaches used to achieve the research objectives

Chapter 2 Methods and Approaches

2.1 Fieldwork and Community Engagement

My research with OA communities involved a fair amount of planning and engagement with various stakeholders ranging from assembling a group of field investigators, liaising with JAKOA (Department of Orang Asli Development) officers, and actually meeting the various communities. Due to my main supervisor's previous work over the past decade, there was already a certain amount of trust from local authorities, knowledge of OA communities and familiarity with elders and community engagement.

This study was approved by the Ministry of Health Malaysia (MOH) under National Medical Research Registry, MNDR ID #09-23-3913, JAKOA and Monash University Human Research Ethics Committee (MUHREC – ID #11794).

We engaged and recruited participants from three OA communities, namely (i) Temuan (urban) from Bukit Lanjan, Selangor; (ii) Temiar (suburban) from Gua Musang, Kelantan; (iii) Jehai (semi-nomadic) from Kampung Sg Tiang, Royal Belum Rainforest, Perak. Their geographical locations are depicted on the map (Figure 2.1). In addition, I included a Temiar (rural) community from Kampung Pos Piah, Perak, where part of the data has been published (Yeo et al., 2019).

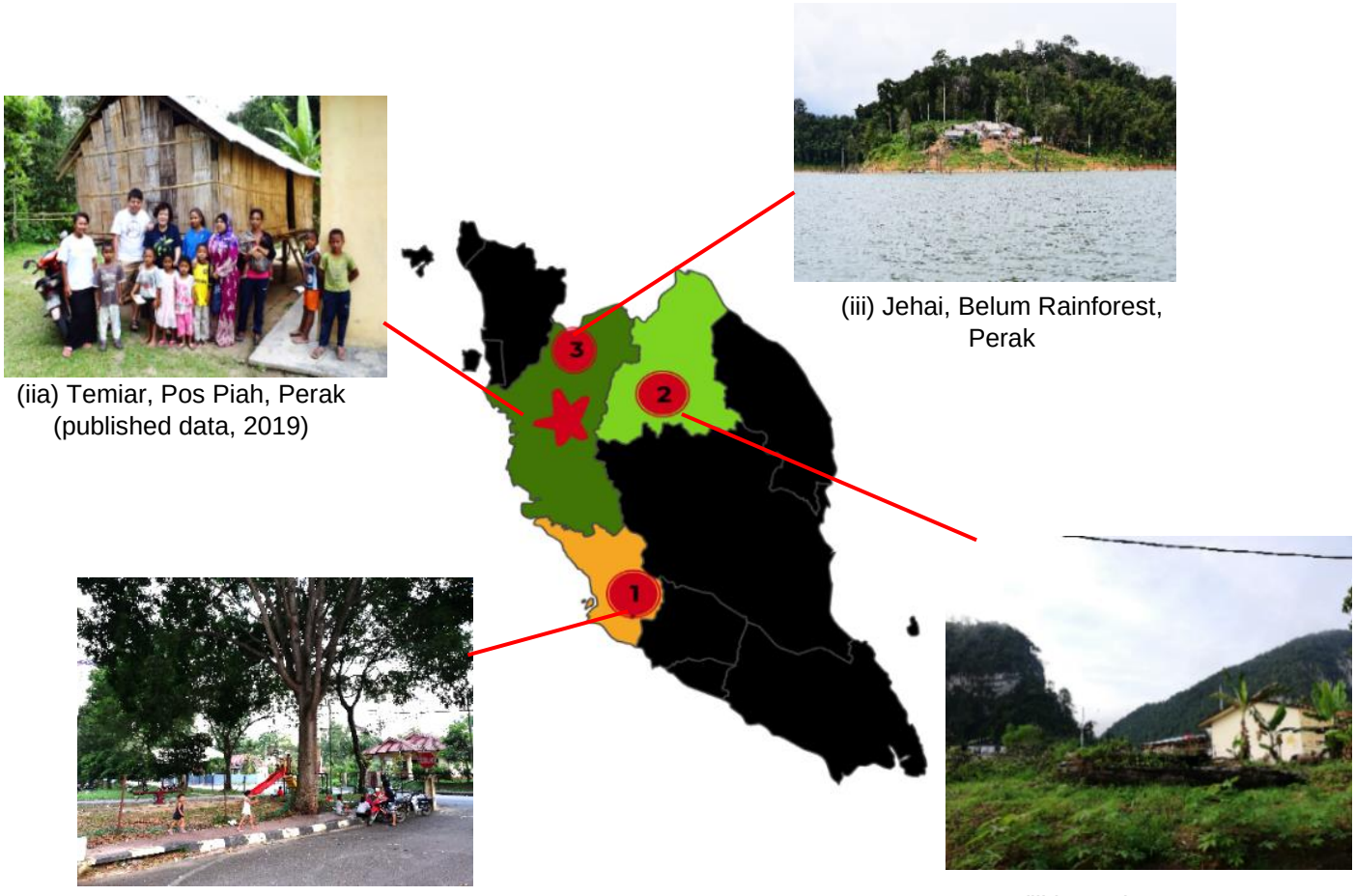


Figure 2.1 Map of Peninsular Malaysia depicting location of OA villages involved in this study.

Every field trip started with a courtesy visit accompanied by a JAKOA officer to seek permission from the headman (Tok Batin) and village elders (Figure 2.2). We spoke Malay, introduced ourselves and explained the project aims, with the aid of simple posters, to them (Figure 2.8). If time and space permitted together with consent from the elders, more villagers were invited, and an explanation was provided. Once a community had agreed to participate, we would return on dates they had suggested with field investigators, including a medical team. The underlying ethos for our OA research engagement is 'benefits sharing'. Medical attention is always welcomed, especially by villagers who live far away from government health facilities. We provided them with medical check-ups, lice medication, lunch and a gift pack consisting of daily necessities (dependent on sponsors) as small recompense in gratitude for their participation.

Figure 2.2(Left clockwise) (a) Initial self-introduction to one villager; (b) explanation of our study, how sampling would proceed, what to expect during sample collection; (c) sharing session about basic microbiology and their role (the OA) in our study.

2.1.1 Engagement with the Temuan

The Temuan are an urban OA community from Bukit Lanjan, Selangor, a fast-developing metropolis. Interviews revealed that they mostly consumed store-bought food that resembled a Malay diet, usually consisting of rice, fried fish, and some vegetables. This Temuan community was relocated twenty years ago when their ancestral land was designated for urban development. As such, they no

longer possessed lands to tend to and live off as they did in the past. Instead of living in a rural area in traditional village homes, they now lived in concrete, single-storeyed terrace houses provided by developers as compensation. The middle-aged workforce appeared to be employed in retail, factories, or as security guards. This community seemed rather disconnected from the forest, living in urban Selangor. However, some men were seen riding on motorbikes carrying fishing rods while some still knew how to hunt monkeys (personal communication), although quite rare nowadays. A few women were seen carrying plastic bags that, upon asking, turned out to be betel nuts. The practise of betel nut chewing as a social activity is still prevalent in rural OA and Indian communities (Ghani et al., 2011, Anand et al., 2014). These Temuan women were on their way to their friend's for a social gathering.

2.1.2 Contact with the Temiar

The Temiar from Pos Piah lived in a more rural place in the state of Perak. They consumed a more traditional diet consisting of plants such as tapioca ("ubi kayu"), fish and sometimes bush meat such as monkeys, river terrapin and small mammals. They supplemented their diets with store-bought rice, biscuits, and bread. Carbonated drinks and fast food were consumed infrequently as the nearest town was a 90-minute drive on a dirt track or a 4-hour walk away from the village.

On the other hand, the Temiar community from Gua Musang, Kelantan lived in a sub-urban area relatively close to town (~20-40 minutes drive away from Gua Musang town). They lived near a few Malay villages. Small stores were seen selling daily necessities such as bread, biscuit, cooking oil, sugary drinks, condensed milk, toys, and other trinkets.

2.1.3 Getting to know the Jehai

The Jehai we worked with were originally hunter-gatherers who eventually adopted a more sedentary, semi-nomadic[4] lifestyle. They lived in a remote riverbank village within the Royal Belum Rainforest State Park. JAKOA officers and park rangers have tried to persuade them to take up swidden farming to improve food security and income. However, the community we engaged with from Kampung Sungai Tiang believed that farming and plantations would attract elephants and wild boars. As a result, small parts of the forest that were initially cleared for farming looked abandoned. Attempts to convince the Jehai to plant and tap rubber failed, and those trees were quickly abandoned too. Similar attempts to introduce OA to farming and cash crops have failed among another hunter-gatherer group known as Batek. "It is not the way of our people and our ancestors.", as well as farming being boring and tedious compared to gathering for a wide variety of different food, depending on what was in season, were reasons some of the OA never picked up farming (Endicott, 2012).

The Jehai preferred to hunt (rarely now), fish and gather forest produce such as 'Gaharu', also known as 'agarwood' with highly fragrant resin used in incense that can be sold for a good price. The Malaysian rainforest apparently produces the purest quality of Gaharu (personal communication with Jehai). Gaharu develops when the heartwood of *Aquilaria malaccensis* trees become infected with a type of mould (*Phialophora parasitica).* The Jehai also collected "petai", a stinky bean that grew in tall trees (as high as 30 metres) and wild honey found in the rainforest. They travelled by motorboats to the nearest town, called Gerik, to sell forest produce and buy rice, flour, and diesel to power their

[4] Members of the community may choose to move away and form new settlements elsewhere, sometimes if there is a dispute among the people. The population in Kg Sg Tiang was much bigger prior to 2017. Park ranger Ridzuan pointed out a new settlement along the riverbanks of people who had left Kg Sg Tiang during our courtesy visit on 16th Sept 2017.

boats for the ride back. The trips could take 1 ½ hour or more by boat, depending on the water level. Some also relied on short-term employment by Chinese-owned factories looking for workers, while the state park employed a few as rangers, guides, and teachers.

Kampung Sungai Tiang was a sizable community complete with a primary school that recently had Wi-Fi connection (2017), a small prayer house, a playroom that doubled as a kindergarten and a government clinic that remained locked throughout my periodic visits in 2017-2019. The Jehai diet mainly consisted of greens obtained from the forest, rice, river fish and occasionally meat from hunting or bought from town. Elephants were frequently sighted near the village and were largely feared by the Jehai because they would trample their houses. Every morning, the Jehai women gathered by the river, exchanged news while doing their laundry, and washed the previous night's cooking utensils and bathed (Figure 2.3). The younger children were seen playing in the shallow waters. In 2019, we saw piping that directed water from the river nearer to their village, allowing people to wash their dishes and cooking utensils without having to walk too far.

Figure 2.3 We came across Jehai women and their children washing up by the river near the village.

2.1.4 Challenges, sample size and reality

The conduct of fieldwork and biomedical research amongst indigenous communities is not commonly undertaken. Such endeavours remain constant challenges for most, if not all, investigators. Several key factors and realities must be taken into account. Most of the challenges stemmed from inconveniences such as no electricity supply and inaccessibility (driving on dirt roads with unstable GPS, seasonal floods or droughts, and sometimes perilous boat rides, treks through leech-infected rainforests and so on). Fieldtrips to the Jehai community were the most challenging and the source of many adventures.

(i) Weather

We avoided fieldwork completely during the monsoon season (usually Nov – Feb). This is because the dirt roads on the east coast in Kelantan were often muddy and prone to flooding. Meanwhile, boat rides into Belum Rainforest and reaching the village jetty will be challenging. Besides, many Jehai would have left their homes to visit family elsewhere.

Unpredictable weather and lack of reliable communication signals remained a constant challenge in the interior. There, hefty rainfalls resulted in authorities draining the Temenggor Dam. This caused unusually low water levels in the Temenggor rivers and lakes beside the Jehai settlement inside Belum Rainforest. We took speed boats that had pointed bottoms that were suitable for cutting through water resistance. However, with low water levels and heavy equipment on board, our boats were constantly stuck in the muddy riverbed. Some Jehai on their way out passed us unhindered on their flat-bottomed rafts and fishing boats. Receding water levels also meant the boats had to stop further and for us to trudge through mudflats and wild grasslands to get to the village (Figure 2.4).

Figure 2.4 The water level was too low so the boat docked at quite a distance as we carried our things barefooted in the mud.

(ii) "Is anyone coming to participate?"

This is always the first question I ask during our fieldwork. The reality was that, despite constant communication with JAKOA officers, rangers, and OA liaison, in addition to visiting the villagers the day before to remind them that we were there as agreed, participant turnouts were never guaranteed. Some communities, such as the Jehai are egalitarian and lack strict social hierarchy. Therefore, no one person had much prolonged authority over the rest of the community. This translates to "You can tell me, but I may or may not do it." Usually, some villagers would be already hanging around outside

while we quickly set up our work stations. Sometimes, we were all set up and ready with not a soul in sight. That can be a nerve-racking time while I mentally replayed our conversations to look for anything I could have misunderstood or said that caused them confusion about the time and date. This usually happened with the Jehai community. I would then search for the village teacher who happens to be my age. She could provide information about her villagers' whereabouts. That being said, it is almost impossible to control how big the sample size will be. Hence our recruitments were usually ~30+ people per visit.

(iii) Timing is everything.

We tried our best to coordinate our dates with the villagers and paid participants a small sum to compensate them for their time. However, often sudden opportunities for sporadic jobs in town with lucrative pay drew the young men away from our study, resulting in us recruiting more women participants.

After reading and explaining the informed consent to the participants, some became wary and hesitant when asked to thumb-print their consent. Having heard stories of the OA being coerced into signing off their lands and other things, the person in charge or I would take time to go over the project again slowly. We would explain precisely what each sentence meant, what they were agreeing to and emphasise that they could say no and withdraw at any time. I took extreme care in making sure the field investigators who joined us, usually freshly graduated doctors and research assistants, understood that making the participants comfortable was of utmost importance. The project always took second place to the people.

Interviews took the longest, from 20-50 minutes, depending on the interviewer's experience and rapport with the participant. When working with the Jehai, we started packing up by 3pm to make it out of the village before sundown at 5pm. The rangers would otherwise be navigating the speedboats in the dark. Sitting in a boat for 1.5 to 2 hours in the dark, surrounded by dark waters and giant silhouettes of the virgin rainforest, was quite frightening. Not to mention the danger of hitting an immersed tree root in the dark (we almost did, the ranger stopped the boat at the nick of time as I watched the outline of a dead tree root loomed before my eyes).

Figure 2.5 (Top) Racing against looming rain clouds and the setting sun. The boat was too heavy with our equipment it could only carry two people. (Bottom) Immersed dead tree trunks are a trademark of Belum Rainforest, as if a reminder of a time before the forest was flooded to build the Temenggor dam, when trees had covered every inch of this place.

(vi) Far from the comforts of home

It is always best to assume there is no water and electricity supply and overpack rather than reach a fieldwork site only to find that electricity had been cut off. Therefore, battery-powered, portable equipment was preferred. Yet, this also limited the types of tests we could conduct (hence choosing the *H. pylori* stool antigen kit). Accordingly, we packed extra batteries and brought generators and diesel to power equipment such as the HbA1c machine and portable microcentrifuge.

In short, fieldwork could get tricky. It's the simple things you did not realise that brought much comfort, such as having clean feet or having clean anything. Working in Belum Rainforest with the Jehai was the most challenging because it was the furthest and the most remote community. Plus, it is one of the last virgin rainforests in Peninsular Malaysia where snakes, leeches , the last of the Malayan tiger, Malayan Sun Bear, Asian elephant, and wild boars call home. That is not to say that fieldwork with other communities was a piece of cake. They were just relatively less physically and mentally demanding.

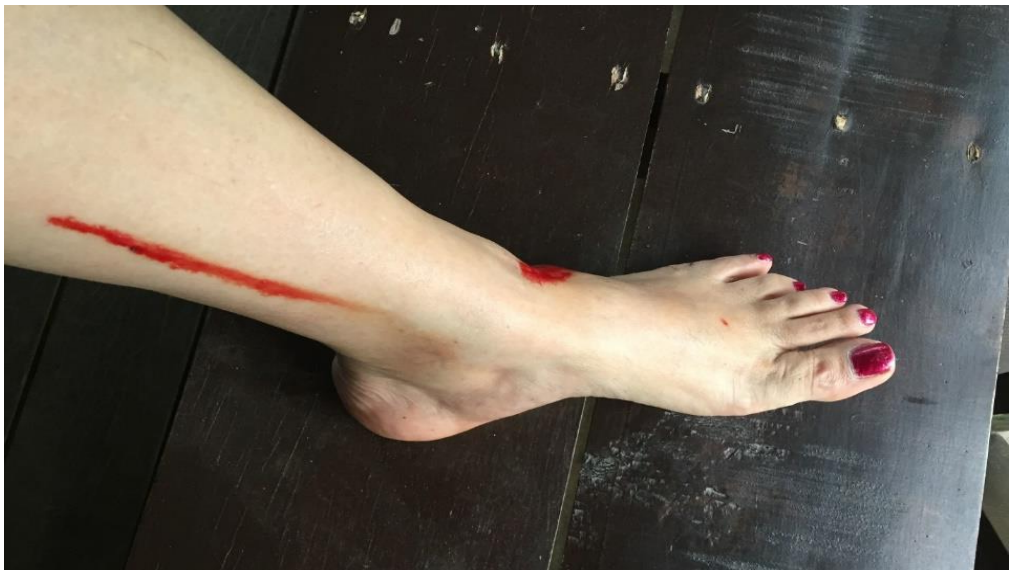


Figure 2.6 From someone with some experience with Malaysian rainforests, Belum Rainforest has got one of the longest leeches I have ever seen. They are so long they can actually stand upright in the thick undergrowth, waiting for the next animal to hitch a ride and a meal. Long and thin leeches mean hungry leeches.

2.2 Sample Collection and Processing

Participant recruitment and data collection at each village followed a typical workflow summarized in the flowchart (Figure 2.8).

2.2.1 Recruitment and interviews

Participants who were over 18 years old, with no visible health ailments and could provide informed consent were recruited. Participants who were pregnant, had a history of alcohol/drug abuse, or with chronic illness (i.e. kidney failure, cancer, heart disease) were excluded. Details of each eligible participant were recorded, and a de-identified laboratory ID assigned. The project aims, risks and the choice to withdraw at any time were explained carefully to each participant before the informed consent form was signed or thumb-printed. Participants were then interviewed on their dietary habits, food and water source, and medical history (Figure 2.7).

Figure 2.7 (Left clockwise) Interview sessions at (a) Temuan community hall on 16th Aug 2018. (b) Temiar Gua Musang community hall 6th Oct 2018. (c) Jehai village 28th July 2019.

2.2.2 Anthropometrics and Metabolic Syndrome

Anthropometric measurements collected include height, weight, waist circumference and blood pressure. We used a finger-prick method to measure HbA1c and blood lipid levels. Accutrend Cholesterol Test Strips were used to measure total cholesterol (TC), Triglyceride (TG), HDL (High-Density Lipoprotein) and LDL (Low-Density Lipoprotein). TG/HDL was used as a surrogate biomarker for insulin resistance with a cut-off point of 0.9-1.7 (Mostafa et al., 2012). TC/HDL ratio (Risk classified as >5.0 for men, >4.5 for women) and LDL/HDL ratio (risk classified as >3.5 for men, >3.0 for women) were used cardiovascular risk markers (Millán et al., 2009). Studies have shown that LDL can build up in arteries, leading to atherosclerosis, whereas HDL removes LDL from the blood to the liver (Escobar et al., 2018). Hence high LDL to HDL ratio was used to indicate cardiovascular risk. On the other hand, TC levels have been highly correlated with LDL levels, thus used as a risk marker for atherogenic risk (Quispe et al., 2020).

Lipoprotein ratios were better predictors for cardiovascular risk than single lipoprotein cut-off values and have been well utilised in Asian cohorts such as Japanese (Katakami et al., 2011) and Chinese (Zhu et al., 2015; Wen et al., 2017). BMI was classified according to WHO Asian criteria* (Anuurad et al., 2003). HbA1c and Waist circumference were classified according to WHO (WHO, 2011, Safety, 2011). Metabolic syndrome (MetS) was present if three or more of the following criteria were met (Grundy et al., 2005):

i) Waist circumference[5] >90 (men); >80 in (women) (Organization, 2011);
ii) Blood pressure > 130/85 mmHg;
iii) Triglyceride (TG) level > 8.3 mmol/L;
iv) HDL level < 2.2 mmol/L (men); <2.8 mmol/L (women);
v) Fasting blood glucose >100 mg/dL or HbA1c[6] >5.7% (Park et al., 2012).

**The cut-off points are detailed in Table 5.*

Framingham Risk Score (FRS) which is a simplified tool to estimate for cardiovascular disease for 5-year risk was calculated. Risk factors including gender, age, total cholesterol, HDL, systolic blood pressure, smoking and diabetes were considered. A recalibrated calculator validated in an Australian indigenous cohort was chosen as this might have been more suited for the OA communities than the conventional calculator commonly used by Western, mainly Caucasian populations (Hua et al., 2017).

[5] Waist circumference cut-off was modified to be appropriate for Asian populations.
[6] HbA1c was measured instead of fasting blood glucose.

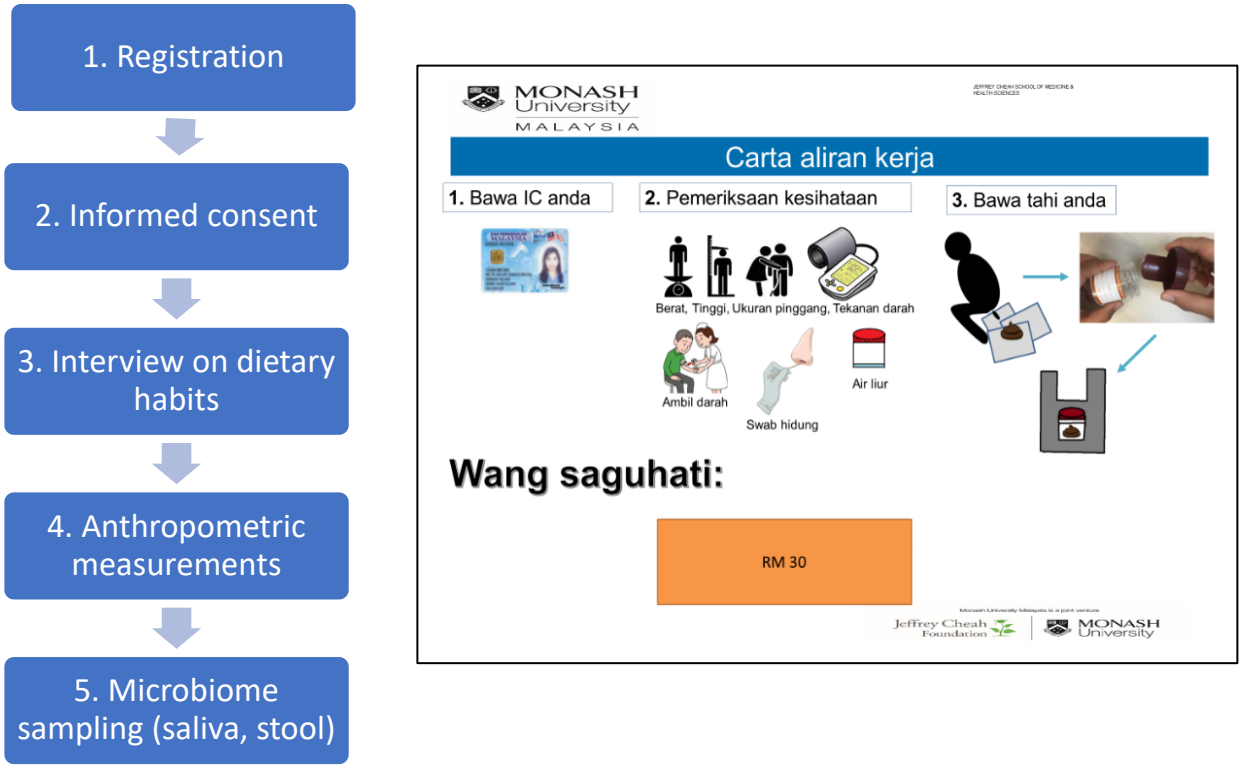


Figure 2.8(Left to right) Flowchart during data collection; A poster used to inform the community our workflow and what to expect.

2.3 Microbiome Analysis

Saliva and stool samples were collected for microbiome analysis. Targeted 16S rRNA sequencing amplified either the V3-V4 (in-house, Monash University Malaysia Genomics Facility) or V4 (uBiome Inc) hypervariable region. The flow chart (Figure 2.9) summarizes the process flow for in-house microbiome sequencing. The subsequent paragraph details DNA extraction and sequencing, combining the V3-V4 and V4 datasets and bioinformatics analysis.

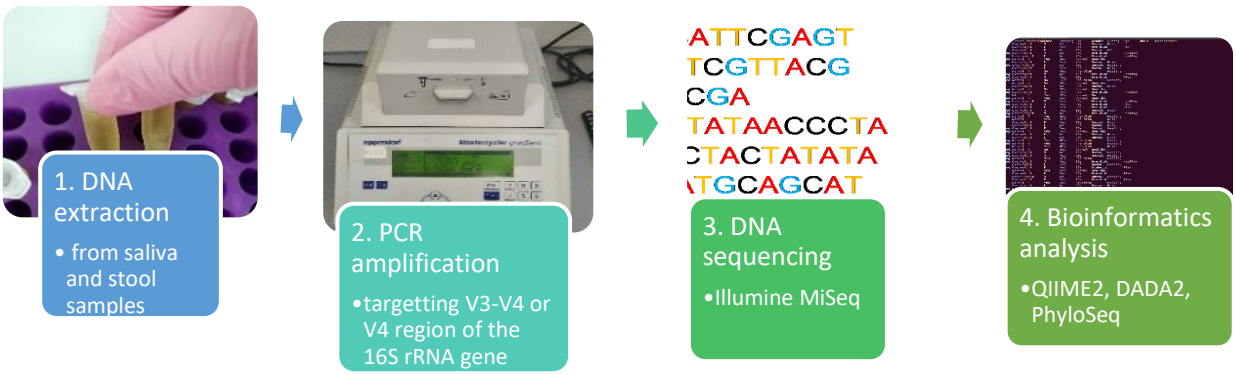


Figure 2.9 Process flow for in-house microbiome analysis

2.3.1 DNA extraction and 16S rRNA sequencing

uBiome method

Stool and saliva samples from Temuan and Temiar (Gua Musang) communities were collected using commercial uBiome kits (sponsored by Ubiome 2017-2019). Targeted 16S rRNA gene sequencing was used to identify and classify microbes. Extraction of DNA and sequencing were done by uBiome (Almonacid et al., 2017). Briefly, samples were lysed using bead-beating. DNA was extracted using guanidine thiocyanate silica column-based purification method. PCR amplification was performed using primers for the V4 variable region of the 16S rRNA gene (515F: GTGCCAGCMGCCGCGGTAA and 806R: GGACTACHVGGGTWTCTAAT) (Caporaso et al., 2011). PCR products were column-purified and size-selected using microfluidic DNA fractionation before being quantified using real-time PCR. The samples sequenced on Illumina NextSeq 500, producing 2x150bp paired-end sequences.

In-house method

Saliva samples from Temiar, Pos Piah (Yeo et al., 2019) and Jehai communities were extracted using a modified salting-out method (Quinque et al., 2006) (Appendix I). Stool samples from Temiar (Pos Piah) were extracted using the Macherey-Nagel NucleoSpin Soil kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany). Stool samples from the Jehai community were extracted according to IHMS protocol Q (Costea et al., 2017) using Qiagen QIAamp DNA mini kit. The V3-V4 hypervariable region of the 16S rRNA gene was targeted using PCR amplification and sequenced on Illumina MiSeq, 2x250bp paired-end reads. Primer sequences used were from Klindworth et al. (2013).

F: TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG

R: GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC

Microbiome control

Saliva and stool kit negative controls consisted of ultrapure water that underwent the same extraction and PCR amplification processes as the biological samples. The positive control was an even mix off 20 strains mock community (ATCC MSA-1002) was amplified at the V3-V4 region. The positive control was sequenced in 2017 and 2020 to investigate batch effect. PCR negative control was ultrapure water amplified at the V3-V4 region to determine contaminants introduced at the PCR step.

2.3.2 Combining V3-V4 and V4 dataset

Typically V3 and V4 regions of 16S rRNA may be amplified for microbiome characterization. The V4 reads were trimmed to 125bp by uBiome that initially sponsored a portion of this work by providing the kits and sequencing services. Only the forward reads were used because the reads were too short to be merged. The V3-V4 datasets were trimmed to only the V4 region using CutAdapt (Martin, 2011) before the paired-end reads were stitched together using USEARCH V11.0667 (Edgar, 2010). The merged reads were then truncated at 125bp using USEARCH to match the V4 reads. Datasets are summarized below in Table 1.

Table 1. Summary of 16S rRNA datasets used for microbiome analysis.

	Community, Year collected	16S rRNA region	No. of stool samples	No. of saliva samples	Publication status
1.	Temiar Pos Piah, 2016	V3-V4	51	68	Saliva dataset is published in Yeo *et al., (*2019) BioProject ID: PRJNA515166
2.	Jehai, 2017	V3-V4	22	47	Unpublished
3.	Temuan, 2018	V4	33	36	Unpublished
4.	Temiar Gua Musang, 2018	V4	25	29	Unpublished
5.	Jehai, 2019	V3-V4	36	36	Unpublished

2.3.3 Targeted 16S rRNA analysis

Feature table, taxonomy table and phylogenetic tree were built using QIIME2 (version 2021.4) plugins. The files were then further analysed on Phyloseq (McMurdie and Holmes, 2013) in R.

Raw reads targeting the V4 region at 125bp were first imported into QIIME2 (Estaki et al., 2020) and denoised on DADA2 (Callahan et al., 2016). A phylogenetic tree was built through open-referenced method via SATé-enabled phylogenetic placement (SEPP)(Janssen et al., 2018). Taxonomy was assigned with a trained Naïve Bayes Classifier using Silva reference database version 12_8. This version was selected to match the SEPP reference database used to build the phylogenetic tree through fragment insertion.

The following were done on R 4.0.3 (R Core Team, 2020) Alpha diversity was calculated using microbial richness, Shannon Index (evenness), which describes the species distribution in a community (Martin, 2015) and Pielou's evenness (microbial richness + evenness). Core microbiome species were identified at 90% sample prevalence at 0.1% detection threshold (relative abundance). Beta-diversity was measured using centred-log-ratio (clr) transformed Euclidean distance and plotted on PCoA and supervised CAP ordination plots. Multivariate analysis PERMANOVA (Anderson, 2017) on adonis was conducted using clr-transformed Euclidean distance with covariates BMI and age-groups. Homogeneity of sample distribution was measured using PERMDISP. Univariate differential abundance analysis was performed using ALDEX2 (Fernandes et al., 2014). Differentially abundant taxa were determined as taxa that tested statistically significant with Benjamini-Hochberg corrected p-value <0.01 and effect size >1.

2.3.4 Whole genome sequencing (WGS) and Bioinformatics Analysis of gut samples

Recognizing the limitations of using targeted 16S rRNA sequencing, absence of WGS sequencing facilities and microbiome bioinformatics pipeline in Monash University Malaysia and Covid-19 national lockdowns, we collaborated with Rob Finn at EMBL-EBI and Alexandre Almeida. They curated the Unified Human Gut Genome (UHGG) catalogue consisting of >200,000 bacteria genomes isolated from human guts (Almeida et al., 2020). They had also established a state of the art bioinformatics pipeline to identify MAGs (metagenome-assembled genomes) of *novel gut bacteria (Figure 2.10).

Twelve Jehai stool gDNA were sequenced on Illumina NovaSeq at 60 million reads for 2x150bp paired-end. *Novel bacterial species were identified using the methods developed (Almeida et al., 2019, Almeida et al., 2020). WGS analysis was performed at EBI, UK to discover the gut microbiome composition, core microbiome and identify *novel bacterial species. The analytical steps are briefly summarized in a flow chart (Figure 2.10).

Quality check

Adapters and low-quality bases were removed using trim_galore (Krueger). Sequences were aligned against the human genome (hg38) using BWA (Li and Durbin, 2009). Host contamination was identified and removed.

Metagenome assembly, binning, and taxonomic classification

Paired-end reads were merged to produce contigs by de-novo assembly using MetaSPAdes (Nurk et al., 2017). Contigs underwent genomic binning (draft genome extraction) using MetaWRAP (Uritskiy et al., 2018). Contigs 'binned' together, also called MAGs (metagenome-assembled genomes), may belong to the same species. MAGs were de-replication using dRep (Olm et al., 2017) at 95% nucleotide identity (species level) to remove strain redundancy.

The quality of MAGs, indicated by the level of genome completeness and contamination, was checked using CheckM (Parks et al., 2015). High quality or near-complete MAGs were genomes with 90% completeness and <5% contamination. Medium quality MAGs were genomes with at least 50-90% completeness and <10% contamination. MAGS with a quality score (QS) above 50 were chosen for further analysis.

(QS = completeness –(5 x contamination))

MAGs species were annotated to the Genome Taxonomy Database using GTDB-Tk (Chaumeil et al., 2019). Phylogeny tree was built with iTOL. MAGs were mapped against Unified Human Gut Genome (UHGG) catalogue (Almeida et al., 2020) to identify microbiome composition, species phylogenetic diversity and prevalence, *novel species. Additionally, the metagenome quality of Jehai MAGs was assessed against UHGG reference genomes. Read mapping was conducted using Kraken2 (Wood et al., 2019). Statistical analysis was done using python and R scripts.

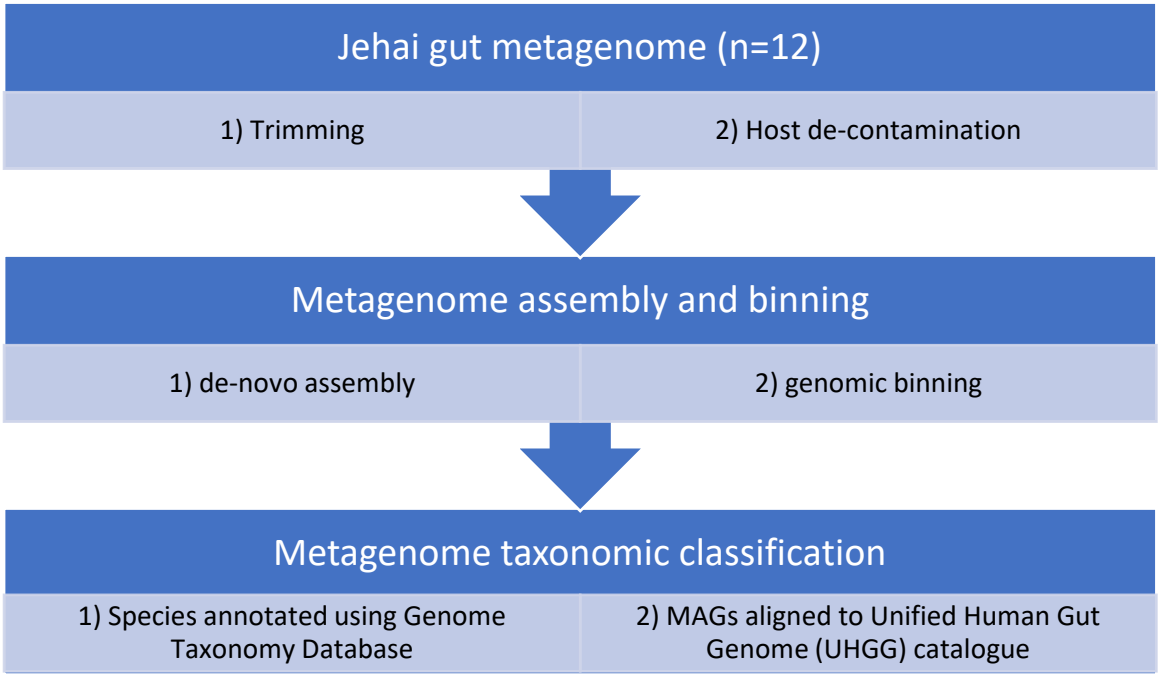


Figure 2.10 Flowchart of Jehai gut metagenome analysis

2.4 Intestinal carriage of Antibiotic-Resistant (ABR) bacteria

Intestinal carriage of ABR bacteria, specifically ESBL (extended-spectrum beta-lactamase) producers and VRE (vancomycin-resistant enterococci), was investigated only in the semi-nomadic Jehai gut samples due to research grant limitations. ABR genes were predicted in the Jehai gut metagenomes.

2.4.1 Chromogenic agar for ESBL and VRE culture

Jehai stool samples were collected and immediately stored in 15% glycerol and 0.9M saliva solution and shipped back to Monash University Malaysia laboratories in dry ice. These stool samples were stored at -80℃ while awaiting the shipment of chromogenic agar from the UK. Brilliance Chromogenic Agar from Oxoid, UK, was selected for ESBL and VRE detection, as Gouliouris et al. (2016) reported.

Samples were revived in Tryptic Soy Broth in a shaking incubator at 37℃ for up to 48 hours. Following this, stool samples were inoculated on MacConkey agar. Stool samples that failed to grow were excluded from the study because that would indicate the bacteria was no longer viable after a long freeze. Next, 200ul of Tryptic Soy Broth was pipetted onto selective, chromogenic Brilliance VRE and ESBL (Oxoid, UK) agar plates and incubated at 37℃ for 24 and 48 hours. Coloured colonies that presumptively suggested VRE and ESBL based on the manufacturer's instructions were further purified on MacConkey agar. Pure colonies were stored in 30% glycerol and Tryptic Soy Broth at -80℃ for future analysis (Gouliouris et al., 2016).

2.4.2 ABR genes prediction

ABR genes were predicted using the software ABRicate v1.0.1 (Seemann, 2020) and CARD 2020 database (Alcock et al., 2019) with shotgun metagenomes assemblies as input. The proportion of predicted gene coverage and predicted gene identity were compared to reference sequences at 90% similarity (Doyle et al., 2020).

2.5 *Helicobacter pylori* screening

Stool samples collected from 2018-2019 (Temuan, TemiarGM and Jehai 2019) were transported back to the lab in dry ice. Detection of *H. pylori* antigen in stool specimens was performed using BIONEXIA® *H. pylori* Stool Antigen test, according to manufacturer's instructions. A small amount of stool was dissolved in buffer. Four drops of the stool-buffer mix were applied to the cassette provided. After 15 minutes, two lines indicated positive infection; one line indicated negative infection (Figure 2.11).

Figure 2.11 BioNexia H.pylori test cassette: One line indicates negative; two line indicates positive infection.

Chapter 3 Results

3.1 Saliva microbiome: composition, diversity, and core constituents

Targeted 16S rRNA at the V4 region was used to detect and characterize a total of 210 OA saliva microbiomes. Six samples failed to pass the QC threshold. An overall view could be glimpsed from the taxonomy bar plot illustrating major saliva microbial composition. Next, bacteria genera shared across majority of the OA samples forming the saliva microbiome core constituents were described. Alpha diversity that characterised microbial diversity within each group and beta diversity that characterised microbial diversity between groups were investigated. Multivariate analysis PERMANOVA was conducted to study the overall difference in microbiome compositions. Finally, univariate differential abundance analysis using ALDEX2 was performed to study individual bacteria genus that had significantly different abundance between any two OA groups. Temiar and TemiarGM were analyzed as separate groups despite both identifying as Temiar. The two communities had different levels of urbanization. Temiar were from a more rural area that was a 90 mins drive from the nearest town in Perak, while TemiarGM lived approximately 30 mins drive away from Gua Musang town in Kelantan. In addition, the ordination plot (see below) suggested their microbiomes correlated more to the level of urbanization rather than ethnicity and cultural identity.

3.1.1 Saliva microbial composition

The top seven most abundant bacteria classes found in OA saliva microbiomes included Bacilli, Bacteroidia, Betaproteobacteria, Gammaproteobacteria, Actinobacteria, Negativicutes and Fusobacteriia. Sphingobacteriia (in black at the bottom of the barplot in Figure 3.1) was observed mostly in Jehai and some rural Temiar, but mostly absent in TemiarGM and Temuan.

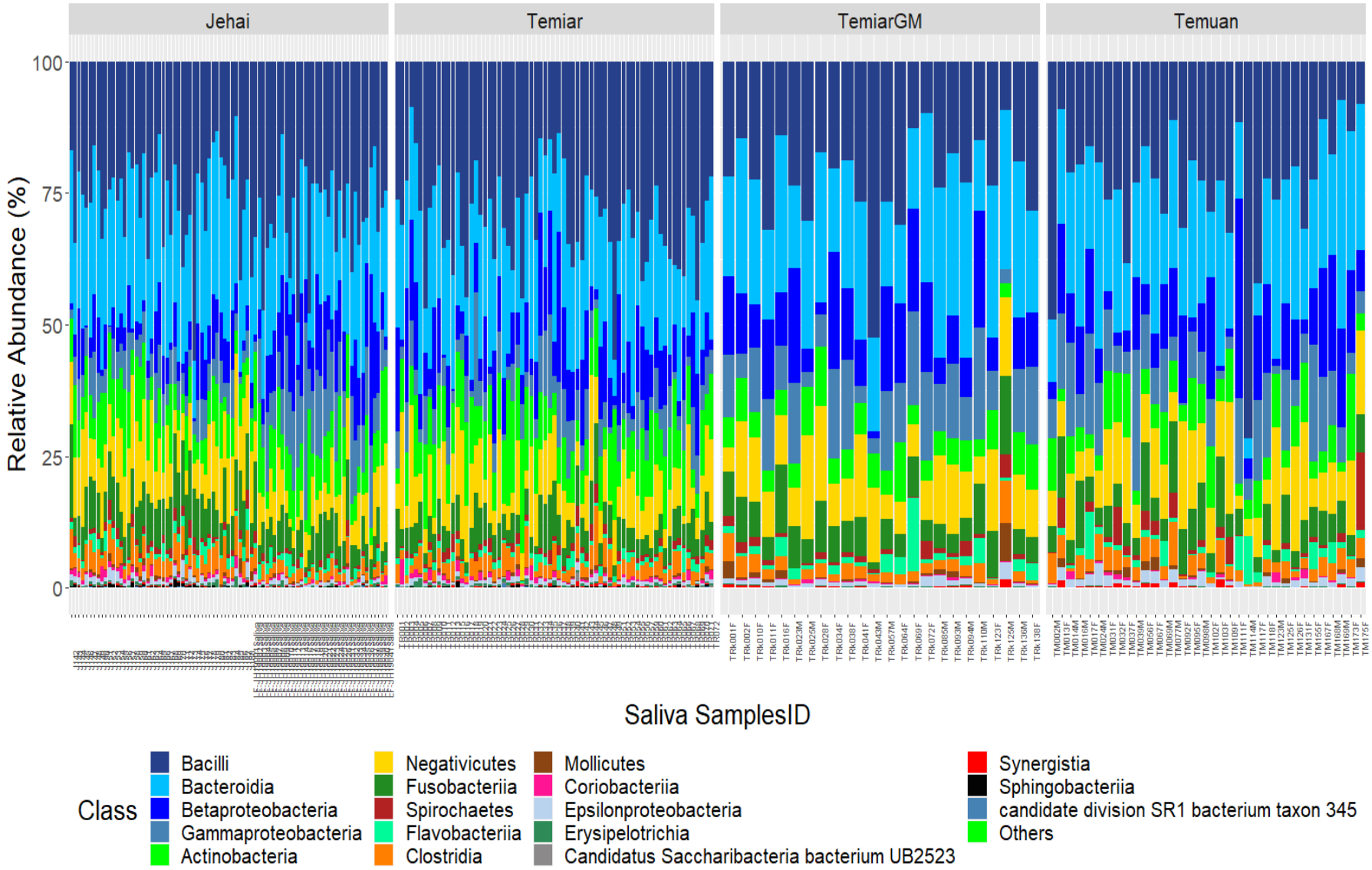


Figure 3.1 Saliva taxonomy bar plot at the Class level

3.1.2 Saliva microbial diversity: core constituents, alpha, and beta diversity

Core constituents in saliva microbiomes

Saliva core genera were investigated at various prevalence and thresholds. At 90% sample prevalence at 0.1% detection threshold (relative abundance) revealed 15 genera. Core constituents included many common oral bacteria such as *Streptococcus, Veillonella, Haemophilus, Prevotella, Gemella, Neisseria, Lactobacillales, Actinomyces* and *Fusobacterium* (Figure 3.2). The adjustment of sample prevalence to 99% revealed four genera, namely *Prevotella, Lactobacillales, Streptococcus* and *Gemella.* Notably, quite a few 'uncultured bacteria', bacteria identified through sequencing but yet to be cultured, were detected as core constituents in the saliva microbiome.

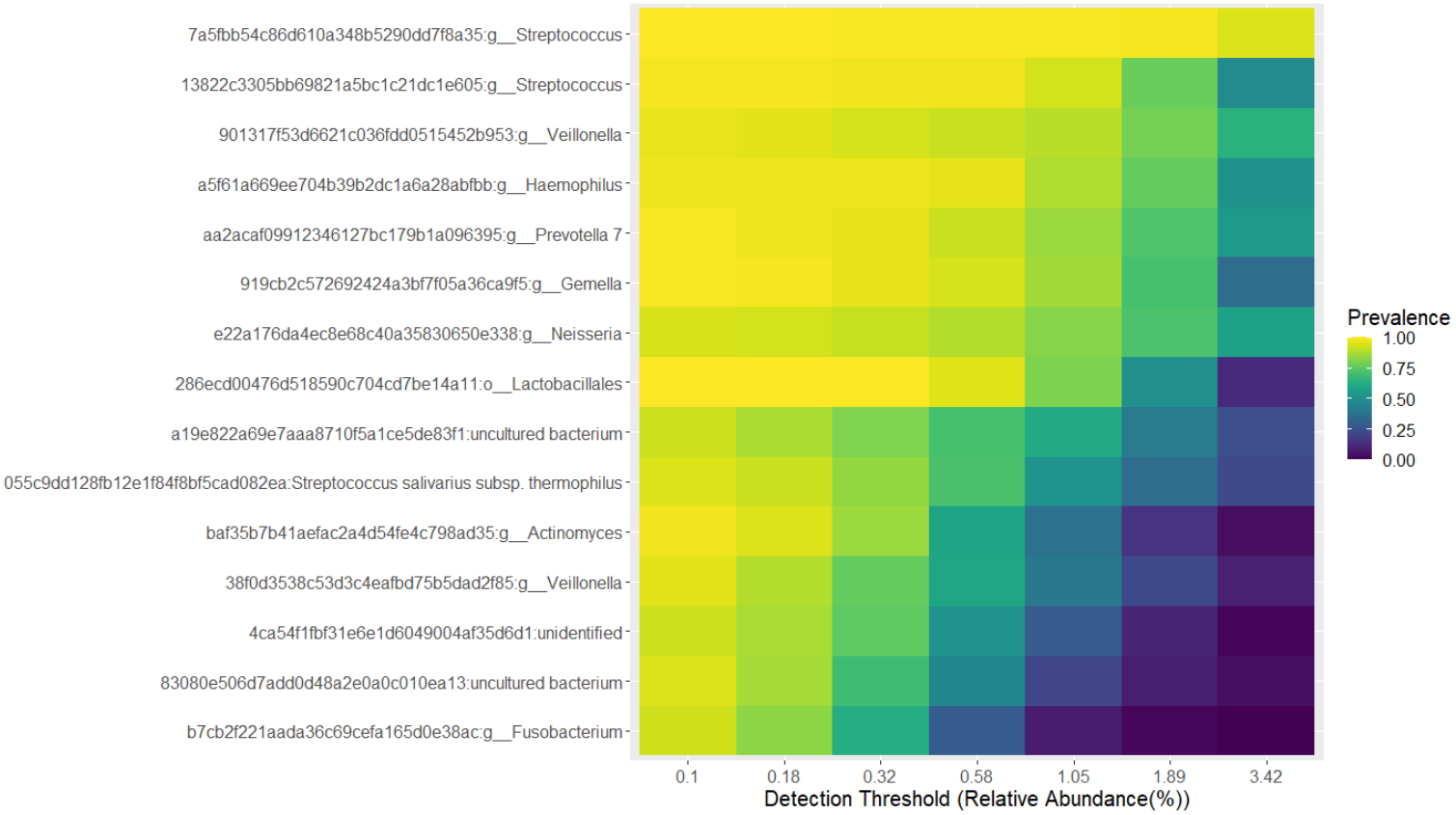


Figure 3.2 Heatmap of 15 saliva core genera detected in at least 90% of the samples.

Alpha diversity

Alpha diversity metrics summarize the structure within an ecological community, in this case, OA saliva microbiomes. Three metrics were measured to indicate within-group diversity, namely Pielou's evenness (richness + evenness), microbial Richness (no. of species observed), and Shannon index (indicating species evenness). High richness indicates a high number of species available to inhabit sub-niches, while high evenness denotes that all species are of equal proportion instead of a few species dominating the environment (Pyron, 2010). Pieolou's evenness takes both richness and evenness measures into account. High Pielou's evenness, which denotes high alpha diversity, suggests higher resilience to temporal shifts in pathogenic microbes' microbiome and colonization resistance (Wade, 2021).

Microbial richness in the middle plot (Figure 3.3) showed that Jehai saliva microbiomes were richer in bacteria species compared to Temiar and Temuan ($p<0.05$). Shannon index on the extreme right plot also indicated Jehai had significantly higher evenness than Temiar and Temuan. High evenness means that the salivary microbes in Jehai are of more 'equal' proportion where each species are 'equally' dominating, instead of just a few species. Conversely, Pielou's evenness indicated that Jehai, sub-urban TemiarGM and urban Temuan had similar alpha diversity, meaning they were similarly homogenous. Temiar's alpha diversity was the lowest compared to the three OA communities ($p<0.05$). It was rather surprising to see that the saliva microbiomes of Temuan appeared to have almost equal intra-diversity compared to Jehai based on Pielou's evenness measure. All in all, Jehai only measured marginally higher alpha diversity than urban Temuan using richness and Shannon index, whereas Pieolou's evenness indicated similar diversity.

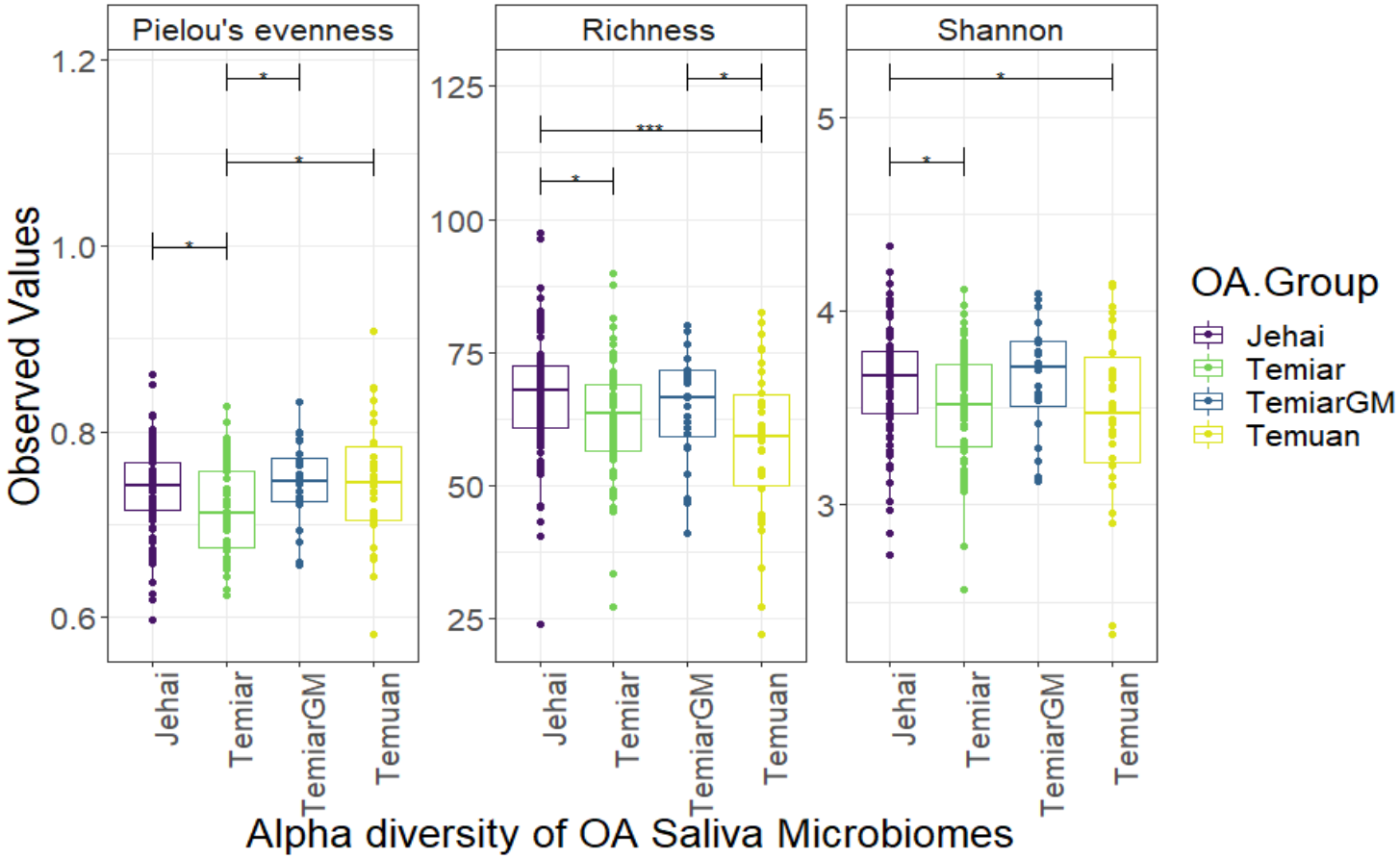


Figure 3.3 Alpha diversity boxplot for OA saliva microbiome. Temiar appeared to have lowest alpha diversity denoted by Pielou's evenness.

Beta diversity

Beta diversity, which measures dissimilarity between samples, was visualized on two ordination plots, PCoA and supervised CAP, using clr-transformed Euclidean distance at the ASV level (Figure 3.4). Sample ordination on unsupervised PCoA plot showed substantial overlap between the Temuan, TemiarGM and Temiar saliva microbiomes. Semi-nomadic Jehai overlapped with rural Temiar saliva microbiomes. In the supervised CAP plot, the saliva microbiomes formed four distinct clusters, with appreciable overlaps between Jehai and Temiar. Saliva microbiomes of TemiarGM still mapped as a subset of the Temuan samples despite different lifestyles and geographical locations. The saliva microbiomes of rural Temiar formed a distinct cluster away from sub-urban TemiarGM with no overlap. Temiar and TemiarGM were treated as distinct communities henceforth instead of forcibly grouping them together. Temiar saliva microbiomes seemed to cluster nearer to semi-nomadic Jehai, whereas TemiarGM saliva microbiomes seemed more similar to urban Temuans' microbiomes.

PERMANOVA multivariate test was utilised to determine community-level differences between the OA saliva microbiomes. The centroids of the OA groups were significantly different after controlling for BMI and age group (PERMANOVA Pseudo-F = 8.1576, R^2=0.104, p-value =0.001). PERMDISP was used to test for homogeneity of sample dispersion and revealed significant within-group variation (PERMDISP F-stats = 8.15, p-value =0.003). This result is in line with the ordination plots where sample dispersions were visibly different. This means that significant differences between the centroids of OA groups as tested by PERMANOVA may well be affected by non-homogenous sample dispersion (shape of the ellipses) instead of differences in saliva microbiomes (Anderson and Willis, 2003). Therefore further univariate differential abundance testing using ALDEX2 was required to investigate saliva microbial differences between the OA groups.

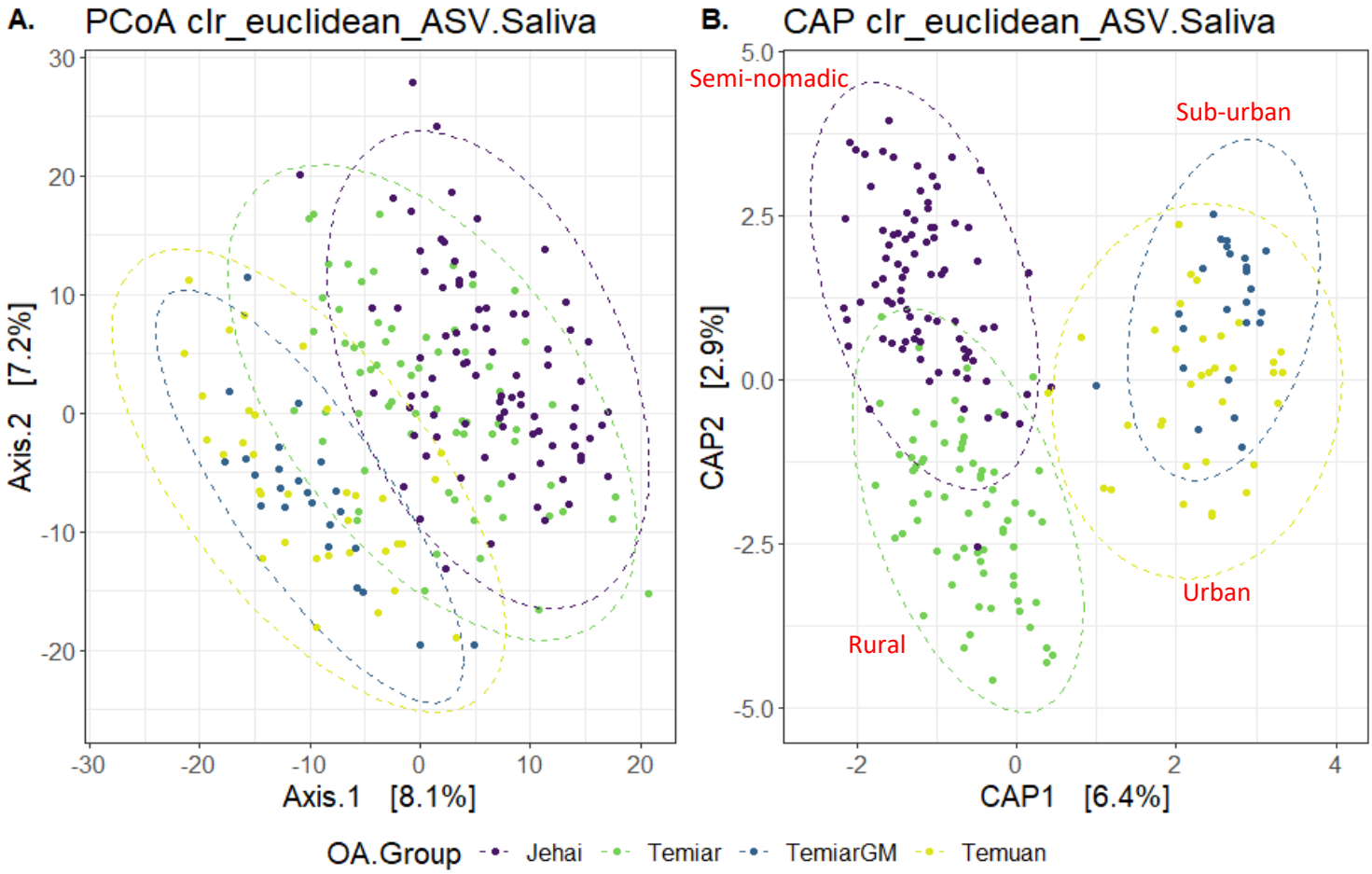


Figure 3.4 Ordination plots - PCoA (Left) and CAP (Right) depicting how dissimilar each saliva microbiome is from another. The more dissimilar the microbiome, the further away the samples will be from each other. The more rural OA groups (Jehai and Temiar) are clustering away from the urban-suburban OA groups (Temuan and TemiarGM).

3.1.3 Saliva differential abundance testing: OA group, obesity (BMI), gender, smoking and age-groups

The saliva microbiomes of Jehai were distinct from Temuan and TemiarGM by five bacteria genera. Amongst the five differentially abundant bacteria, *Corynebacterium, Actinomyces, Prevotella* and *Mogibacterium* (BH-corrected p-value <0.01, effect size >1) were increased in the saliva microbiomes of Temuan and TemiarGM compared to Jehai (Figure 3.5).

On the other hand, the saliva microbiomes of Temiar were differentiated from Temuan and TemiarGM by four bacteria genera (Table 2). Amongst the four genera, the relative abundance of *Corynebacterium, Prevotella,* and *Mogibacterium* were found to be increased in the saliva microbiomes of Temuan and TemiarGM compared to Temiar.

Saliva microbiomes of Jehai had no statistically distinct bacteria taxa from Temiar. The ordination plots (Figure 3.4) reflected this result, whereby their saliva microbiomes had some overlap indicating similarity. Temuan and TemiarGM saliva microbiomes had no differentially abundant taxa as well. There was no difference in the saliva microbiomes when tested for age groups, gender, BMI, and smoking habits. Sphingobacteriia that had been observed to be found in Jehai and Temiar saliva microbiomes in the taxonomy bar plot (Figure 3.1) has so far only been identified to the Family level, Lentimicrobiaceae and was not significantly different when investigated in Jehai against Temuan group.

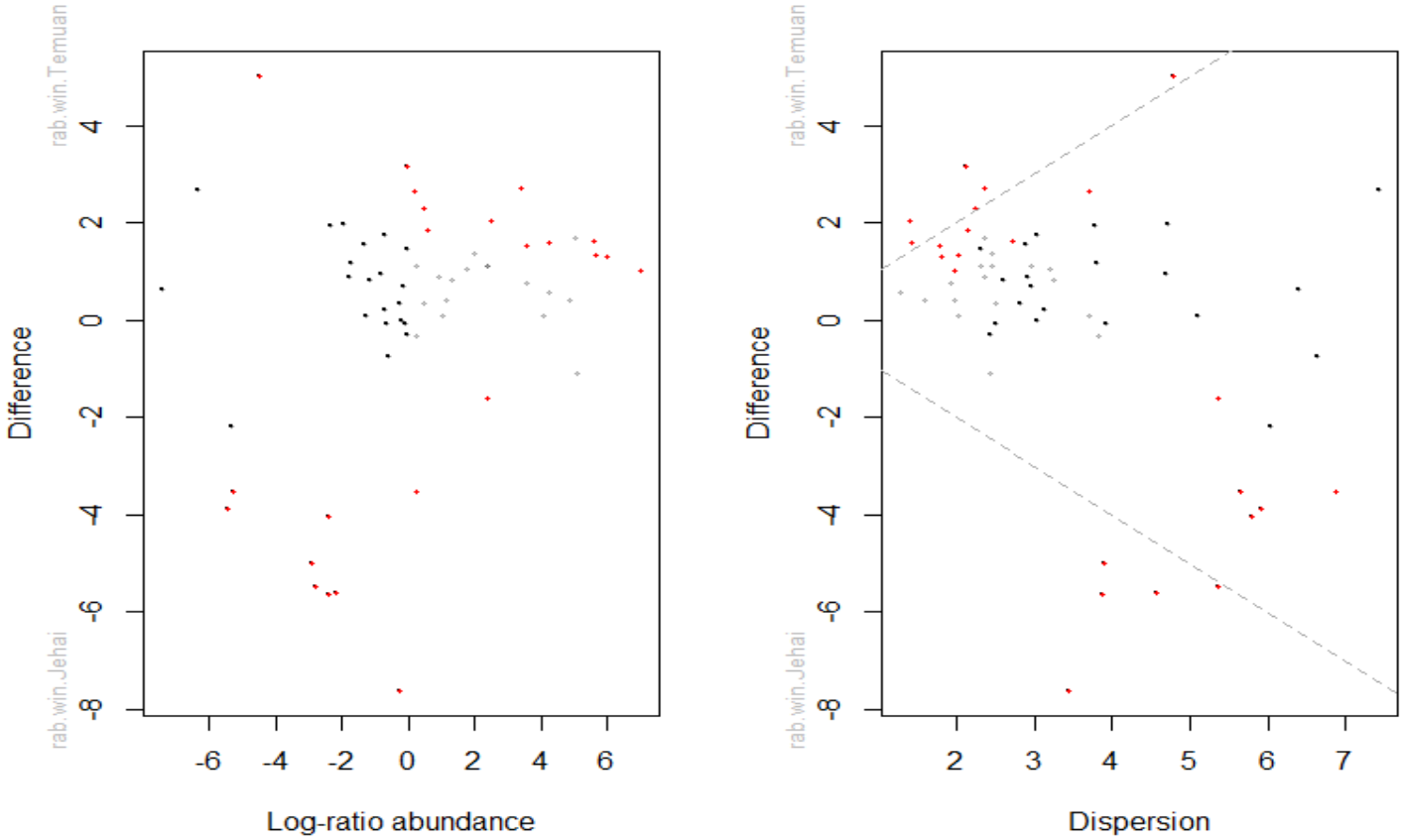


Figure 3.5 (Left) MA plot and (Right) Effect plot depicting differential abundance of bacteria taxa. The red dots are bacteria genera that significantly different in abundance (BH-corrected p-value <0.01), while grey dots are differentially abundant taxa but not statistically significant. Black dots are those that were not differentially abundant. The dotted diagonal line indicates dispersion equals difference (effect size = 1). Red dots found above the dotted line represents bacteria genera that have BH-corrected p-value <0.01 and effect size >1.

Table 2. Pairwise comparison of OA communities for differentially abundant salivary bacteria genera via ALDEX2 (BH-corrected p-value <0.01 and effect size >1).

Group	Genus[#]	effect	we.eBH	wi.eBH
Jehai vs Temuan	*Campylobacter*	1.407222	1.77E-10	5.61E-13
	Corynebacterium	1.286858	0.000525	3.63E-10
	Actinomyces	1.105451	9.47E-10	2.54E-10
	Prevotella	1.06383	5.24E-09	2.01E-10
	Mogibacterium	1.010736	1.50E-06	3.47E-09
Jehai vs TemiarGM	*Corynebacterium*	1.728029	0.001233	1.23E-09
	Actinomyces	1.004074	3.54E-07	3.08E-07
	Prevotella	1.385742	3.61E-09	7.06E-10
	Peptostreptococcus	1.036108	3.60E-06	5.01E-07
	Mogibacterium	1.6123	6.58E-20	2.23E-11
Temiar vs Temuan	*Campylobacter*	1.314972	5.97E-12	1.20E-11
	Corynebacterium	1.306579	0.000431	1.33E-09
	Prevotella	1.268683	8.42E-12	6.87E-11
	Mogibacterium	1.152272	3.07E-06	2.61E-09
Temiar vs TemiarGM	*Corynebacterium*	1.73868	0.000585	3.11E-09
	Alloprevotella	1.448045	5.42E-13	2.86E-09
	Prevotella	1.506326	4.28E-12	9.01E-10
	Mogibacterium	1.844286	1.02E-18	3.41E-11

[#]It should be noted that all differentially abundant saliva bacteria were detected in the negative controls, albeit at 0.03-1.46% abundance. *Mogibacterium* was the only bacteria not detected in the negative controls.

3.2 Gut microbiome: composition, diversity, and core constituents

Targeted 16S rRNA sequencing of the V4 region was used to characterize 165 OA gut microbiomes. Two samples failed to pass the QC threshold. An overall view was evident using bacterial composition, core constituents, alpha (intra-group) and beta diversity (inter-group) diversity. Multivariate analysis PERMANOVA (overall community-level) and differential abundance analysis (genus level) using ALDEX2 were performed to investigate differences in gut microbiome compositions between OA groups.

3.2.1 Gut microbial composition

The taxonomy bar plot (Figure 3.6) illustrates the overall gut bacterial composition at the class level for the four OA communities, namely semi-nomadic Jehai, rural Temiar, sub-urban TemiarGM and urban Temuan. Bacteria were arranged from most to least abundant class (Top to bottom). The top five bacteria classes of OA gut microbiomes were Clostridia, Bacteroidia, Actinobacteria, Gammaproteobacteria, Erysipelotrichia. Lower abundant bacteria class, Betaproteobacteria (dark brown) was more readily observed in the Temuan and Temiar communities while it was least observed in Jehai.

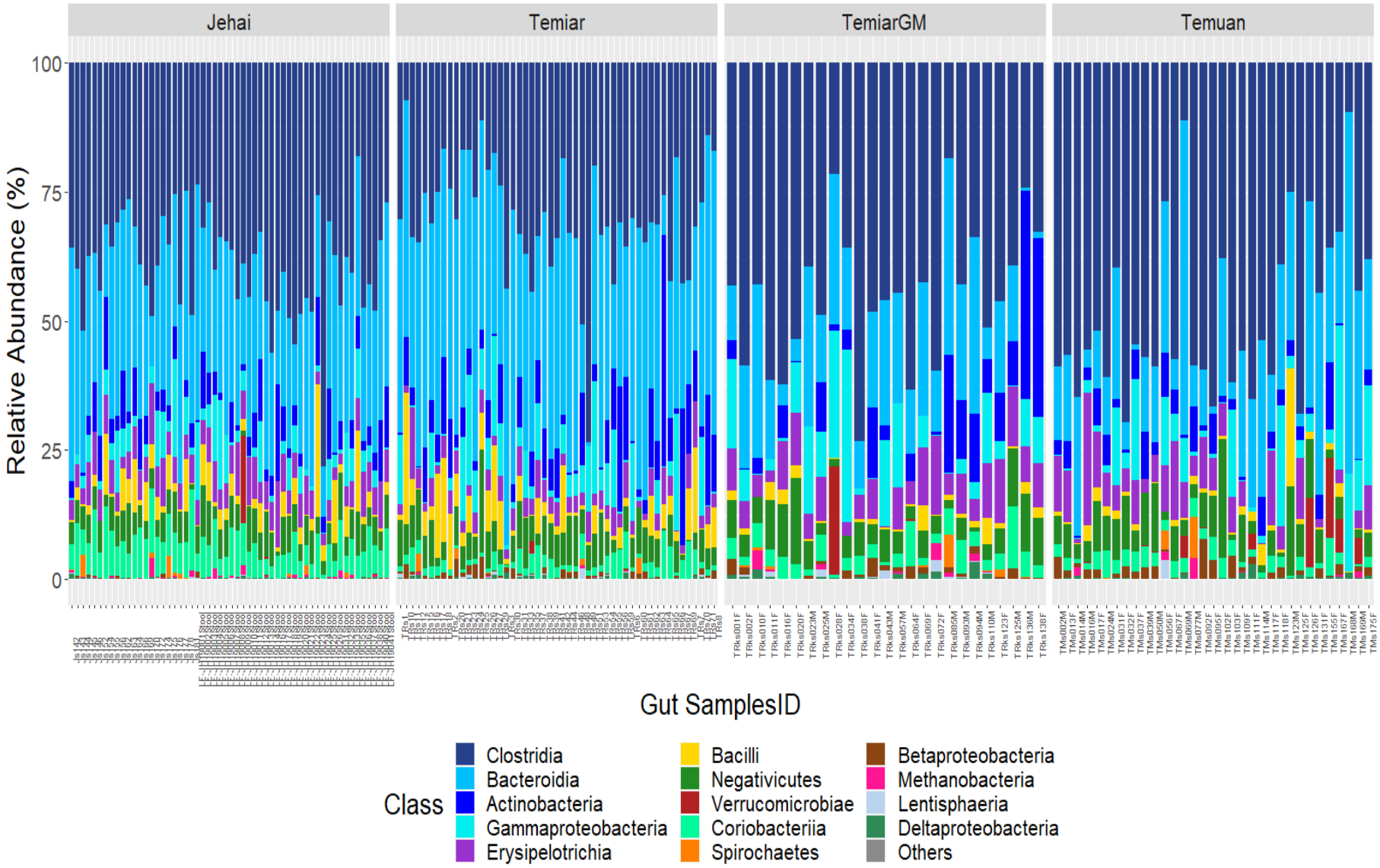


Figure 3.6 Gut taxonomy bar plot at the Class level

3.2.2 Gut Microbial diversity: core constituents, alpha, and beta diversity

Core constituents in gut microbiomes

Bacteria shared by at least 90% of the samples at 0.1% detection threshold revealed ten core genera (Figure 3.7). Nine out of ten of the core genera are from the phylum Firmicutes, class Clostridia Clostridia was one of the most abundant class observed in OA as evidenced in the taxonomy bar plot in (Figure 3.6). Other genera, specifically *Collinsella, Dorea, Blautia, Faecalibacterium, Lachnospiraceae,* and *Ruminococcaceae* were detected in at least 80% of the samples at 0.32% relative abundance.

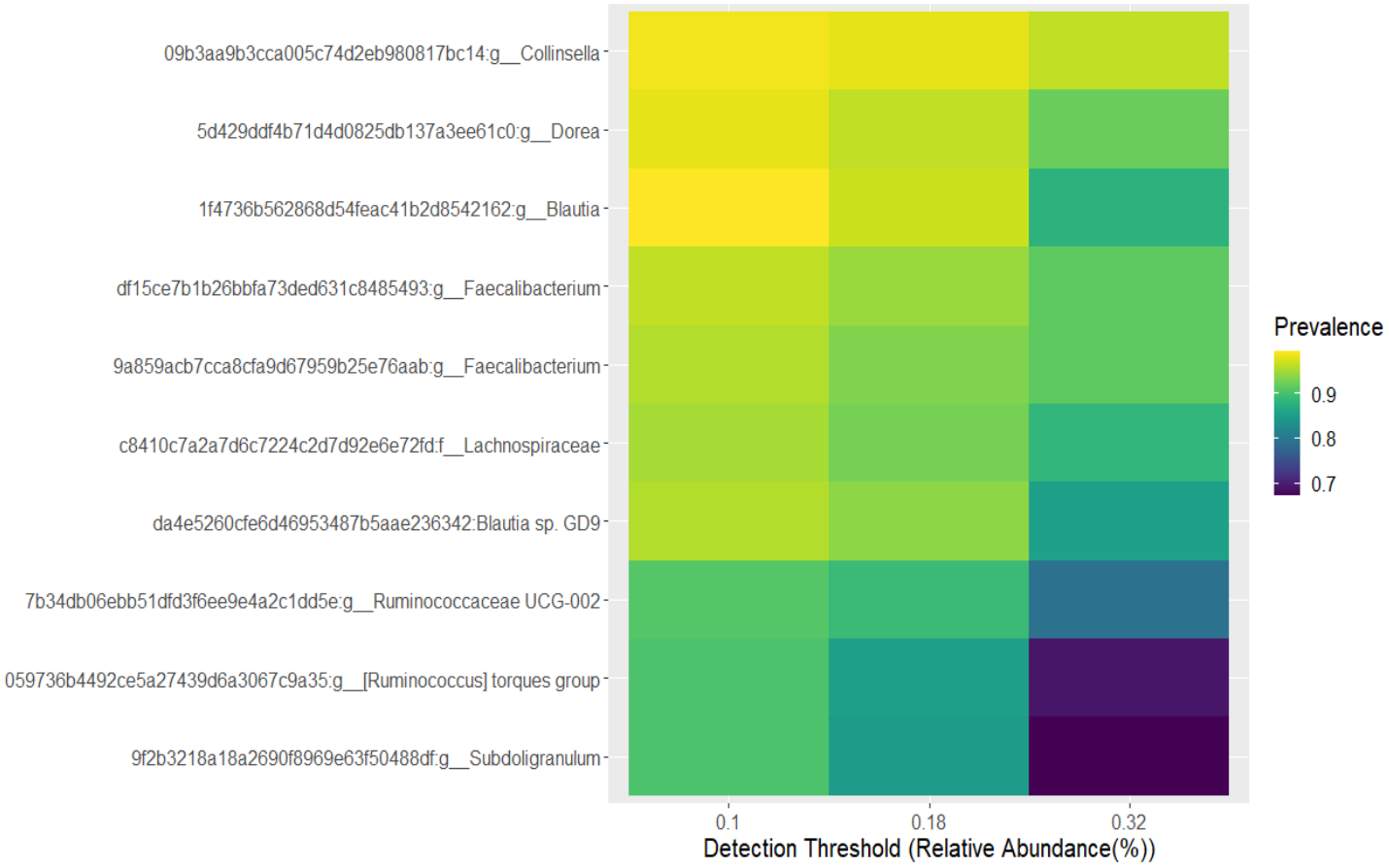


Figure 3.7 Core gut constituents at 0.1% relative abundance revealed nine core genera shared across 90% of the samples.

Alpha diversity

Alpha diversity metrics summarize the structure within an ecological community, in this case, OA gut microbiomes. Studies have reported higher alpha diversity among traditional communities, associated with high variation in traditional diets and higher resilience to temporal shifts in the microbiome (Dubois et al., 2017, Ruggles et al., 2018).

The alpha diversity measures chosen were Pielou's evenness which takes into account both species evenness and richness (left plot, Figure 3.8), microbial richness (middle plot) and Shannon index (right plot). Semi-nomadic Jehai exhibited the highest alpha diversity among the four OA communities for all three diversity measures ($p<0.001$)(Figure 3.8). OA communities who lived nearer to urban areas had observably decreased alpha diversity. Urban Temuans had the lowest alpha diversity for all three measures. Alpha diversity of both Temiar communities appeared to be an intermediate between the

Jehai and Temuan. Rural Temiar scored higher in microbial species richness than sub-urban TemiarGM, although it was not statistically significant.

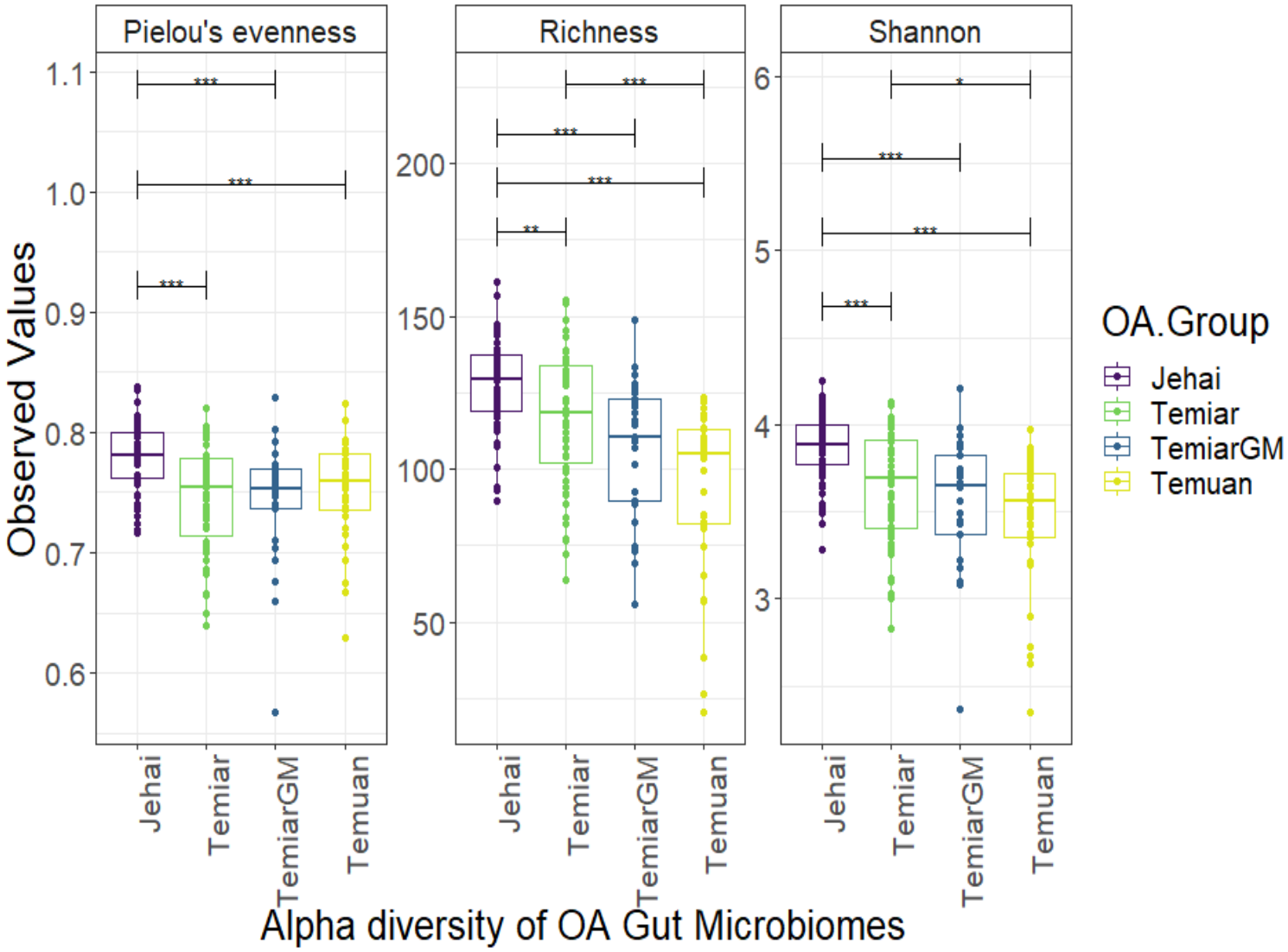


Figure 3.8 Alpha diversity box plot showing the forest-fringe Jehai with the highest alpha diversity.

Beta diversity

Beta diversity measures the microbial diversity between two groups and visualized on two ordination plots using centred log-ratio (clr) transformed Euclidean distance at the ASV (amplicon sequence variant) level.

The Jehai, Temiar and TemiarGM samples appeared to overlap, as seen in PCoA plot (Figure 3.9A). Temuan and TemiarGM samples appeared to have some similarities too. A Supervised CAP ordination plot was chosen to uncover possibly masked ordination in the unsupervised PCoA plot (Anderson and Willis, 2003). The gut microbiomes of OA communities separated into four distinct clusters with some overlaps (Figure 3.9B). Sub-urban TemiarGM appeared to have a mixture of 'microbiome types'. Most of the TemiarGM gut microbiomes overlapped with urban Temuans but a small portion appeared to

be leaning towards a Jehai-like microbiome type. Temiar and TemiarGM appeared separated on the ordination plot.

PERMANOVA multivariate analysis was utilised to investigate overall differences in gut microbiomes. It revealed the centroids of the OA groups were significantly different after controlling for BMI and Age-group (PERMANOVA Pseudo-F = 8.7755, R^2 = 0.141, p-value = 0.001). PERMDISP was used to test for homogeneity of sample dispersion and revealed significant within-group variation (PERMDISP F-stats = 16.476, p-value = 0.001). This result is in line with the ordination plot in Figure 3.9, where the sample dispersion of each OA community was visibly different (see the shape of the ellipses). This means that, although the gut microbiomes tested appeared to be significantly different in OA groups using PERMANOVA, the proportion of differences observed may well be affected by differences in sample dispersion (ellipses) as shown in PERMDSIP (Anderson and Willis, 2003). Therefore, further investigation, namely using univariate differential abundance analysis with ALDEX2, will be required to unequivocally determine true, biological differences in the gut microbiomes.

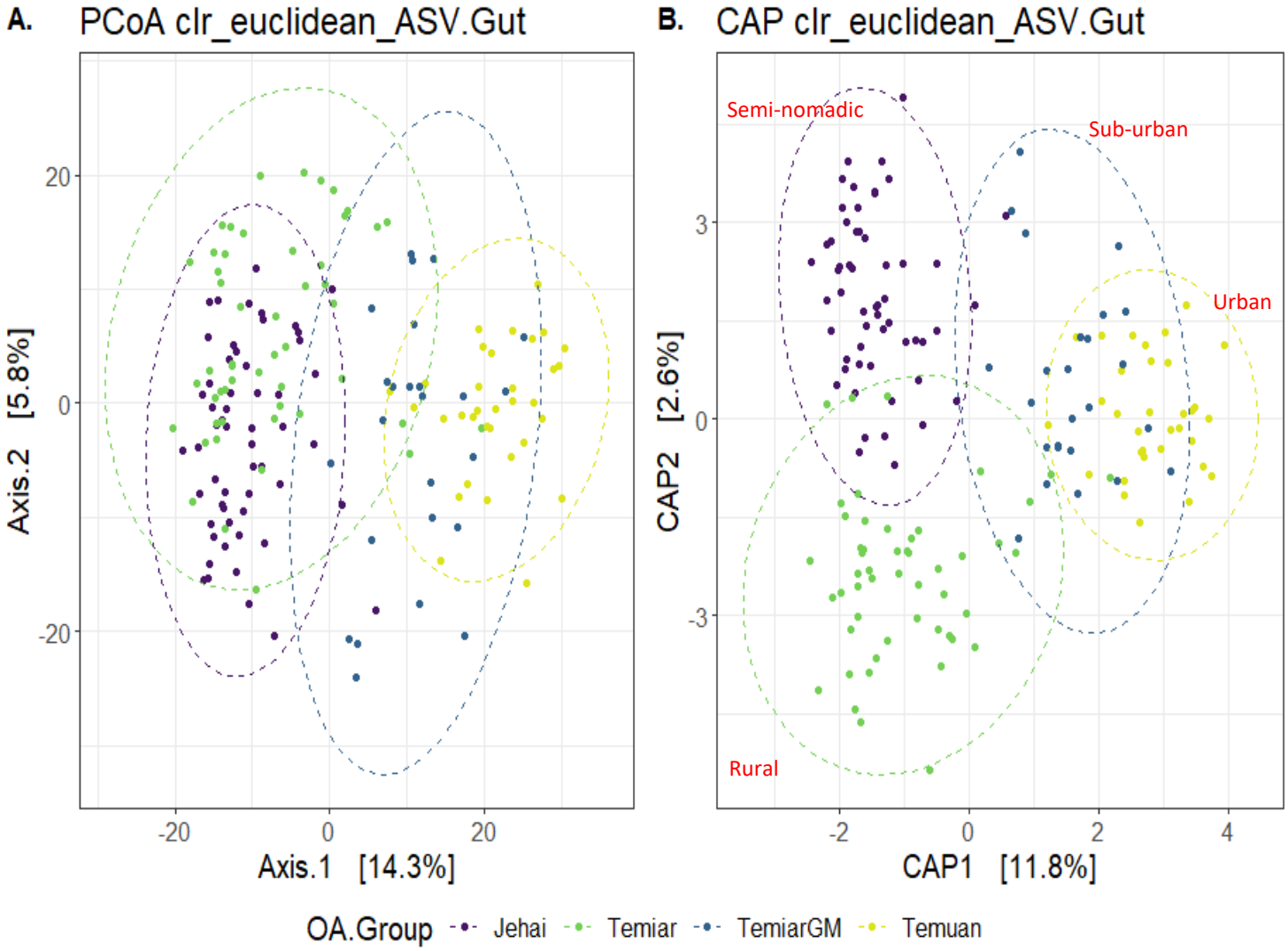


Figure 3.9 Beta-diversity unsupervised PCoA (A) and supervised CAP (RB) plots illustrating the dissimilarity of OA samples relative to each other. The OA communities formed four separate clusters in the CAP plot with some overlap.

3.2.3 Gut differential abundance testing: OA group, obesity (BMI), gender and age-groups

Differential abundance testing was done with ALDEX2 using centred log-ratio (clr) transformed data at the Genus level. The MA plot on the left in Figure 3.10 depicted different bacteria and their relationship to log-ratio abundance level. The Effect plot on the right showed differentially abundant bacteria and its dispersion where the dotted lines meant dispersion was equal to the difference (effect size = 1).

Significantly differential abundant bacteria were determined as the red dots above the dotted lines which indicated statistical significance (BH-corrected p<0.01) with an effect size >1. Effect size which is believed to be more robust and reproducible (Fernandes et al., 2018) may be more suited to microbiome studies to determine significantly different taxa between groups. P-value may be highly influenced by sample sizes, where the larger the sample size the higher the possibility of obtaining a significant p-value.

The relative abundance of seven genera was uncovered to be significantly depleted among the gut microbiomes of semi-nomadic Jehai compared to urban Temuans (Figure 3.10). These genera included *Sutterella, Odoribacter, Blautia, Lachnoclostridium, Parabacteroides, Bacteroides* and Ruminococcaceae UCG-013 (Table 3). The taxonomy bar plot in Figure 3.6 indicated that low abundance Betaproteobacteria appeared to be more abundant in the gut microbiomes of urban Temuan compared to semi-nomadic Jehai. The distribution of genus *Sutterella,* from the Betaproteobacteria class was depleted in gut microbiomes of Jehai.

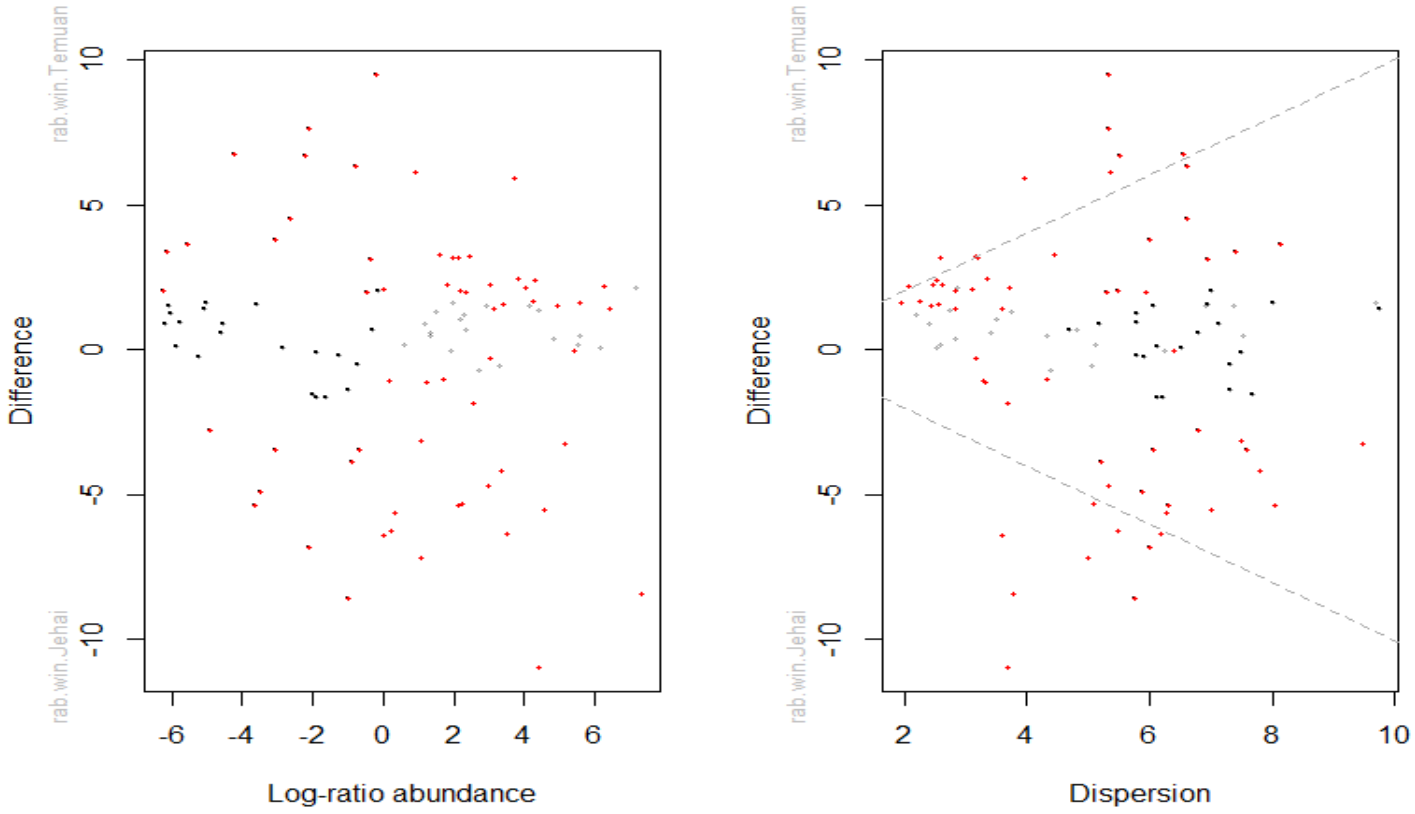


Figure 3.10 (Left) MA plot and (Right) Effect plot depicting differential abundance of bacteria taxa. The red dots are bacteria genera that significantly different in abundance (BH-corrected p-value <0.01), while grey dots are differentially abundant taxa but not statistically significant. Black dots are those that were not differentially abundant.

There were two genera, *Parabacteroides* and *Bacteroides,* that were also found to be decreased in gut microbiomes of Jehai compared to sub-urban TemiarGM. Rural Temiar and sub-urban TemiarGM had slightly distinct gut microbiomes where *Coprococcus 3* was lower in Temiar guts. Rural Temiar had lower abundance levels of five genera compared to urban Temuans, namely *Odoribacter, Blautia, Parabacteroides, Bacteroides,* and *Ruminococcaceae UCG-013.* The gut microbiomes of Jehai were not significantly different from rural Temiar.

There did not seem to be appreciable taxa differences in gut microbiomes when stratification was performed for age group, smoking habit, or gender. However, overweight individuals had higher abundance of two *Prevotella* genera, *Prevotella 9* and *Prevotella 2,* than obese individuals. The gut microbiomes of normal weight individuals were not substantially different from the overweight or obese individuals, which was rather unexpected. Four genera had significant p-values (BH-corrected p-value <0.01) for normal weight compared to obese individuals. However, the effect size was <1 which indicated that the differences observed may have been anecdotal rather than reproducible. It is also plausible that this may be a product of random chance. As such, those four genera were excluded.

Table 3 Pairwise comparison of OA communities for differentially abundant gut bacteria genera via ALDEX2 (BH-corrected p-value <0.01 and effect size >1).

Group	Genus[+]	effect	we.eBH	wi.eBH
Jehai vs Temuan	*Sutterella*	**1.1010**	**5.86E-07**	**2.56E-08**
	Odoribacter	**1.3725**	**1.42E-08**	**6.29E-10**
	Blautia	**1.0473**	**1.20E-07**	**1.73E-09**
	Lachnoclostridium	**1.0701**	**0.001156482**	**4.96E-08**
	Parabacteroides	**1.7592**	**2.11E-18**	**4.48E-13**
	Bacteroides	**1.4719**	**6.93E-11**	**5.03E-11**
	Ruminococcaceae UCG-013	**1.1493**	**2.87E-07**	**2.83E-08**
Jehai vs TemiarGM	*Parabacteroides*	**1.1132**	**1.22E-05**	**1.54E-07**
	Bacteroides	**1.1062**	**1.39E-06**	**1.68E-07**
Temiar vs TemiarGM	*Coprococcus 3*	**1.0654**	**1.87E-06**	**1.53E-06**
Temiar vs Temuan	*Odoribacter*	**1.1618**	**8.87E-08**	**2.52E-08**
	Blautia	**1.0481**	**5.80E-08**	**1.12E-08**
	Parabacteroides	**1.2430**	**9.39E-12**	**1.85E-10**
	Bacteroides	**1.0601**	**7.17E-08**	**3.06E-08**
	Ruminococcaceae UCG-013	**1.2154**	**3.69E-08**	**1.84E-08**
Overweight vs Obese	*Prevotella 9*	**1.1615**	**4.53E-11**	**2.52E-09**
	Prevotella 2	**1.1329**	**3.30E-10**	**3.24E-09**

[+]It should be noted that *Blautia* (2%)*, Prevotella 9* (6.4%) and *Prevotella 2* (1.04%) were detected in gut negative controls.

3.3 Microbiome controls: mock community, negative controls, and repeats

Earlier studies (Edmonds and Williams, 2017, Bukin et al., 2019, Hornung et al., 2019) have extolled the limitations of Next Generation Sequencing (NGS) which include batch effects, the 'kitome' – contamination from commercial and in-house extraction kits, as well as the effect of sequencing different 16S rRNA hypervariable region. Currently there is no way to separate contamination from actual biological sequences in NGS datasets. Nor is there any standardized method to analyze datasets sequenced at different 16S rRNA hypervariable regions. This section describes the negative and positive microbiome controls used.

The taxonomy bar plot (Figure 3.11) shows the composition of 24 bacterial classes. From the far left of the bar plot, are two mock communities (ATCC MSA-1002) that underwent the same PCR amplification procedure of the V3-V4 region as all the samples in 2017 and 2020 respectively. There appears to be a batch effect for the same mock community sequenced using the same protocol. The PCR blank was the PCR negative control consisting of ultrapure water. The most dominant bacteria in the negative control was Gammaproteobacteria, followed by Betaproteobacteria. Saliva kit and stool kit blank were ultrapure water that underwent the same DNA extraction methods as the DNA samples respectively. The "kitome" appeared to be quite similar despite saliva being extracted using a home-made kit while stool was extracted using a commercial kit. The top five genera found in the kit and PCR negative controls were *Pseudomonas* (50.96%), *Streptococcus* (17.57%), *Neisseria*(14.70%), *Prevotella 9* (10.21%) and *Prevotella 7* (9.67%) (full table of contaminants in supplementary).

In order to ensure that the bacteria that tested differentially abundant in ALDEX2 were true, biological differences instead of contaminants, ALDEX2 results were manually checked against the list of contaminant bacteria found in the negative controls. Table 4 is a summary of the differentially abundant bacteria that was also identified in the negative controls as well as the extent of contamination. With the exception for *Blautia* and *Prevotella 9* found in the stool-negative control with contaminant abundance of 2% and 6.4% respectively, all contaminant genera were less than 2%.

The samples labelled with "_trim" were samples that were randomly chosen to be re-sequenced on V3-V4 before being trimmed to V4. Its counterpart was sequenced at the V4 region. There were a total of three stools and three saliva samples that were re-sequenced to be used as a comparison of the microbiome composition using two different hypervariable regions. Stool appeared more susceptible to different extraction method, 16S rRNA hypervariable region and sequencing platform as the re-sequenced "trimmed" gut samples appeared quite different from their V4 counterparts, as previously reported (Sinha et al., 2017, Videnska et al., 2019). Saliva samples, on the other hand, had more similar profiles to their counterparts despite differences in extraction method, 16S rRNA hypervariable region and sequencing platform. This observation was also previously reported (Lim et al., 2017, Rosenbaum et al., 2019).

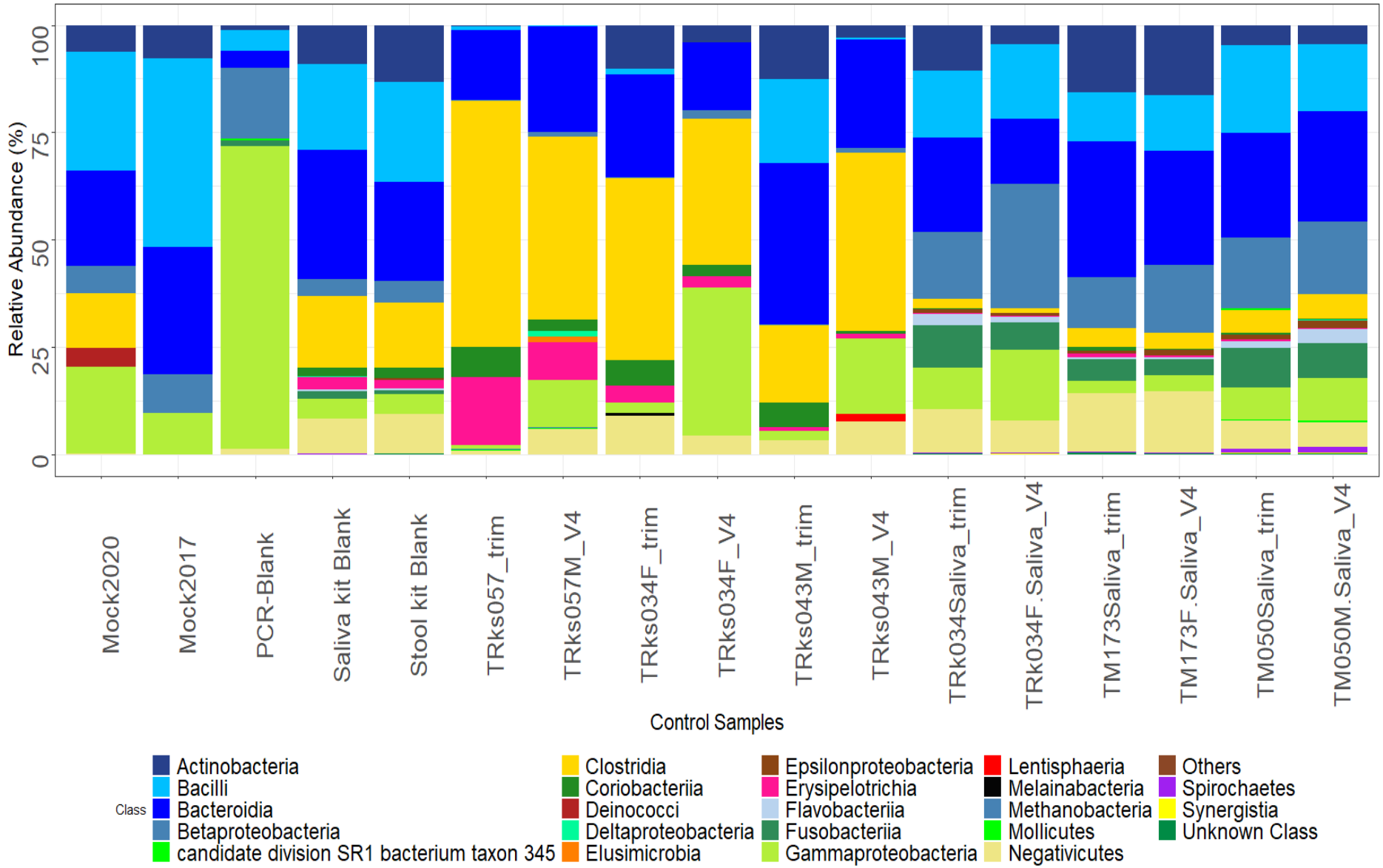


Figure 3.11 Taxonomy bar plot of control samples including Mock community (positive), PCR and extraction kit blanks (negative) and re-sequenced stool (n=3) and saliva (n=3) samples.

Table 4. Summary of all bacteria genera that were found in the negative controls and tested differentially abundant in ALDEX2 analysis.

Saliva OTU_ID	Genus	Sample ID	Sample Type	Body Site	Abundance (%)
1363	*Campylobacter*	Saliva-neg	PCR	NA	0.158277
1967	*Corynebacterium*	Saliva-neg	Kit	Oral	0.033322
1624	*Actinomyces*	Saliva-neg	Kit	Oral	0.897599
565	*Prevotella*	Saliva-neg	Kit	Oral	1.464065
1137	*Peptostreptococcus*	Saliva-neg	Kit	Oral	0.05623
1655	*Alloprevotella*	Saliva-neg	Kit	Oral	1.412
1662	*Alloprevotella*	LF-Negative	Kit	Oral	1.06383
Gut					
295	*Blautia*	Stool-neg	PCR	NA	2
1913	*Prevotella 9*	Stool-neg	Kit	Gut	6.438462
2206	*Prevotella 2*	Stool-neg	Kit	Gut	1.046154

3.4 Comparison of methods for mixed hypervariable region analysis

The various datasets that were obtained (V4) and generated (V3-V4) required a suitable analytical approach to combine them. SATé-enabled phylogenetic placement (SEPP) technique by Janssen et al. (2018) was utilised which allows mixed-region analysis. Sequencing reads at one nucleotide difference, also known as Amplicon Sequence Variant (ASV) was used to construct a phylogenetic tree via SEPP. Sequences with less than 75% similarity to the reference tree were removed. SEPP constructed tree can mitigate differences in sequencing region, length and platform, thus allowing analysis of biological differences in mixed-region datasets over technical differences (Janssen et al., 2018).

SEPP-method

Beta diversity was measured using unweighted UniFrac distance metric with SEPP constructed tree and clr-transformed data because Euclidean distance, which does not take into account the phylogenetic distance between samples, resulted in the V4 dataset (Temuan and TemiarGM) to be poorly differentiated (Figure 3.13). However, even using UniFrac and SEPP-generated phylogenetic tree, there was a random separation (Figure 3.12). It is most obviously seen in the purple Jehai samples on the left in the PCoA plot, which appeared to be mitigated in the CAP plot.

Based on the CAP plot using unweighted UniFrac, the OA communities formed four clusters with some overlap. However, it was important to ascertain whether the apparent clustering was based on actual biological differences or merely a reflection of differences in sequencing regions. This was because the V3-V4 datasets, including Jehai and Temiar, appeared on the left side of the plot, while the V4 datasets, Temuan and TemiarGM, appeared on the right side. There was no overlap between the V3-V4 and V4 datasets.

To test this result, three random TemiarGM gut samples were re-sequenced using V3-V4 primers. I found that these failed to cluster with their V4 counterparts, supposedly appearing on the right. Instead, they clustered with the V3-V4 samples on the left. These results suggested that SEPP method may not be sufficient to resolve most differences in variable gut microbiome datasets. This calls to question its suitability for meta-analysis as promoted by the QIIME2 community and the authors.

Testing for suitability of unweighted UniFrac distance metric

When using the SEPP-method, there was a separation horizontally across the PCoA plot (Figure 3.12). Reducing taxonomic resolution from ASV level, where bacteria taxa may be distinguished by single nucleotide level, to genus level seemed to make the separation more prominent (Supp Figure 2). To further investigate this, the V3-V4 dataset was trimmed to only V4 region, generated SEPP-tree, and plotted using unweighted UniFrac (Supp Figure 3). The separation was still observed in both ordination plots, suggesting perhaps unweighted UniFrac was not suitable for this analysis. Therefore, Euclidean distance was finally chosen to calculate dissimilarity for all data in this study.

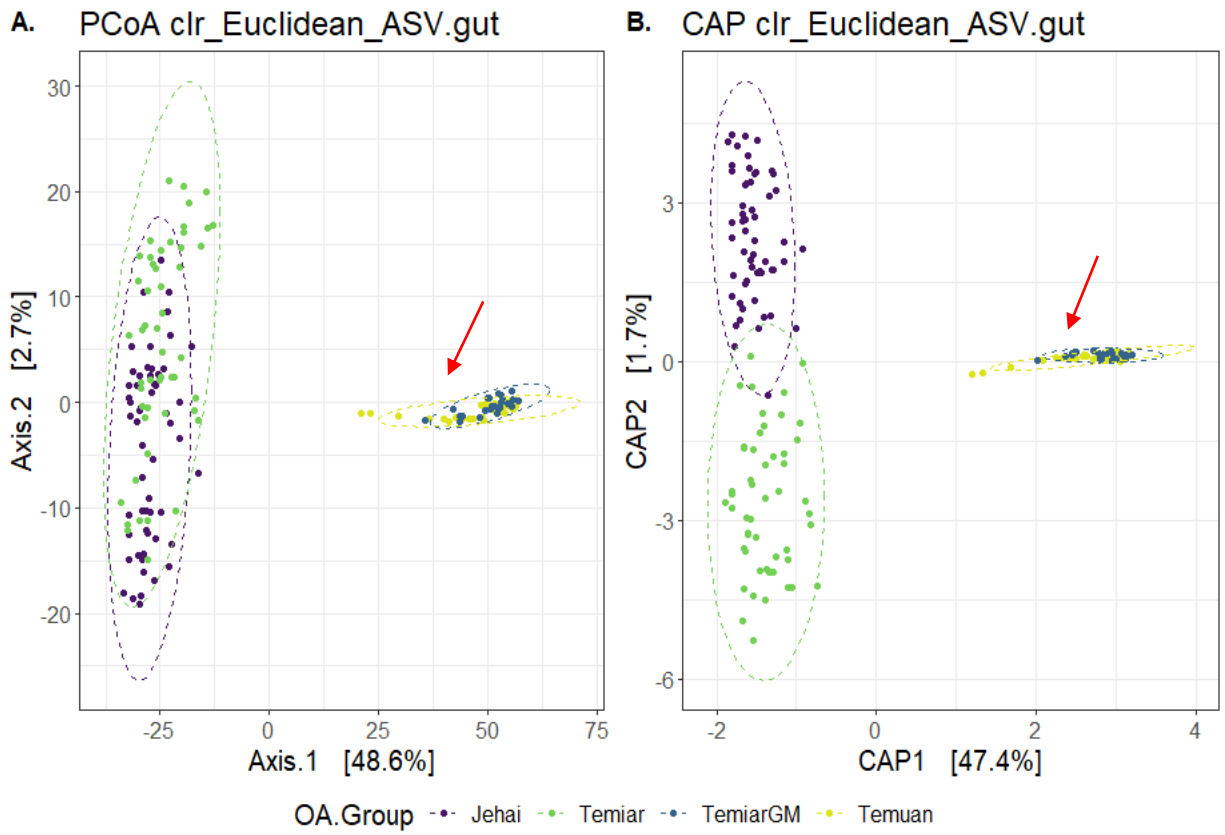


Figure 3.13 Ordination plots of V3-V4 and V4 data plotted using Euclidean (non-phylogenetic) metric showed poor differentiation for V4 samples (red arrow)

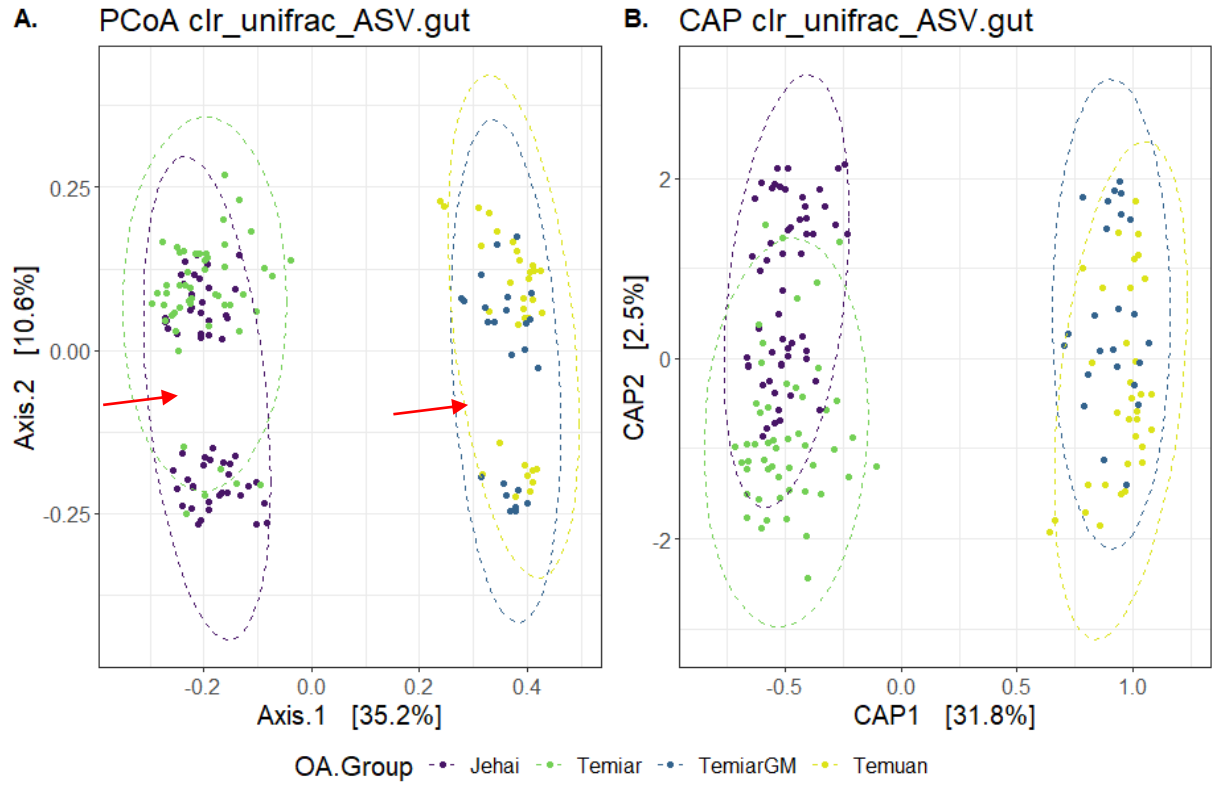


Figure 3.12 Ordination plots of V3-V4 and V4 data plotted using SEPP method and unweighted UniFrac metric shows a seemingly random separation across the samples.

3.5 Shotgun Metagenomics of Jehai Gut

The semi-nomadic Jehai community was the most isolated OA group in this study. They still rely mostly on traditional food from the forest, supplemented by store-bought rice and canned-food when possible. Due to limited funds, shotgun metagenomics was conducted only on twelve Jehai gut samples as a pilot investigation. This section describes, for the first time, the Jehai gut metagenome composition, core bacterial species and identified *novel bacteria species. Lastly, the quality of Jehai metagenomes were compared against the Unified Human Gastrointestinal Genome (UHGG) catalogue (Almeida et al., 2020). The UHGG catalogue is a specially curated database comprising >200,000 high quality genomes from 4,644 gut prokaryotes, of which >70% of the species have yet to be cultured.

3.5.1 Taxonomic classification of metagenomes

Reads were split into 31 *k*-mers and classified against UHGG reference database by the lowest common ancestor using Kraken2. An average of 91% of reads were classified against the UHGG catalogue. This is a good indication which demonstrates the comprehensiveness of the UHGG database for the Jehai reads. The taxonomy bar plot (Figure 3.14) shows the classification of bacterial species found in the Jehai gut metagenome at the order level. “Ambiguous” species shown here in light blue are reads that were mapped to multiple orders. Bacteria that were classified against UHGG but present at very low abundances were categorised as “Other”. Reads that were unclassified (~9%) are potentially *novel species that will be further investigated.

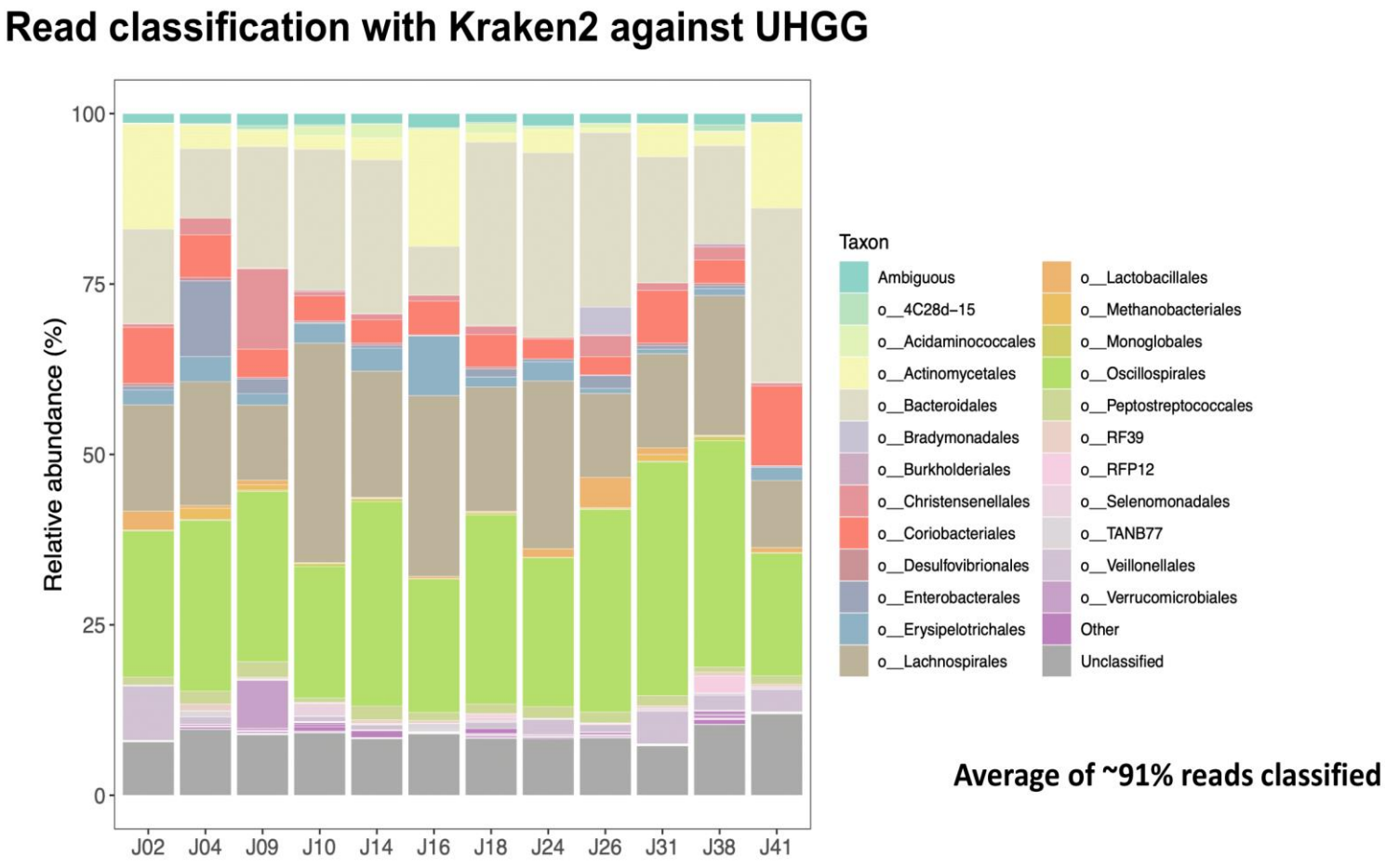


Figure 3.14. Taxonomy bar plot showing bacterial species found in Jehai gut metagenome at the order level.

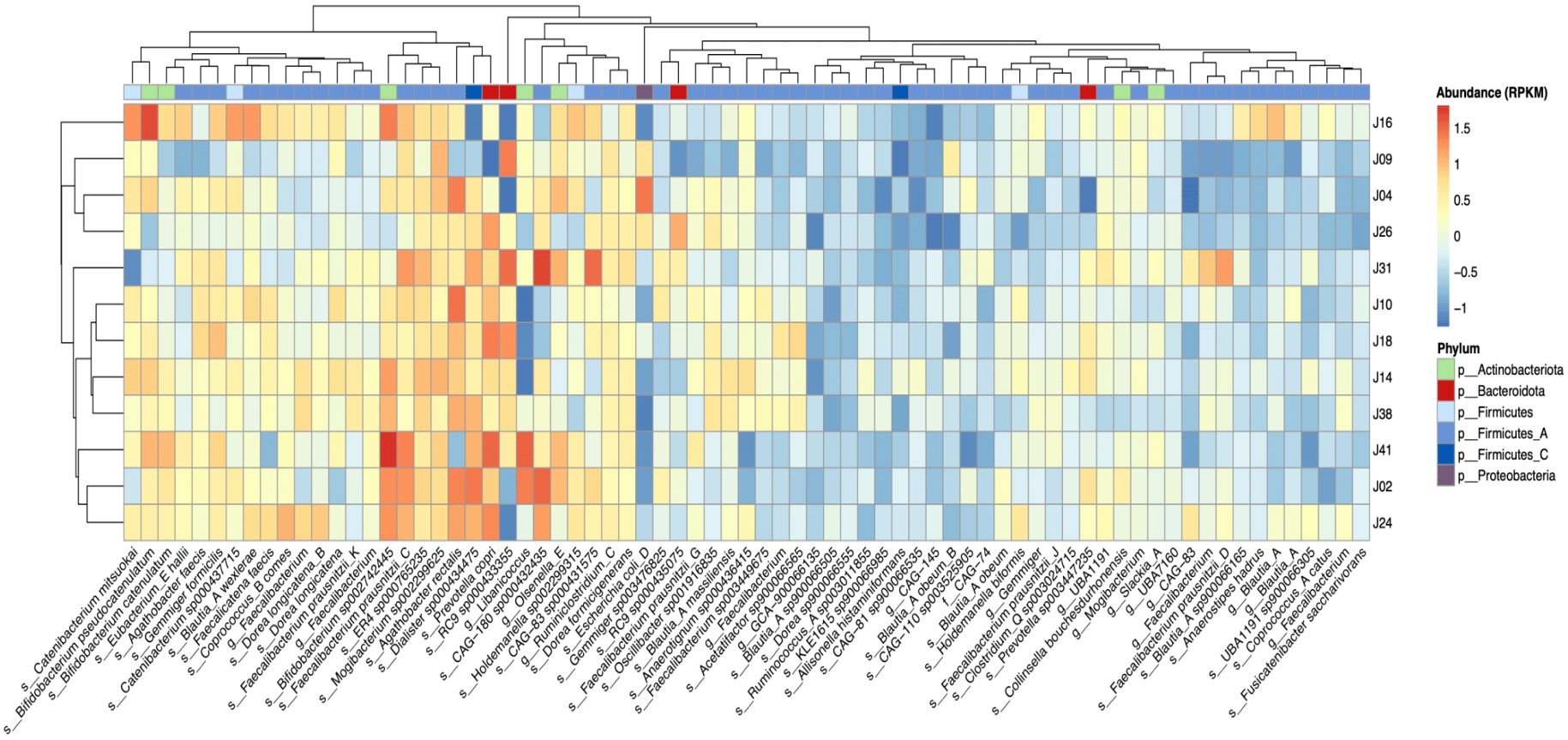
3.5.2 Core species of Jehai gut metagenomes

MAGs (Metagenome-assembled genome) were mapped to UHGG using BWA aligner. To overcome the problem of reads mapping to multiple species, BWA was set to show species that mapped uniquely to UHGG. Species annotation was based on GTDB (Genome taxonomy Database), which is considered as the gold standard due to it mapping taxonomy based on genome relatedness compared to NCBI database which is more prone to classify taxonomy based on "historical" annotations (Coil et al., 2019, Glendinning et al., 2020). There were a total of 73 core bacterial species found in all twelve Jehai gut metagenomes (Figure 3.15).

Strikingly, many core species identified in the core microbiome fell within relatively new bacterial species, in particular those which have yet to be characterised and are still classified by numbers. A high proportion of core species are from the phylum Firmicutes A and Firmicutes, followed by Actinobacteriota and Bacteroidota. Strikingly, *Prevotella copri,* a gut bacteria of recent clinical interest, was found in high abundance in the Jehai gut metagenome (red-orange rectangles).

Another notable finding was that several strains of *Faecalibacterium prausnitzii* were detected, albeit with varying abundance in the core microbiome. *F. prausnitzii* is another gut bacteria of clinical interest, commonly found in the gut of healthy individuals. Similarly, several species of *Blautia,* many uncultivated, have been found at low abundance levels (blue rectangles).

Species prevalence/abundance with BWA against UHGG



73 core species (found in all 12 samples)

Figure 3.15. Heatmap showing 73 core species found in all Jehai gut metagenomes. Many species have yet to be well-characterized and are temporarily delineated using place-card numbers, awaiting further information.

3.5.3 *Novel bacteria species

From 1074 MAGs, 379 bacterial species were identified in the Jehai gut metagenome, of which 28 appeared to be *novel, where the metagenome sequences could not be classified based on any known species in the UHGG reference database. The *novel species identified were found within different phyla. Some of these species clustered together, namely in the phylum Actinobacteriota, Firmicutes A, C and B. One species was observed in Campylobacterota and another in Desulfobacterota A (Figure 3.16). The 'spikes' surrounding the circular tree, depicting the number of genomes per bacteria phyla, were evenly distributed (high evenness). This revealed that the Jehai gut metagenomes had high microbial richness and high evenness, indicators of a diverse and robust microbiome.

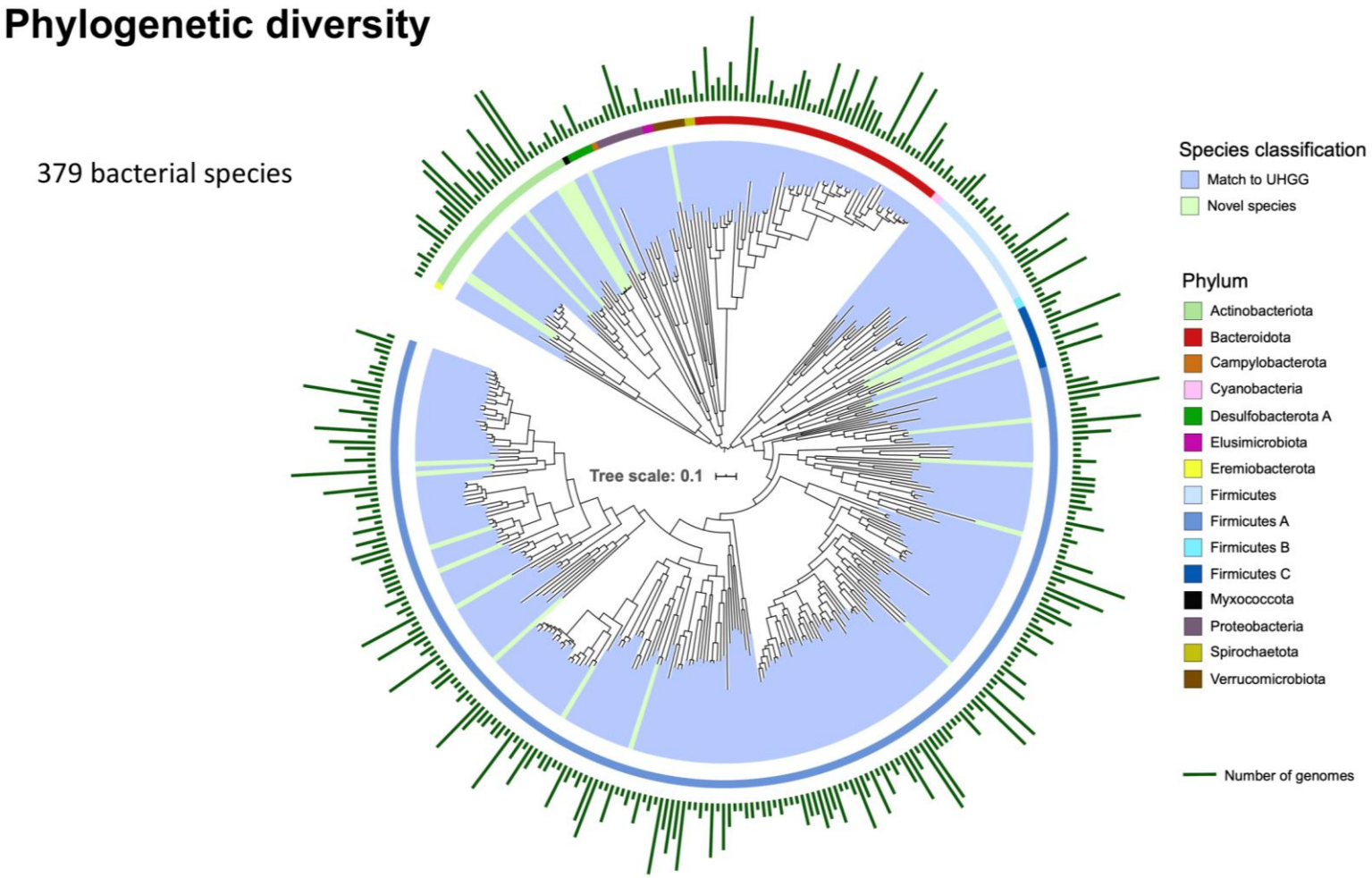


*Figure 3.16. Tree of life depicting phylogenetic placement of the *novel bacteria species among MAGs mapped to UHGG, as well at the number of genomes per bacteria phylum.*

*Abundance of *novel species from the genus* Collinsella

The heatmap in Figure 3.17 shows the abundance of *novel species, focusing on species found on the left of the heatmap, indicated by red squares. Those species were found in high abundance across all Jehai individuals, suggesting those to be "true" *novel species. Putative 'novel' species found at low abundance or in one individual should be treated with caution and not ascribed as 'novel' in the

absence of more evidence. Strikingly, several *novel species classified under the genus *Collinsella* were observed at high abundance levels in all Jehai guts. There were a few bacteria classified only to the family level. These could possibly be new bacteria families or genera waiting to be discovered. However further studies and sequencing of more samples have to be done for more information.

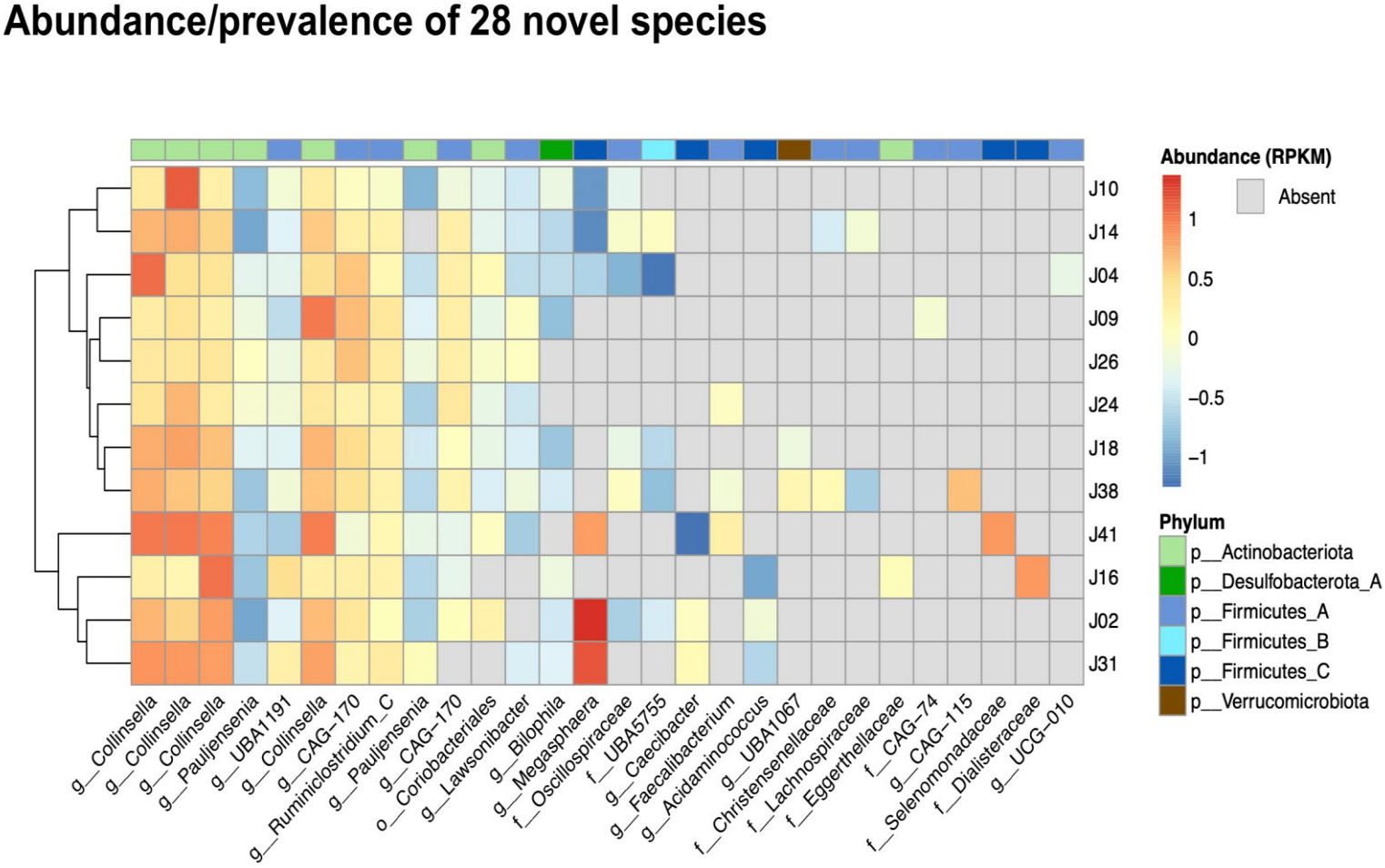


*Figure 3.17. Heatmap of *novel gut bacterial species. *Novel species from the phylum Actinobacteriota, from the genus Collinsella are present at high abundance in all Jehai individuals.*

3.5.4 MAGs versus UHGG genomes

MAGs aligned to UHGG reference genomes

MAGs generated from Jehai guts were compared to UHGG genomes using a pairwise method using dRep at a threshold of ANI (Average nucleotide index) ≥95% and AF(Alignment Fraction) ≥ 30%, both measures of nucleotide-level genomic similarity. MAGs were dereplicated to remove strain redundancy and plotted against UHGG in a scatter plot to visualise the percentage of MAGs that aligned to UHGG (Figure 3.18).

Each coloured dot on the graph represents a MAG. The x-axis indicates how well the MAGs align to the UHGG reference. CheckM, a tool that assesses the level of genome completeness and contamination was utilized. High quality or near-complete MAGs visualized as solid red dots have a quality score of >90% completeness and <5% contamination. Medium quality MAGs are solid blue

dots, scored at 50-90% completeness and <10% contamination. Three quarters of the total MAGs were aligned to more than 50% of reference genomes in UHGG database. The majority of the high quality MAGs mapped very well to the reference genomes.

Remarkably, 1074 unique MAGs were discovered in the Jehai gut metagenomes. Of these, 38 MAGs (3.8%) failed to match to the UHGG database.

Assessment of MAG quality

Quality scores were calculated using (level of genome completeness – 5 x contamination). The diagonal line on the graph means that the quality score is the same for the reference genome and MAG. MAGs that cluster above the diagonal line are of lower quality compared to the reference while those that cluster below the diagonal line are of higher quality.

As observed, 953 MAGs (88.7%) which represent 330 species were of lower quality than the reference. This is to be expected because the UHGG database was generated from over 300,000 MAGs from many whole genome sequencing datasets hence have a higher chance of obtaining a high quality or near-complete sequence. Surprisingly, 54 MAGs (5%) representing 43 species were of higher quality than the reference. Notably, one MAG reached a quality score of 100, exceeding the reference genome score of 70.

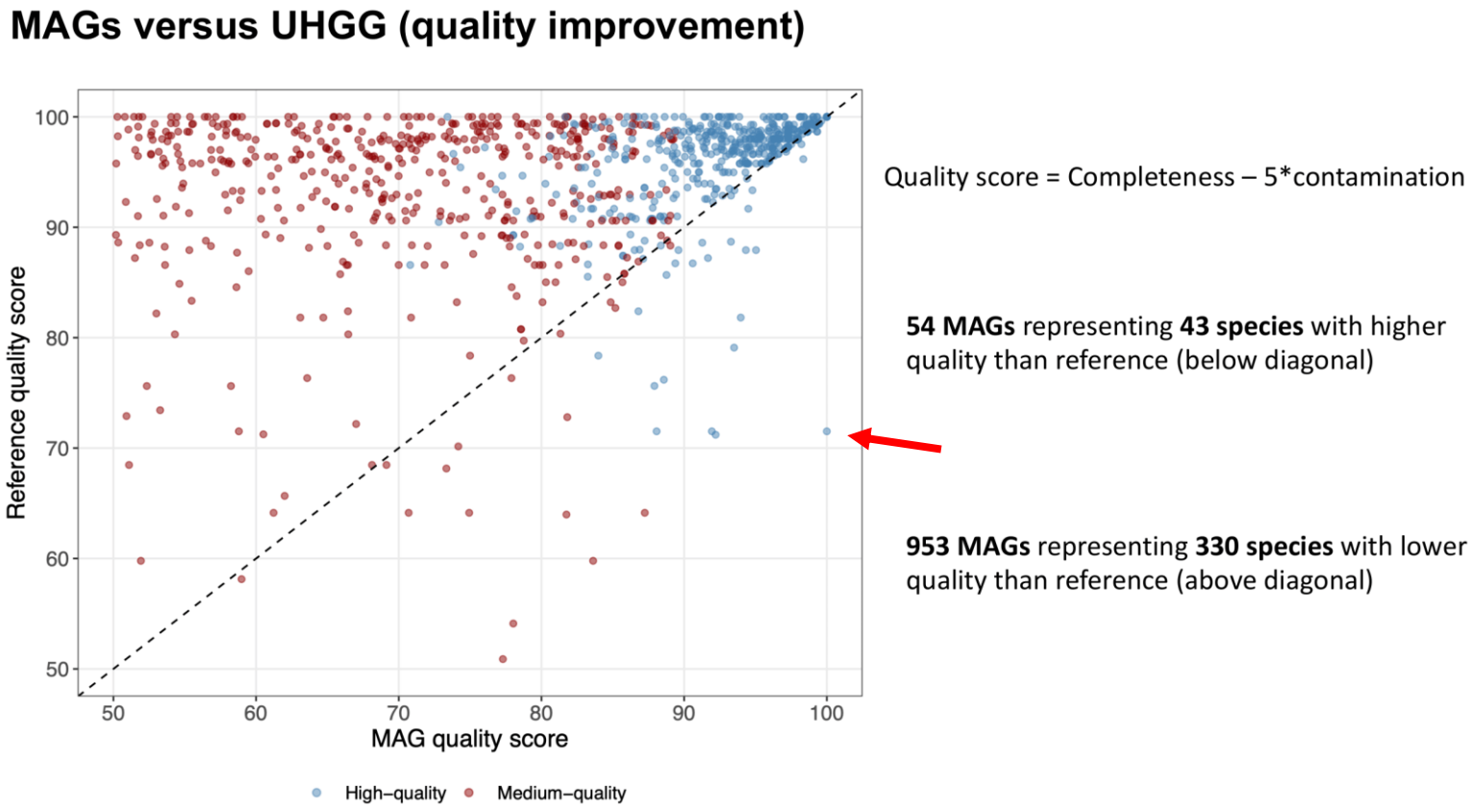


Figure 3.18 Scatterplot indicating how well Jehai MAGs mapped to UHGG reference genomes. The red arrow points to a MAG with a quality score of 100, which exceeds the reference genome score of 70.

3.6 Cardio-metabolic health trends in three OA groups

Aspects of cardio-metabolic health were appraised in 215 OA individuals, of which there were 135 women with a mean age of 33 ±1.9, and 80 men with a mean age of 40 ± 3.4. Health factors assessed include general and abdominal obesity, blood pressure, blood glucose, insulin resistance using a surrogate marker TG/HDL and self-reported smoking habits. Framingham Risk Score for cardiovascular risk, in the next five years was calculated using a modified calculator that was recalibrated for an Australian indigenous cohort (Hua et al., 2017). Metabolic syndrome was also determined by having three or more of the five risk factors (Grundy et al., 2005).

3.6.1 General and abdominal obesity

General obesity, body mass index (BMI ≥30 kg/m^2), prevalence was the highest among the urban Temuans at 70.6%. Surprisingly, the semi-nomadic Jehai had a slightly higher percentage of obese individuals compared to rural Temiar (25.3% obese Jehai vs 20.6% obese Temiar). Nevertheless, Temiar had the largest proportion of overweight individuals (42.3%) among the three groups (Figure 3.19). The semi-nomadic Jehai not only displayed the highest percentage of normal weight individuals (42.2%) but also the highest percentage of underweight individuals (15.7%). Less than half of the Jehai who participated had healthy weight. Almost 60% of them were either overweight/obese or underweight. From the obesity graph, it appears that obesity rate increased among OA groups who lived closer to urban areas, exemplified by the Temuan who live in a metropolitan area in bustling Selangor. Abdominal obesity was used as an additional risk indicator for cardio-metabolic diseases. More than half of each OA group had waist circumference indicated they were at risk (Figure 3.20).

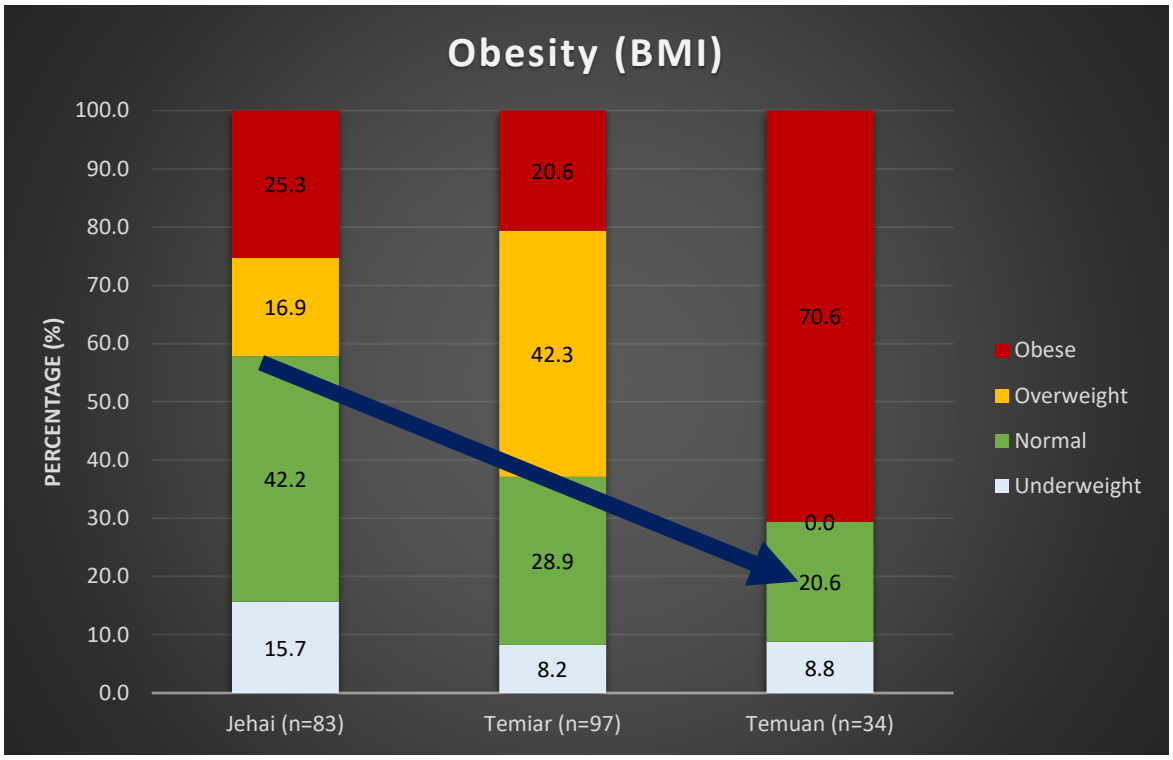


*Figure 3.19 The total proportion of overweight and obese individuals was lowest in semi-nomadic Jehai and highest in urban Temuan. *Sample size insufficient for statistics*

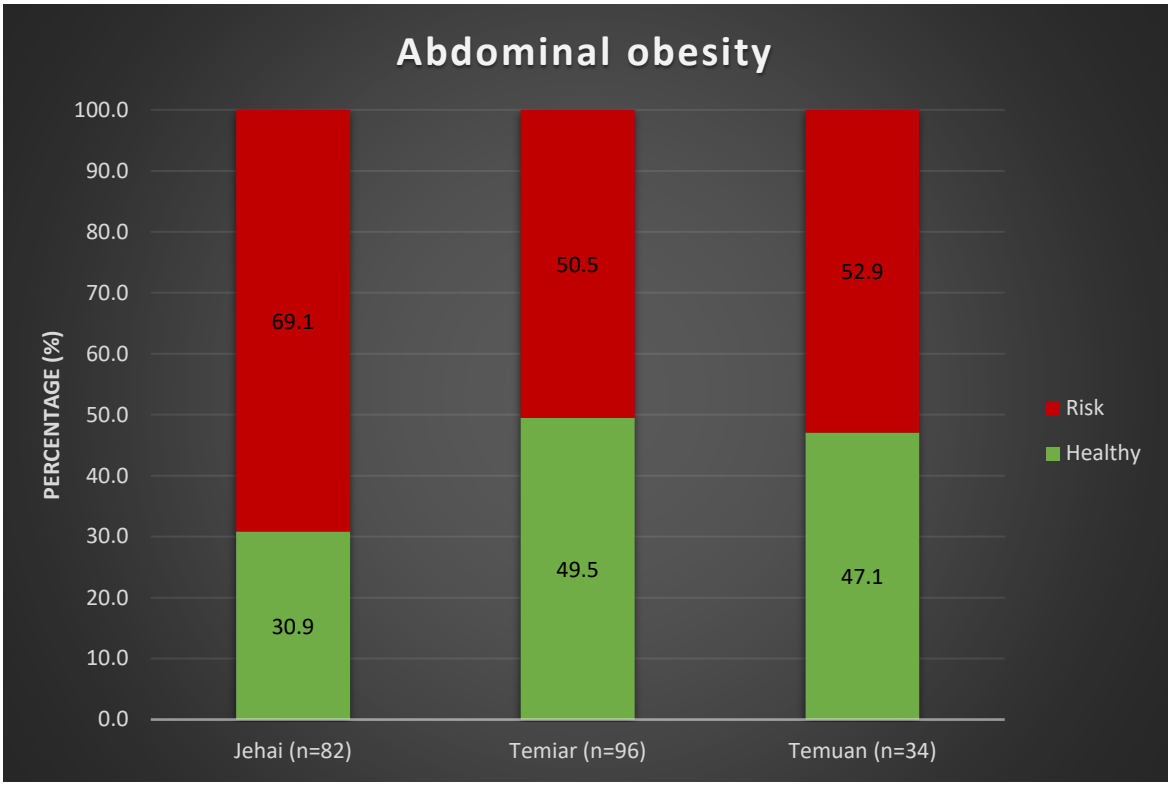


*Figure 3.20 More than half of the OA suffered from abdominal obesity. *Sample size insufficient for statistics.*

3.6.2 Blood pressure

A third of the OA from each community were normotensive while another third were pre-hypertensive. Jehai had the largest proportion of individuals suffering from hypertension. Even more strikingly, Jehai was the only group to have individuals (2.4%) presenting with stage 3 hypertension (Figure 3.21). Urban Temuans had the second highest proportion of individuals with stage 1 hypertension and above, whereas 20.6% of Temuan had stage 1 hypertension and 2.9% had stage 2 hypertension. Temiar had the highest percentage of individuals with pre-hypertension (46.4%), while 10.3% had stage 1 hypertension and 4.1% had stage 2 hypertension.

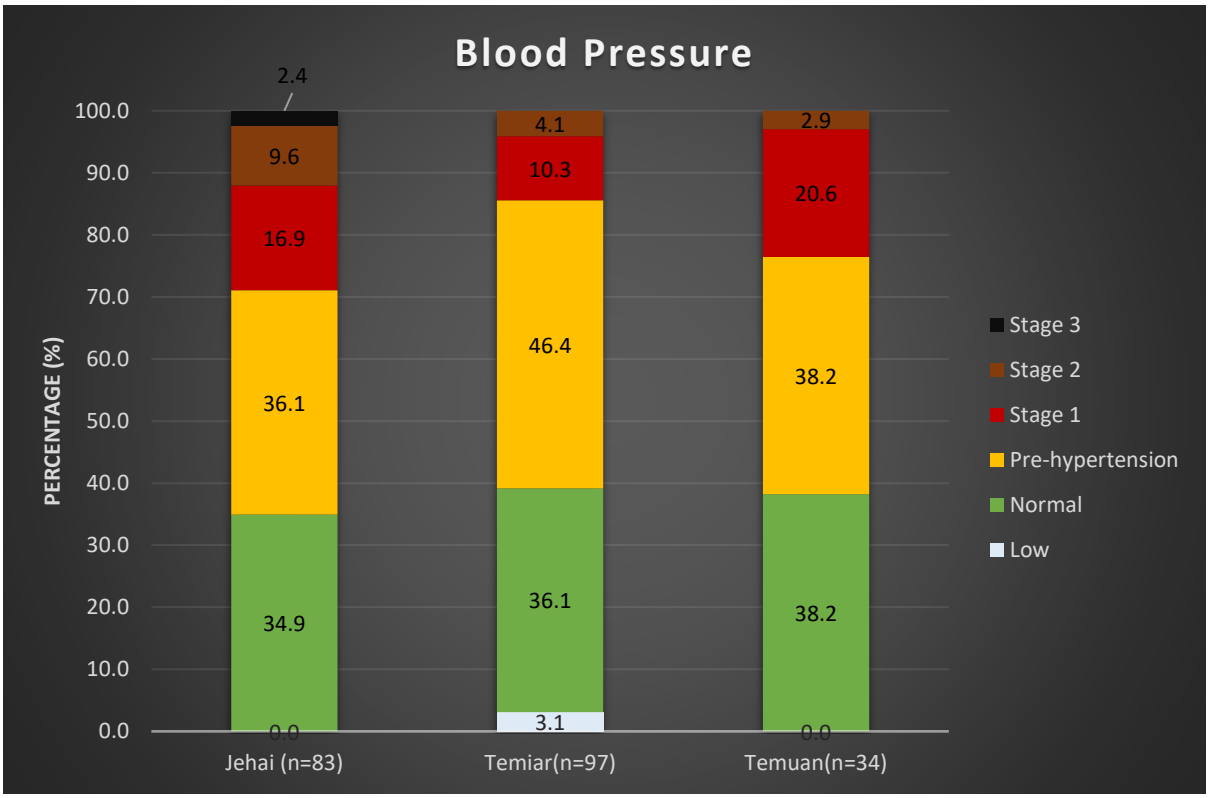


*Figure 3.21 Jehai had the largest proportion of hypertensive individuals and the only group with stage 3 hypertensive individuals. *Sample size insufficient for statistics.*

3.6.3 Type 2 Diabetes(T2D)

HbA1c which refers to glycated haemoglobin (haemoglobin with a sugar molecule attached) measures average blood sugar over 2-3 months and is used as an indicator of T2D (Sherwani et al., 2016). Data was collected only for the Jehai and Temiar groups. We engaged the Temuan community in collaboration with the National Diabetes Institute (NADI) who were conducting a longitudinal health project on diabetes education and prevention. Due to limitations with working with this external agency, we could not perform this test among the Temuan.

The Jehai reported 87.8% individuals with normal HbA1c levels, 7.3% were pre-diabetic and 4.9% (n=4) were diabetic. One diabetic Jehai was aware of his condition and mentioned that he was on medication whilst waiting for a finger prick test (personal communication between the nurse, Jehai man and myself). T2D was relatively low among the Jehai. However, half of the Temiar were pre-diabetic and 6.8% diabetic. Overall, T2D was relatively low in both OA communities.

Insulin resistance (IR) was predicted using a surrogate biomarker TG/HDL ratio. The surrogate biomarker TG/HDL ratio was a suitable predictor of IR as it was found to be strongly associated with IR identified through HOMA-IR (Homeostatic Model Assessment of Insulin Resistance) (Yeh et al., 2019). The prevalence of IR among the OA communities was calculated to be 25%. When segregated by OA group, presumptive IR prevalence among Jehai was 28.4% while Temiar was 21.9% (Figure 3.22). When investigating IR prevalence by gender, 28% men and 15% women presumably had IR.

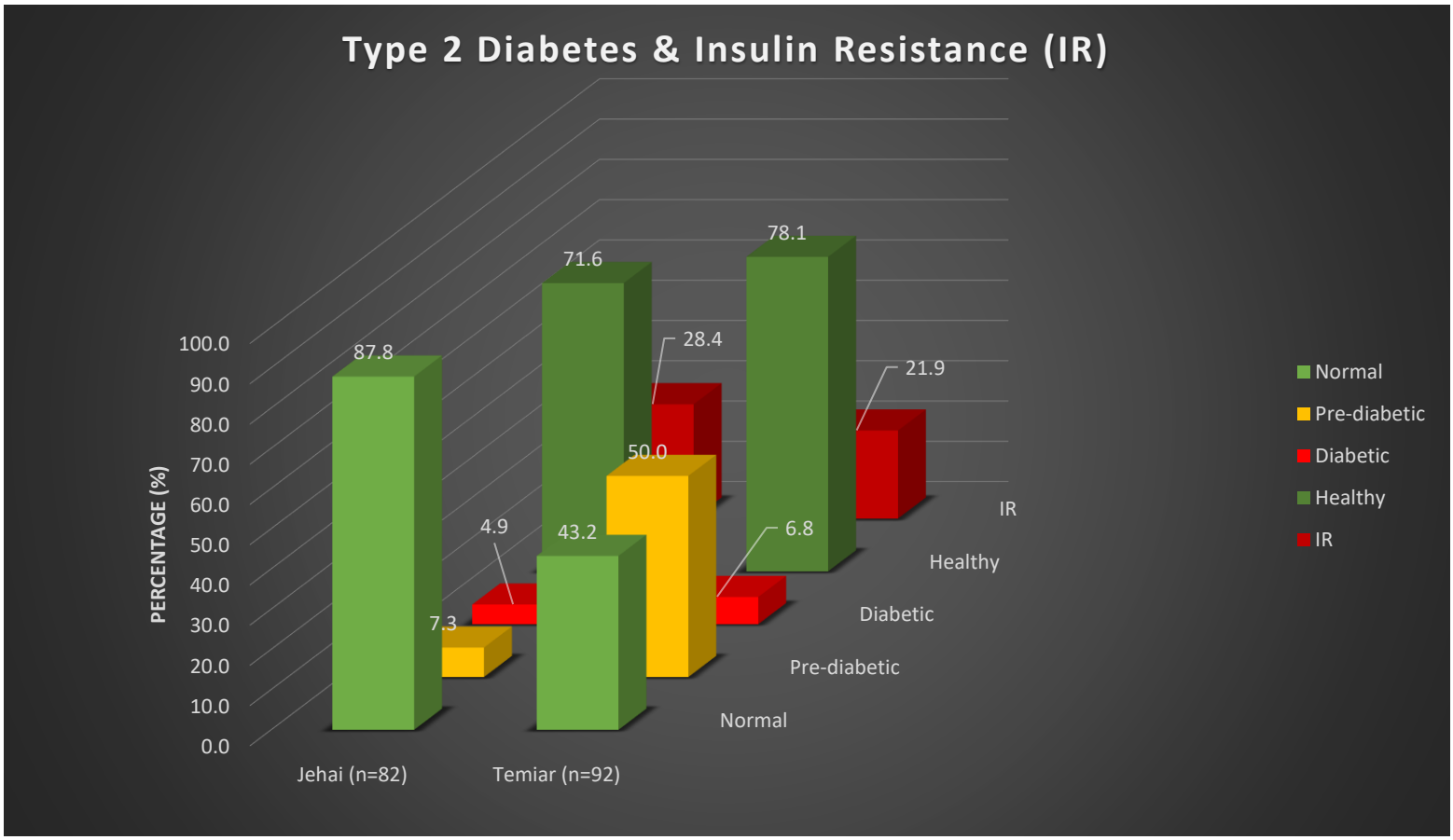


*Figure 3.22 Most of the Jehai had normal blood glucose level, whereas more than half of the Temiar were pre-diabetic. Overall diabetes was still relatively low among the two OA groups. *Data not available for Temuan. **Sample size insufficient for statistics*

3.6.4 Lipoprotein ratio

High LDL levels have been suggested to be atherogenic, while high HDL have been proposed to be cardioprotective. Hence LDL/HDL ratio was used to estimate the overall cardiovascular risk (Escobar et al., 2018). Individuals with increased TG tended to present with higher TC (Total Cholesterol) levels (Lamarche et al., 1996; Lemieux et al., 2001; Mathews et al., 2012; Quispe et al., 2020). Since LDL/HDL ratio alone may underestimate dyslipidaemia, TC/HDL was included to assess atherogenic risk (Lemieux et al., 2001; Quispe et al., 2020).

The total OA cardiovascular risk using TC/HDL ratio and LDL/HDL ratio was comparable at 26.5% and 24%. The Jehai appeared to have double the prevalence of dyslipidaemia and coronary heart risk compared to Temiar. The lipoprotein ratios, TC/HDL and LDL/HDL were 38.27% and 35.80% for Jehai and 16.67% and 15.63% for Temiar respectively (Figure 3.23).

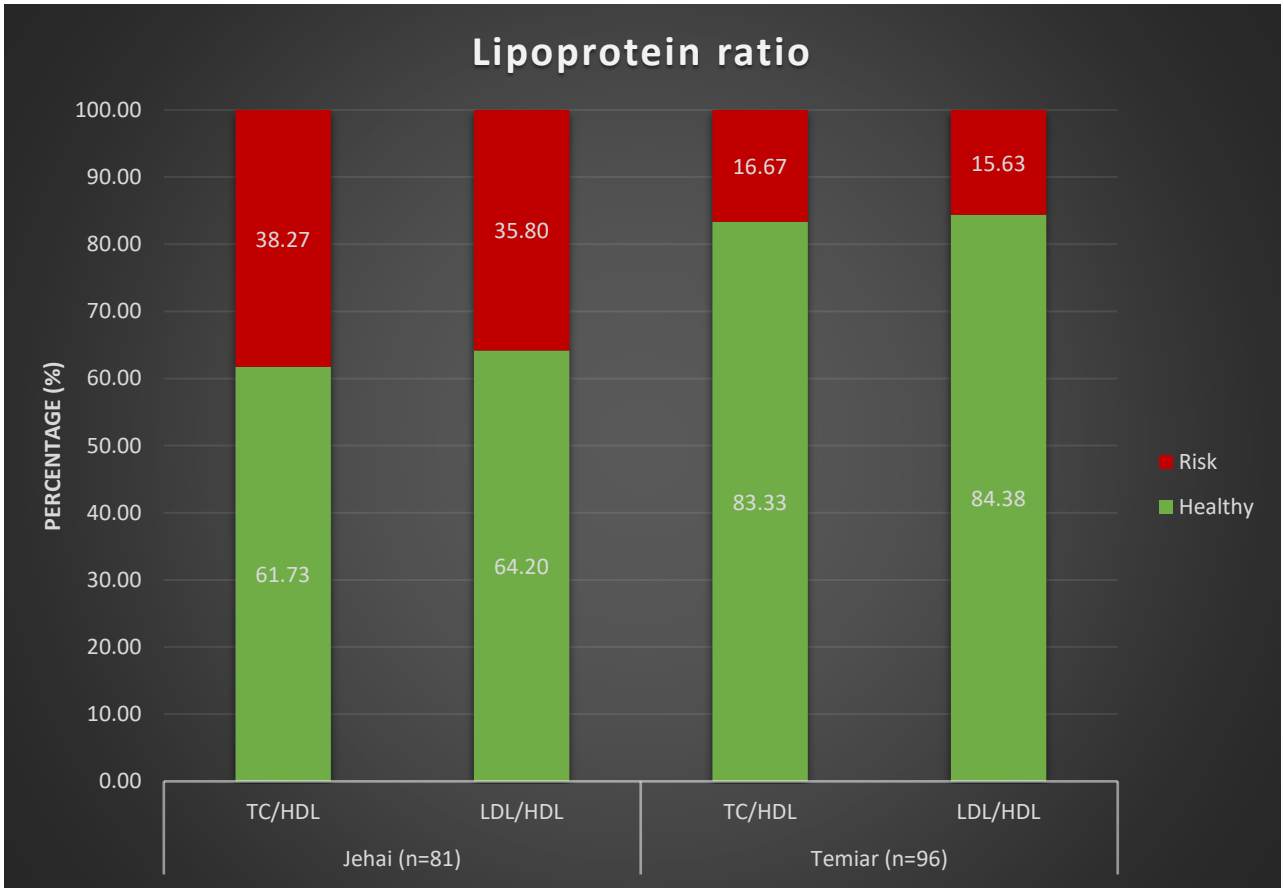


Figure 3.23 Jehai had double risk prevalence for cardiovascular disease compared to Temiar.

3.6.5 Smoking

Smoking was recorded as a variable in the effort to investigate OA oral microbiomes. It appeared that most of the OA were non-smokers. It should be noted that this was self-reported data. Temiar recorded the highest percentage with 37.8% active smokers and 11.2% former smokers (Figure 3.24). Jehai had the second highest percentage of smokers at 30.3% and 5.3% former smokers. Temuan recorded the lowest percentage with 20.6% active smokers and 14.7% former smokers. More men identified as smokers than women (Table 5).

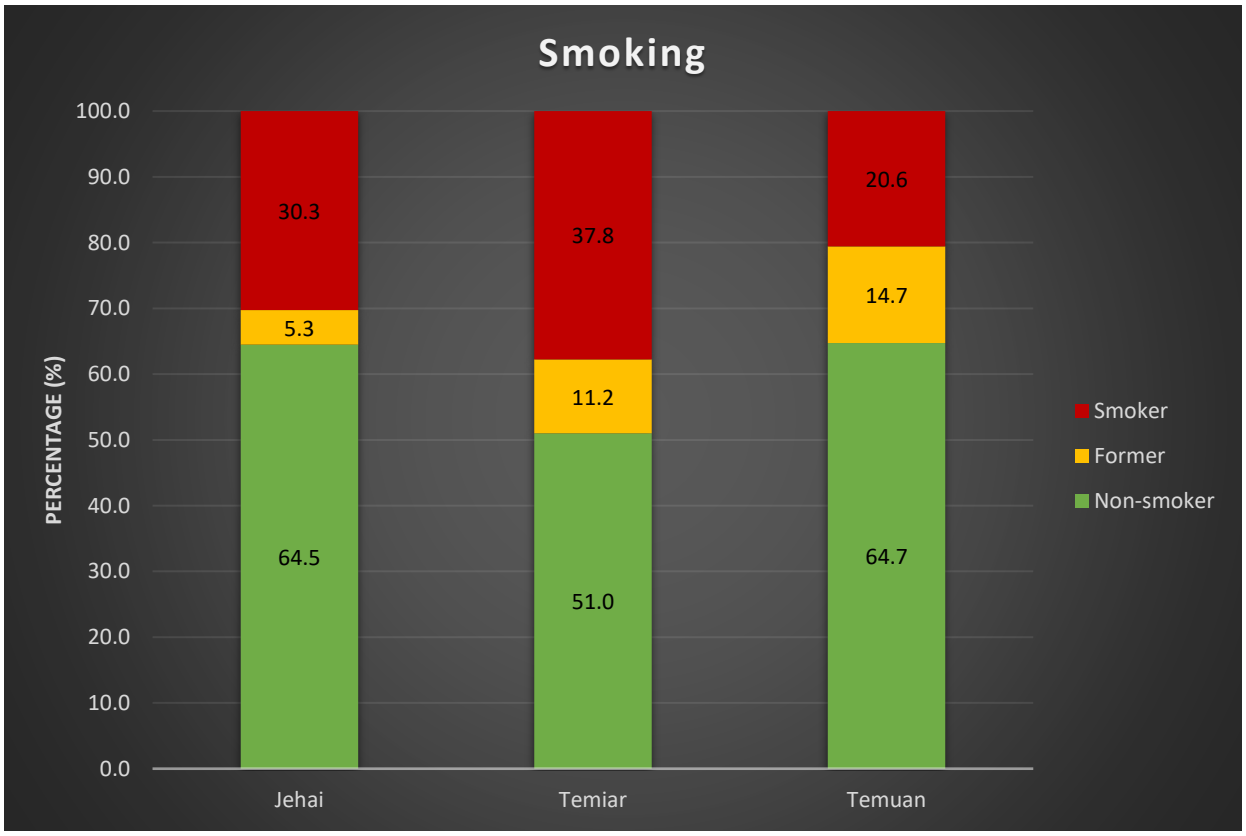


Figure 3.24 Roughly one third of the OA self-reported as smokers, of which a higher proportion were men.

3.6.6 Framingham Risk Score (FRS) and Metabolic Syndrome (MetS)

The Framingham Risk Score (FRS) is a simplified tool used to calculate risk for developing cardiovascular disease, usually over the next ten years. FRS considers risk factors including gender, age, total cholesterol, HDL, systolic blood pressure, smoking and diabetes. FRS for this study was calculated using a recalibrated calculator for an indigenous Australian cohort over the next five years (Hua et al., 2017). This recalibrated calculation may be more suitable for the OA communities than the conventional calculator used by Western and mainly Caucasian cohorts.

FRS was calculated for Jehai and Temiar communities. The Jehai and Temiar had comparable risk scores of 4.06 ± CI 1.17 and 4.59 ± CI 1.30, respectively. There was insufficient data for Temuan to proceed with the calculations.

The overall FRS score for the OA communities studied was 4.35% ± CI 0.88. The OA men had a risk score of 5.94% ± CI 1.69, while the OA women had a 3.37% ± CI 0.93.

Metabolic syndrome (MetS) was diagnosed if three or more of the five risk factors were present (waist circumference >90cm for men, >80cm for women, diabetes, low HDL, high TG levels and pre-hypertension or above). Overall, 44.63% of the OA were observed to have MetS, of which 36.76% were men and 49.54% were women. Temiar were seen to have a higher prevalence (54.17%) of MetS than the Jehai (33.33%).

Table 5 Anthropometrics of OA as a whole and segregated by gender and OA group.

	Category	Asian Criteria	Total (%)	Male (%)	Female (%)	Jehai (%)	Temiar (%)	Temuan (%)
BMI	Underweight	<18.5	11.21	16.25	8.21	15.66	8.25	8.82
	Normal	18.5-22.9	32.71	35.00	31.34	42.17	28.87	20.59
	Overweight	23-29.9	25.70	26.25	25.37	16.87	42.27	0.00
	Obese	≥30	30.37	22.50	35.07	25.30	20.62	70.59
Waist CM	Healthy	<90cm (M), <80cm (F)	41.98	64.56	28.57	30.86	49.48	47.06
	Risk	>90cm (M), >80cm (F)	58.02	35.44	71.43	69.14	50.52	52.94
HbA1c (%)	Normal	4.0-5.6	64.37	53.73	71.03	87.80	43.24	
	Pre-diabetes	5.7-6.4	28.74	40.30	21.50	7.32	50.00	
	Diabetes	≥6.5	6.90	5.97	7.48	4.88	6.76	
IR (TG/ HDL)	Healthy		74.60	71.88	84.97	71.60	78.13	
	IR	0.9-1.7	25.40	28.13	15.03	28.40	21.88	
Blood Pressure	Low	<90/60 mmHg	1.40	0.00	2.22	0.00	3.06	0.00
	Normal	<120/80 mmHg	35.98	27.85	40.74	42.17	31.63	33.33
	Pre-hypertension	<130/80 mmHg	41.12	41.77	40.74	39.76	44.90	33.33
	Stage 1	<140/90 mmHg	14.49	21.52	10.37	9.64	15.31	24.24
	Stage 2	<180/90 mmHg	6.07	7.59	5.19	7.23	4.08	9.09
	Stage 3	>180/120 mmHg	0.93	1.27	0.74	1.20	1.02	0.00
TC/ HDL	Healthy	<5.0(M), <4.5 (F)	73.45	75.00	72.48	61.73	83.33	
	Risk	>5.0 (M), >4.5 (F)	26.55	25.00	27.52	38.27	16.67	
LDL/HDL	Healthy	<3.5 (M), <3.0 (F)	75.14	78.13	82.66	64.20	84.38	
	Risk	>3.5 (M), >3.0 (F)	24.86	21.88	17.34	35.80	15.63	
Smoking	Non-smoker		58.17	23.08	79.23	76.92	40.63	64.71
	Former		9.62	16.67	5.38	2.56	13.54	14.71
	Smoker		32.21	60.26	15.38	20.51	45.83	20.59
FRS	5-year risk ± 95% CI		4.35 ± 0.88	5.94 ± 1.69	3.37 ± 0.93	4.06 ± 1.17	4.59 ± 1.30	
MetS	Yes		44.63	36.76	49.54	33.33	54.17	
	No		55.37	63.24	50.46	66.67	45.83	

3.7 Antibiotic resistance and the gut resistome

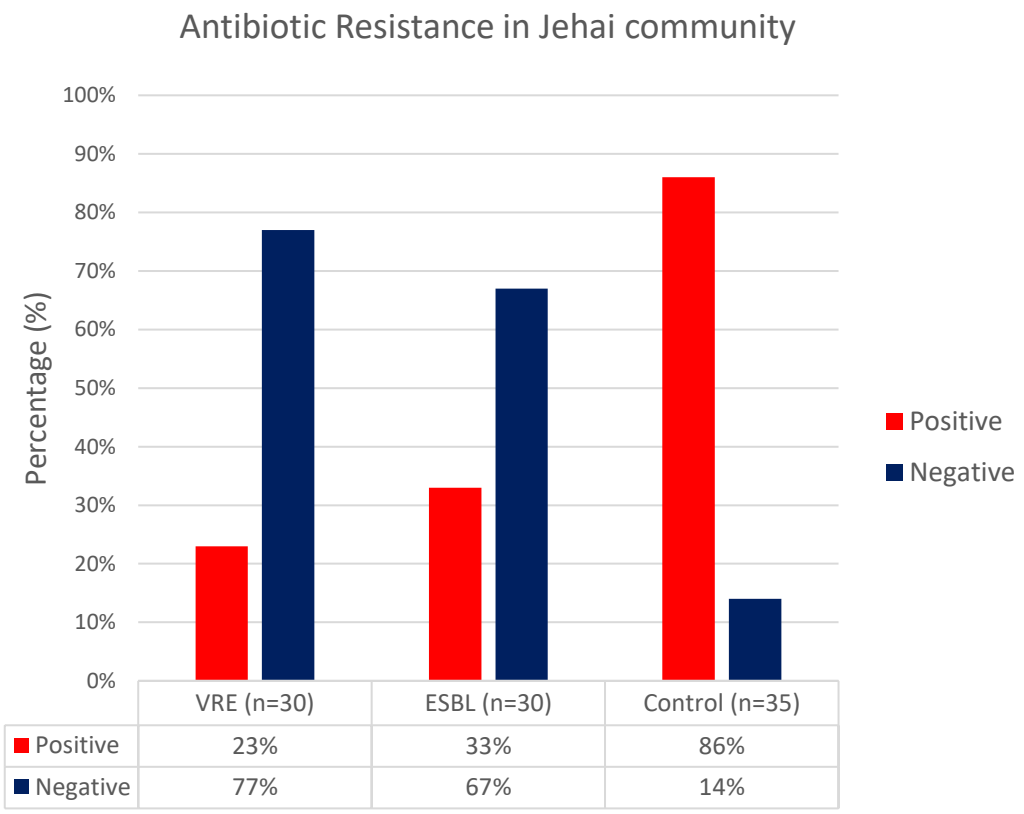
3.7.1 Intestinal carriage of VRE and ESBL in semi-nomadic Jehai community

Frozen samples were revived in Tryptic Soy Broth for two days before being plated on MacConkey (MAC) agar to assess the viability of stool bacteria. The revival rate was 86% (n=30), of which 14% (n=5) failed to form any colonies, thus excluded from the study.

VRE (Vancomycin-Resistant Enterococci) presented as purple/indigo coloured colonies on agar in 23% of Jehai stool (Figure 3.25). The purple/indigo coloured colonies indicated *E. faecium.* Light blue colonies characteristic of *E. faecalis* were not detected. ESBL (Extended-spectrum Beta-Lactamase) *E.coli* mostly presented as pink/blue colonies in 33% of Jehai stool as described in the manufacturer's manual.

The most common bacteria isolated were purple *E. faecium* on VRE agar. Conversely, pink/blue *E.coli* were the most common colonies that grew on ESBL agar, with the exception of two samples. One sample had colourless colonies growing on ESBL agar, indicating *Salmonella* or *Acinetobacter*. Interestingly, one other sample presented with a mixture of pink/blue *E.coli,* green (suggesting either *Klebsiella, Enterobacter, Serratia* or *Citrobacter)* and brown colonies (suggesting either *Proteus, Morganella* or *Providencia)*. In other words, all the ESBL that could grow on this ESBL chromogenic agar were present in this individual's gut.

Coloured pigmentation on the agar, in the absence of formed colonies, was often observed in many of the agar plates.



	VRE (n=30)	ESBL (n=30)	Control (n=35)
Positive	23%	33%	86%
Negative	77%	67%	14%

Figure 3.25 Intestinal carriage of VRE in Jehai was 23%, ESBL was 33%. Control indicated that 86% of the frozen samples were viable.

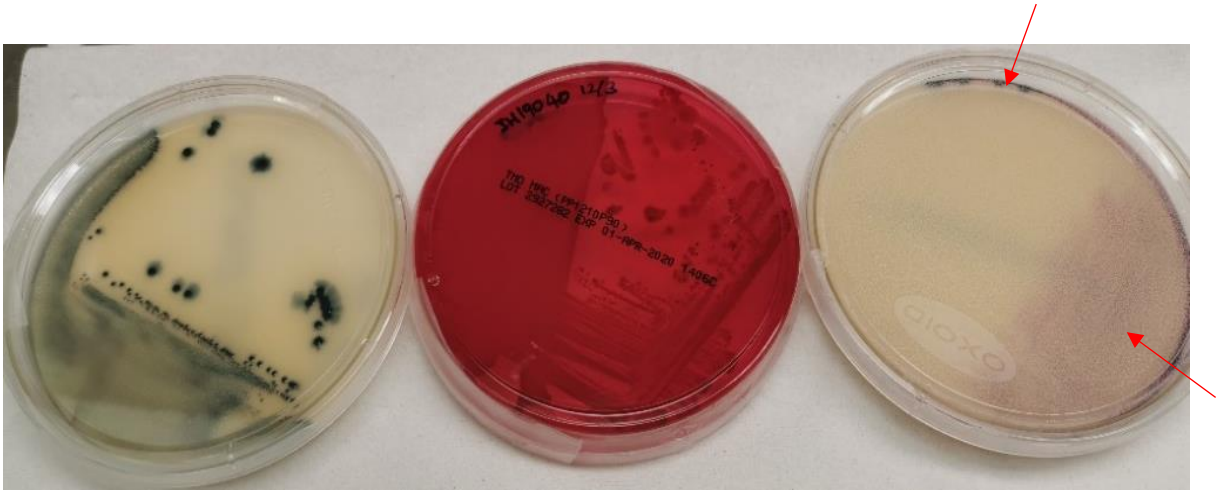


Figure 3.26 .(Left to right) ESBL positive with blue colonies (E.coli); MAC acts as a control for viability; VRE positive with clumps of purple colonies (E.faecium) growing at the edge of the plate. Note purple colour change on rightmost agar, suspected due to interactions between residual enzymes and chromogen on the agar but without formation of colonies.

3.7.2 Gut resistome of Jehai

Due to the uniqueness of these Jehai samples and the fact that this was done for the first time, the metagenomes were sequenced up to 100x. This far exceeded the 30x sequence depth recommended for most studies. The accuracy of ABR gene prediction increases with sequencing depth (Doyle et al., 2020). Jehai metagenomes seemed to be different. For example, sample J38 (file size at 9MB) was found to harbour 14 ABR genes whereas J31 (file size at 5.25 MB) harboured 44 ABR genes with sequence identity >90% to the ABR gene sequences in CARD database.

Half of the samples (n=6) had less than 15 ABR genes (range from 10 to 14 ABR genes per sample), while the other half (n=6) identified 43 ABR genes and more (range from 43 to 71 ABR genes). It is unlikely that increasing sequencing depth beyond 100x would result in an increase in ABR gene detection.

ABR genes detected in Jehai gut

A few ABR genes found in >80% of the samples included (*rpoB* – resistance to rifampicin, *CfxA6* – resistance to beta-lactamases, *dfrF* – dihydrofolate reductase found in *Streptococcus pyogens* and tetracycline resistance genes *tetM*, *Q* , *W*). ABR genes previously isolated from *E.coli* such as *acrA, ampC, ampC1-beta lactamase, ampH, emrE* and *mdfA* were mostly identified in samples that harboured high abundance of ABR genes. These *E.coli* ABR genes were mostly absent in samples with <14 ABR genes.

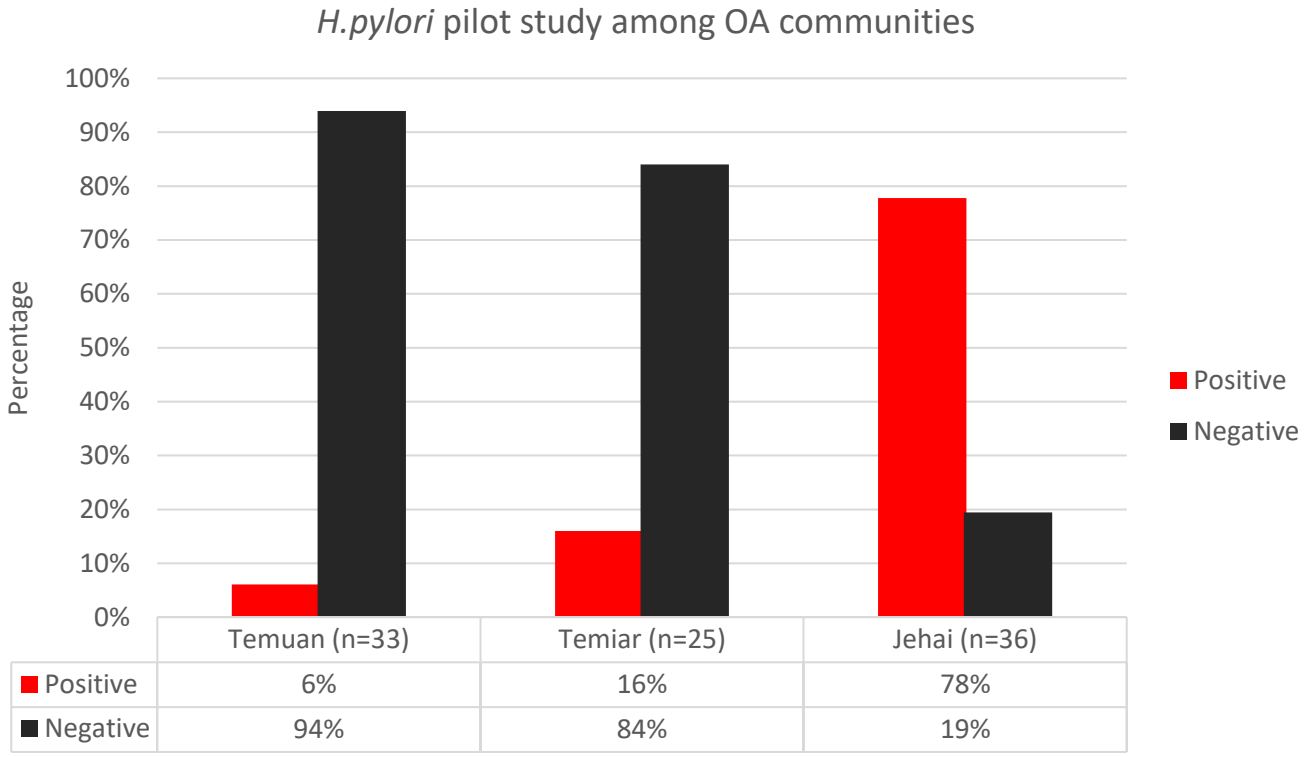
Klebsiella pneumoniae ABR genes (KpnE, KpnF, KpnG, OmpK37, acrA) were scarcely identified at <16% prevalence in Jehai gut metagenome. The *emr* gene family (*emrA, emrB, emrK* and *emrR*) mostly encode for membrane proteins and efflux pumps. The *emr* gene family was observed in samples with high abundance of ABR genes and matched the CARD reference database at 100% sequence identity.

The *mdt* genes (*mdt A, B, C, E, F, G, H, M ,N, O, P*) are scattered across the chromosome and are part of the major facilitator superfamily (MFS) that encodes for various solute transporters and efflux pumps (Hudson et al., 2014). Just like the *emr* gene family, mdt genes were mostly observed in samples that had high abundance ABR genes (>44 ABR genes). A number of *tet* genes (*tet 32, M, Q, W*) were found in >75% of the samples. The *tet* genes were usually present on transposable DNA elements thus facilitating horizontal transfer between gut bacteria as well as possible spread among humans (Lee et al., 2020), especially in a close-knit community with common food and water sources such as the Jehai.

A complete table of the ABR genes detected can be found in Appendix II.

3.8 *Helicobacter pylori* preliminary screening

The semi-nomadic Jehai had the highest prevalence of *H. pylori* infection, with 78% testing positive. The sub-urban TemiarGM had a prevalence of 16%, followed by the urban Temuan with the lowest prevalence of 6% (Figure 3.27).



	Temuan (n=33)	Temiar (n=25)	Jehai (n=36)
Positive	6%	16%	78%
Negative	94%	84%	19%

Figure 3.27 Helicobacter pylori infection among three OA communities. Jehai community had the highest prevalence of positive infections.

Chapter 4 Discussion

4.1 Targeted 16S rRNA microbiome findings

4.1.1 Temiar and TemiarGM microbiomes

The Temiar forms one of the largest OA communities out of the 18 sub-groups, with villages found mostly in the north and northeast part of Peninsular Malaysia. Temiar, Pos Piah participants were recruited from a relatively isolated resettlement village which was an hour and a half car ride away from the nearest town, Sg. Siput in Perak. On the other hand, TemiarGM participants were recruited from resettlement villages in Kelantan that was 30mins away from the nearest town and was surrounded by a higher frequency of land development and deforestation. The difference in distance to town may affect accessibility to store-bought food and plausible differences in diet. Recent studies have suggested that diet and degree of urbanization have a distinct impact on the microbiome (Lokmer et al., 2020, De Angelis et al., 2020). Perhaps differences in diet and degree of urbanization resulted in the microbiomes of Temiar appearing relatively more similar to Jehai microbiomes. Meanwhile, TemiarGM saliva and gut microbiomes seemed to resemble urban Temuans' microbiomes more. This observation may be further investigated with larger sample size and quantitative measurement of differences in diet.

4.1.2 Saliva microbiomes

Saliva microbial constituents and diversity

There were a few uncultured bacteria identified as core constituents of the saliva microbiome. For now, definitive investigations and further characterization would require more suitable techniques such as shotgun metagenomics. The overall alpha diversity of OA saliva microbiomes was comparable across all communities. Jehai saliva microbiomes were significantly higher in microbial richness and evenness, respectively, when compared to Temiar and Temuan, but Pielou's evenness measure was not statistically significant. Interestingly, Temiar scored the lowest in within-group microbial diversity using Pielou's evenness. Long term betel quid chewing has been associated with lower alpha diversity (Hernandez et al., 2017). This community of Temiar has been known to chew betel quid (Yeo PgDip Thesis, 2017, see Appendix II, Supp Table 1) that may have contributed to low alpha diversity. Temuan and Jehai are also known to chew betel quid (personal observation).

Low alpha diversity has also been reported among rural saliva microbiomes from Sierra Leone and DRC Congo inhabitants, whereas higher alpha diversity was reported among Germans and Batwa pygmies (Li et al., 2014). The authors attributed high alpha diversity among Germans to them having access to a wider variety of food sources and a higher frequency of microbe exchange from living in more dense population areas. On the other hand, the Central African populations, including Sierra Leone and DRC Congo, were farmers with relatively less diverse food choices. They lived in sparsely populated rural areas (Li et al., 2014). Another study showed that a shared environment resulted in more similar saliva microbiomes instead of genetics-relatedness (Shaw et al., 2017).

The oral microbiome of Tibetans reflected an alpha diversity that decreased with altitude, where higher altitudes usually have a less dense population and limited food type (Liu et al., 2021a). The combination of betel quid chewing, less diverse food source and low-density population may have resulted in a more homogenous microbial community and lower alpha diversity in the saliva microbiomes of OA in this study. Interestingly, Batwa pygmies, who were former hunter-gatherers, lived in rural areas yet reportedly had high alpha diversity (heterogenous) oral microbiome,

comparable to Germans, which the authors attributed to diverse food sources available to the Batwa pygmies (Li et al., 2014). The Jehai, who are traditionally hunter-gatherers, had saliva microbiomes highly similar in alpha diversity to urban Temuans. This study, supported by past research, seems to suggest that two ends of the spectrum of urbanization, namely urban dwellers and semi-nomadic hunter-gatherers living by the forest, were characterized by higher alpha diversity in their saliva microbiomes.

On the other hand, communities in transition, mostly living in the suburbs and rural areas, appear to be characterized by low alpha diversity, possibly attributed to a less diverse variety of food available to the communities. These findings are very much on the contrary to gut microbiomes of rural populations. It is very interesting because this denotes that perhaps the oral microbiome reacts quite differently from the gut microbiome. This once again illustrates the importance of including rural, traditional populations to understand the whole picture of the human microbiome.

Ordination plots showed that rural OA saliva microbiomes, Jehai and Temiar, were more similar and distinct from urban and sub-urban OA microbiomes, Temuan and TemiarGM. Further analysis using ALDEX2 revealed three genera, *Corynebacterium, Prevotella* and *Mogibacterium*[7]*,* to be increased in urban-like OA saliva microbiomes, Temuan and TemiarGM. *Corynebacterium* is a gram positive, facultative anaerobe and a commensal in the oral cavity. *Corynebacterium* has more than 110 species (Oliveira et al., 2017), with some pathogenic species such as *C. diphtheriae* and some commensal, but opportunistic species such as *C. durum* (Treerat et al., 2020) *and C. matruchotii* (Esberg et al., 2020). *Prevotella* is also a commensal oral bacteria that have been studied in association with gout (Liu et al., 2018), oral cancer (Zhang et al., 2020) and tonsillitis (Yeoh et al., 2019). On the other hand, *Mogibacterium* is a strict anaerobe found in higher abundance in periodontitis and caries (Chen et al., 2018, Chen et al., 2020).

It is premature to draw any conclusion because there are various species under each of these genera. They could be pathogenic, commensals or opportunistic, of which short-read 16S rRNA used in this study cannot discern. However, these bacteria have been associated with oral diseases and are found elevated in urban-like saliva microbiomes of Temuan and TemiarGM. This association may suggest that the more urbanised OA saliva microbiomes are at dysbiosis. Further studies may be warranted to investigate oral health among rural and urban OA communities.

ALDEX2 cut-off point

The cut-off values chosen for ALDEX2 testing for differentially abundant bacteria genera were set quite stringently with Bonferroni-Hochberg corrected p-value < 0.01 (default of ALDEX2 was 0.1) and effect size >1 as suggested by the authors of ALDEX2 (Fernandes et al., 2014).

Effect size was suggested as a more robust and reproducible measure than p-value in determining significantly different taxa between query groups (Lovell, 2013, Fernandes et al., 2018). P-values are susceptible to large sample sizes. The larger the sample size, the higher the possibility of obtaining a small p-value by chance, thus falsely rejecting the null value. A more stringent statistical threshold was selected in this study so that differentially abundant taxa were highly possible to be true, biological differences observed between two groups.

[7] *Mogibacterium* was the only bacteria not detected in the negative controls. The rest were detected at very low levels, <1.5%.

Sphingobacteriia

It was a purely coincidental finding where Sphingobacteriia was observed in Jehai and Temiar saliva microbiomes in the taxonomy barplot at the class taxonomy level (Figure 3.1), but not in the more urbanised microbiomes of Temuan and TemiarGM. It was coincidentally coloured black so its absence in Temuan and TemiarGM samples became quite obvious. Sphingobacteriia is currently identified only to the Family level, Lentimicrobiaceae. Upon further investigation, Sphingobacteriia may not be absent but were actually not detected[8] in Temuan and TemiarGM saliva microbiomes. Since the Temuan and TemiarGM samples were extracted and sequenced using uBiome's method, it is difficult to determine if Sphingobacteriia was not detected because they were not present in OA saliva or simply not extracted. They were also not detected in samples re-sequenced using our in-house methods. Sphingobacteriia is a relatively new class of bacteria found in environmental and marine samples. It has also been found in human guts as a class of bacteria that produces sphingolipids, hence its name Sphingobacteriia (Feng et al., 2019). Collecting and sequencing more OA samples would be needed to discern this finding. In the meantime, more information will be required to discover the role of Sphingobacteriia in the human microbiome.

4.1.3 Gut microbiomes

Gut microbial constituents and diversity

Jehai had the highest alpha diversity, whereas the urban Temuans scored the lowest. Alpha diversity was reportedly higher among traditional populations such as Tanzania's Hadza hunter-gatherers (Hansen et al., 2019), Nepal's Chepang foragers (Jha et al., 2018), traditional Matsés hunter-gatherers and Tunapuco's rural agriculturist population in Peru (Obregon-Tito et al., 2015). The intermediary microbiomes observed in Temiar and TemiarGM communities were also reported in Nepal's traditional, semi-urbanized agriculturist population (Jha et al., 2018). An increasingly higher alpha diversity reported among the gut microbiomes of more rural populations is quite different from the trend observed in saliva microbiomes. The saliva microbiomes of rural, possibly agriculturalist populations seem to exhibit lower alpha diversity, while hunter-gatherer and urban communities appear to have higher saliva alpha diversity.

From the taxonomy bar plot (Figure 3.6), bacteria class Betaproteobacteria appeared less in semi-nomadic Jehai and more in urban Temuan gut microbiomes. Rare and low abundant taxa are interesting because changes in diet, environment, and seasons seem to be reflected more readily in lower abundant taxa than higher abundant taxa (Davenport et al., 2014, Zhu et al., 2021). Upon further investigation using ALDEX2, the only genus that tested significantly different in Jehai and Temuan gut microbiomes from the class Betaproteobacteria was *Sutterella*. It should be noted that the taxonomy bar plot was used merely as a preliminary, visual guide to look for possible differences in gut microbiomes among the OA communities that would warrant deeper investigation. It may be premature to deduce biological significance at any level above species. To illustrate this point, possibly over 70 different bacteria genera were condensed in the brown coloured block in the taxonomy bar plot representing Betaproteobacteria. Only one genus tested significantly different, which is *Sutterella*. With this observation and current limitations in knowledge about the microbiota, it may be more

[8] Bacteria that were absent would appear as "zero" in the abundance table, but Sphingobacteriia did not appear at all in TemiarGM and Temuan samples.

informative and accurate to analyse bacteria abundance level at the finest level possible, which would be the genus level for short-read 16S rRNA studies. Analysis at any higher level, such as Family or Class, may be too broad and should be used as a guide instead of a definitive claim of differential abundance.

With that being said, *Sutterella* was found to be reduced in the Jehai guts. However, it was reportedly elevated in the gut microbiomes of BaAka hunter-gatherers compared to Bantu agriculturalists who had a more Western-like microbiome (Gomez et al., 2016). Another study similarly reported a higher abundance of Sutterella among a transitioning rural South African population compared to urban South Africans (Oduaran et al., 2020). The findings of my study are contradictory to previous findings. However, *Sutterella* was found elevated in relation to obesity in that rural cohort (Oduaran et al., 2020). Furthermore, *Sutterella* was also elevated in gut microbiomes of Burkina Faso children living in urban areas compared to their counterparts in more rural areas (De Filippo et al., 2017). Conversely, these studies are in line with my study where *Sutterella* is more abundant among urban Temuans who coincidentally have the highest prevalence of obesity.

The tree of life illustrates that *Sutterella* has actually got many uncultured species (de Vienne, 2016, Figure 4.1) and that *Sutterella's* role in gastrointestinal diseases is still highly conflicted (Kaakoush, 2020). Another genus commonly associated with traditional communities and healthy guts, *Prevotella*, did not appear significantly higher in Jehai gut microbiomes. All in all, these findings reiterate the point where we are still in the early stages of microbiome research, and more research will be required to explain seemingly contradictory discoveries.

Odoribacter, Blautia[9]*, Lachnoclostridium, Parabacteroides, Bacteroides* and Ruminococcaceae UCG-013 were found more abundant in urban Temuan gut microbiomes compared to Jehai. *Odoribacter* is an anaerobic gut commensal with its name meaning bad smell (Hardham et al., 2008). *Odoribacter* has been reportedly associated with IBD (Morgan et al., 2012) and Alzheimer's disease (Haran et al., 2019). *Odoribacter* has also been indicated in host-cell interactions, which affect inflammation (Hiippala et al., 2020).

Blautia was depleted in rural OA gut microbiomes, Jehai and Temiar. *Blautia* was also found to be decreased among the Hadza hunter-gatherers compared to urban Italians (Schnorr et al., 2014). *Blautia* and *Bacteroides* were found to increase with the degree of urbanization, which appears in line with my study (Yong et al., 2021). However, in an obesity study involving adult Japanese, *Blautia* was reportedly inversely associated with visceral fat (Ozato et al., 2019). *Blautia* was also found depleted in obese children, associated with metabolic inflammation, and possibly leading to insulin resistance (Benítez-Páez et al., 2020). This association of *Blautia* depleted in obese people contradicts my study because Jehai was the leanest and had the lowest prevalence of obesity compared to urban Temuan.

On the other hand, increased *Blautia* abundance was positively related to improved hyperglycaemia and lipid metabolism (Tong et al., 2018). My study has insufficient data and sample size to investigate this link. Nonetheless, there is an ostensible association between *Blautia* and health (Liu et al., 2021b). The seemingly conflicting associations suggest that investigation of *Blautia* needs to be advanced to the species level for further clarity.

[9] *Blautia* (2%)*, Prevotella 9* (6.4%) and *Prevotella 2* (1.04%) were detected in gut negative controls.

Bacteroides and their close relative *Parabacteroides* are major constituents of the human microbiome. Several *Bacteroides spp.* provide nutrients to the human host and other gut residents by metabolizing polysaccharides and oligosaccharides(Patrick, 2015). *Bacteroides* are a predominant gut microbe, especially for people with Western diets enriched in animal fat and protein (Wu et al., 2011). It is highly likely because *Bacteroides* can tolerate bile, often found in the guts of people who favour meat (Tomova et al., 2019). *Bacteroides* have also been found to be depleted in gut microbiomes of Hadza hunter-gatherers (Schnorr et al., 2014). Another study showed that a dietary-fibre deprived gut had more *Bacteroides* (Desai et al., 2016). These studies are in line with my study as Temiar and especially Jehai consumed a lot of plant-based food, including tapioca, a variety of ferns that grow by the forest and other edible shoots . However, *Bacteroides spp.* were reportedly prevalent among vegans too (Ferrocino et al., 2015, Zafar and Saier, 2021). It will be interesting to further elucidate the differences in gut microbiome compositions by examining the differences in OA diets.

There is more information now about the gut microbiota, such as *Blautia, Bacteroides* and *Sutterella*. Some bacteria are relatively new and less-studied such as *Lachnoclostridium*, *Odoribacter* and Ruminococcaceae UCG-013. However, caution must be exercised when making overly simplistic associative claims. Furthermore, relatively recent microbiome studies were chosen as comparisons to this study, when possible and appropriate. Downstream analytical tools have become more sophisticated and new measures are now more relevant and specific to microbiome data (such as reporting significance using effect size over p-value). Some genera observed[10] and previously reported to be significantly different among urban and rural cohorts (Schnorr et al., 2014, Zhang et al., 2014, Dehingia et al., 2015) were not significantly different in this study. It may be difficult to compare recent studies with studies analysed a while back that used methods suitable at that time. Those methods may have become inappropriate and possibly erroneous by today's standards. However, it is not reasonable to have researchers constantly re-analysing and re-submitting their work since new methods will keep coming out as the field evolves and improves. The findings will simply have to be analysed cautiously.

[10] Some genera that previously tested as significant when I analysed this OA data in 2019-2020 have mostly not tested significant in this latest analysis. It could be because I have incorporated a more stringent cut-off value to minimise false positives which could explain why many genera did not pass the significance threshold. Those genera include *Succinivibrio*, *Bifidobacterium, Dorea, Roseburia, Faecalibacterium.*

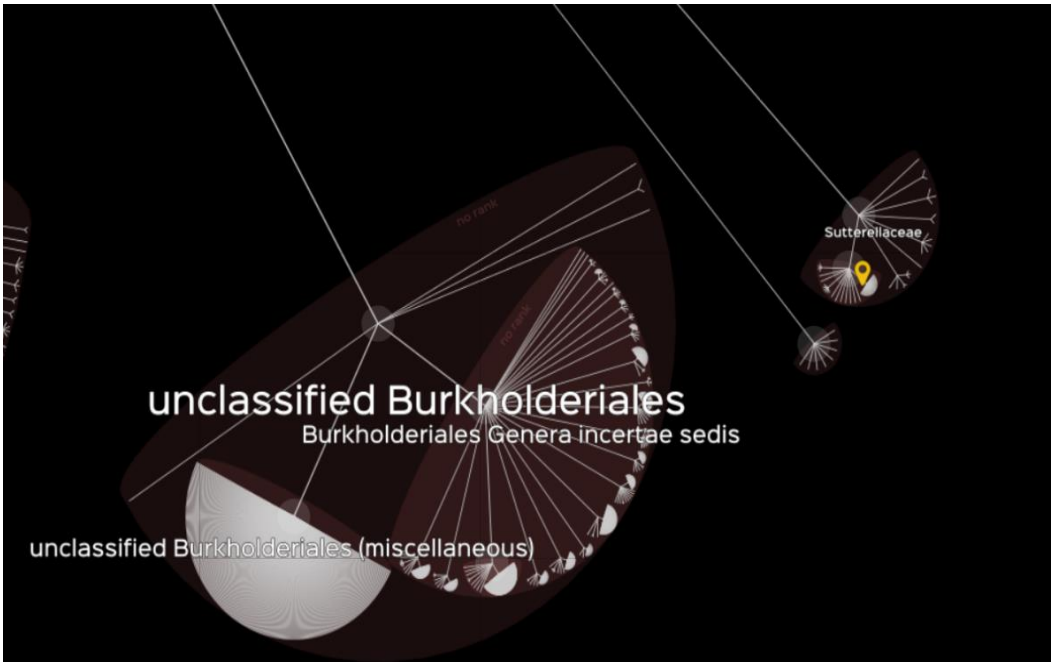


Figure 4.1 A partial screenshot of Betaproteobacteria Class, the yellow pinpoint is where Sutterella genus is located in the tree of life. This screenshot was taken from https://lifemap-ncbi.univ-lyon1.fr/ by Damien de Vienne / CNRS under a Creative Commons Attribution-NonCommercial 4.0 International License.

Microbiome controls

Currently, there is no consensus[11] on how to remove contamination found in the negative control. The PCR negative control had no bands when viewed on an agarose gel. However, sequencing revealed *Pseudomonas* (50.96%), *Streptococcus* (17.57%), *Neisseria*(14.70%), *Prevotella 9* (10.21%) and *Prevotella 7* (9.67%) to be among the top five contaminant genera. *Pseudomonas* can be found in the environment like soil, water, plant, food, and human skin commensal (Neves et al., 2014). Other possible sources of contamination in microbiome studies include, but not limited to, plastic consumables, extraction kits, PCR reagents and laboratory environment (Cosseau et al., 2016, Eisenhofer et al., 2019).

The two highest abundance contaminants were in the gut negative controls, namely *Blautia* at 2% and *Prevotella 9* at 6.43%. *Blautia* is an anaerobic bacteria widely studied in the human gastrointestinal tract and rarely from environmental samples (Liu et al., 2021b). The presence of *Blautia* in the negative samples may have come from cross-contamination during the handling of stool samples in the laboratory. On the other hand, *Prevotella* is a human commensal found in the oral and gut (Murray and Cassese, 2016). Thus, contamination may have come from laboratory personnel. The other contaminants, including *Campylobacter, Corynebacterium, Actinomyces, Peptostreptococcus* and *Alloprevotella,* were present at <1.5% abundance in the saliva negative controls. The abundance of these contaminants may be low, but it is still good to be aware of potential confounding factors affecting the OA microbiomes in this study.

Several times in the QIIME2 community forum, complete removal of contaminants were deemed too drastic. It may artificially skew the data since the contaminants could also be part of the actual microbiome data. Although there was not much downstream analysis that could be done to remove contaminants or correct for batch effect, this section was added to describe the controls as recommended by Hornung et al. (2019). This section acknowledges and informs the reader that there was some sequencing bias, possibly introduced at the extraction and sequencing stage. The OA microbiomes should be interpreted with the contaminants kept in mind.

For the samples that were sequenced on both V3-V4 and V4 regions, stool samples were more affected by differences in extraction protocol, 16S rRNA hypervariable region and sequencing platform. This was expected as the effect of pre-sequencing protocols on recovery of gut microbiome compositions have previously been discussed at length (Sinha et al., 2017, Videnska et al., 2019). On the other hand, saliva microbiomes were less affected by extraction, hypervariable region, and sequencing platforms (Lim et al., 2017, Rosenbaum et al., 2019), as was reflected in this study.

4.1.4 Mixed-region analysis: challenges and methods to overcome.

Combining microbiome datasets generated using different sequencing primers was extremely challenging. To date, there is no downstream method that can completely resolve the differences that arise from sequencing primer bias where certain taxa may be preferentially amplified according to the targeted hypervariable region used (i.e. V1-V2 vs V3-V4).

One of the available methods tried in this study was SEPP fragment insertion (currently only available as a qiime2 plugin). SEPP fragment insertion method claimed on the qiime2 manual

[11] https://forum.qiime2.org/t/filter-controls-or-not-to-filter-controls/1832/4

(https://library.qiime2.org/plugins/q2-fragment-insertion/16/) and in their journal paper to be able to combine mixed sequencing region data and allow for meta-analysis of microbiome data (Janssen et al., 2018). To the best of my knowledge, the following problem has not been reported in the original paper or the QIIME2 community forum.

In my study, the SEPP-generated phylogenetic tree and unweighted UniFrac metric, as recommended by the authors, performed better than using Euclidean distance metric, which does not consider a phylogenetic tree. There were two problems with the SEPP method. Firstly, there was an unexplainable separation horizontally across the ordination plots when using unweighted UniFrac. Reducing taxonomic resolution from the ASV level where bacteria taxa may be distinguished by single nucleotide level to genus level made the separation more profound.

To test this, I tried using data trimmed to the V4 region[12] and observed ordination plotted using unweighted UniFrac. This separation was also observed at the genus level. A possible explanation is perhaps unweighted UniFrac is not suited for microbiome compositional data analysis (Gloor et al., 2016). This may be why we see the apparently random separation in sample distribution. Not excluding other possible explanations, this is what I have observed for now.

From this small analysis, the SEPP method for generating a phylogenetic tree appears reliable based on its ability to resolve samples (compare Figure 3.12 and Figure 3.13), as demonstrated in their paper (Janssen et al., 2018). However, their claim of being able to facilitate mixed region analysis may require further refinement. There is a possibility SEPP method is only suitable for combining V4 and V2 region datasets and not suitable for other regions, as demonstrated by V3-V4 and V4 regions in this study.

Based on this study, I would suggest trimming the dataset to the same hypervariable region and analysing data at the same read length, if possible. Analysis can be done using clr-transformed Euclidean distance at the ASV level. However, differences introduced at the extraction stage currently should be minimised in the prior.

4.1.5 Limitations

One of the biggest limitations of my study was the samples were extracted and amplified using different protocols. Although the sequences were trimmed and analysed at the V4 region only, which was found to be quite comparable (Liu et al., 2020), extraction methods were shown by others (Costea et al., 2017, Sinha et al., 2017) and in the quality control section of this thesis to have an impact on the recovered microbial community. One of the inherent limitations of 16S rRNA microbiome studies is the low compatibility of data generated by different methods, making the comparison of any published 16S rRNA data very difficult. The reality is, more often than not, samples will be extracted and amplified using different protocols.

My doctoral study was initially supported through a research sponsorship grant from uBiome, USA that had provided the sample collection kits and sequencing services. Unexpectedly, the company ceased operations in the midst of my investigations. This was a serious setback that led us to complete the sequencing of Jehai samples using our in-house platform (V3-V4 sequencing on Illumina MiSeq). In another instance, Microbiome Quality Control Consortium recommended using zirconia beads over

[12] A reminder that the entire analysis was eventually completed using data trimmed to V4 region and dissimilarity measured using Euclidean distance metric.

glass beads for stool extraction (Sinha et al., 2017). Zirconia was reportedly better at recovering gram-positive bacteria DNA from stool samples but prohibitively costly given our limited resources. I made a conscious choice to use glass beads that were freely available in our lab instead. This was a deviation from some recommended protocols. Hence, until we can overcome the current limitations of 16S rRNA short reads technology, including sample extraction and processing, as well as downstream bioinformatics analysis, one may consider using 16S rRNA short reads as a pilot study. At the same time, further investigations can be conducted using shotgun metagenomics.

Another limitation of this study was that participant diets could not be accurately quantified within the time we had to complete the initial field investigations. This data would possibly have improved analysis of the microbiome results. Planned follow-up and detailed interviews with the participants had to be cancelled due to the Covid-19 pandemic and several national lockdowns. Had these additional engagements been possible, I would have been able to provide more specific information about each community and lead to a better understanding of the results.

4.2 Shotgun metagenomics

4.2.1 Core species found in Jehai gut metagenome.

A total of 73 core species were shared across twelve Jehai gut metagenomes, which is substantially higher than what has been reported for other populations. Most population samples contain between 30-40 species, and if comparisons are made between different populations, the core species shared may be as low as 20 species (Lozupone et al., 2012, Salonen et al., 2012, Martínez et al., 2013). This number is on the high side but can be explained as related individuals were from three families and two unrelated individuals.

Prevotella copri was found in high abundance in the Jehai core gut metagenome. This bacteria has been associated with host health, such as improved glucose metabolism (Kovatcheva-Datchary et al., 2015), rheumatoid arthritis(Stoll, 2020), and plant-based diets (Tomova et al., 2019). *P. copri* have also been over-represented in communities leading non-westernized lifestyles and generally depleted in westernized microbiomes (Tett et al., 2019). The Jehai consumed a largely plant-based diet consisting of forest ferns and herbs foraged from the forest and some vegetables that are easy to cultivate (Figure 4.2).

An interesting study found that *P.copri* strains associated with omnivore diet had a higher prevalence of *leuB* gene, which is a risk factor for insulin resistance and type 2 diabetes (De Filippis et al., 2019). Their study also found that *P. copri* pangenome of Western individuals had a higher prevalence of genes involved in antibiotic and toxic compound export from cells and membrane proteins that may be associated with drug resistance (De Filippis et al., 2019). *P. copri* strains from non-Western individuals were seemingly more adept at breaking down complex fibres (De Filippis et al., 2019). It would be pretty exciting to explore further the strains of *P. copri* present in the Jehai gut metagenomes.

F. prausnitzii is a bacteria of clinical interest as it has been reported in association with several inflammatory diseases. For example, reduced levels of *F. prausnitzii* have been reportedly associated with Crohn's disease and the development of metabolic syndrome like type 2 diabetes (Verhoog et al., 2019). Studies have reported a low abundance of *F. prausnitzii* associated with *Crohn's disease* and that oral administration of live *F. prausnitzii* in mice reduced TNBS (trinitrobenzenesulphonic acid)-induced colitis (Sokol et al., 2008). However, *F. prausnitzii* was found at varying abundance in the Jehai gut metagenome, many of which were uncharacterized strains. This means that available data and knowledge about the role *F. prausnitzii* plays in human health remains to be fully elucidated. There is currently insufficient knowledge about the nuanced roles different strains and their relative abundances may play in various aspects of health.

Blautia obeum was present at low abundance in the Jehai gut core metagenome. *B. obeum* was negatively associated with colonization of *Vibrio cholerae,* a causative agent for cholera. Studies showed that *B. obeum* could deplete virulence activating signals and produce virulence-suppressing signals within the gut, thereby conferring resistance to *V. cholerae* infection (Alavi et al., 2020, Liu et al., 2021b). One very interesting finding was that high levels of *B. obeum* did not seem necessary to affect *V. cholerae* (Alavi et al., 2020). A recent study found *B. obeum* at higher abundance in urban Zimbabweans than rural individuals using full-length 16S rRNA sequencing (Katsidzira et al., 2019). The Jehai samples also had a low abundance of *B. obeum.* The study sample size was very small, where ten urban samples were compared to ten rural (Katsidzira et al., 2019) and should be inferred with caution.

Evidently, more research is needed, such as understanding the specific roles bacteria play in the human host and conducting strain-level investigations. From this analysis, there were some similarities with other studies where *P. copri* and *B. obeum* have also been reported in rural microbiomes (Katsidzira et al., 2019, Tett et al., 2019). *F. prausnitzii* was detected with various strains at different abundance levels, indicating it is premature to deduce any conclusion with much that still needs to be investigated. This analysis highlights the importance of studying the microbiomes of indigenous communities, especially those such as the OA. This is perhaps the first study to report on the gut metagenomes of a semi-nomadic OA community that still lead a traditional lifestyle in relative isolation in the rainforest.

Figure 4.2 Plants grown by Jehai community shown to us by Saadiah. These were not planted yet in 2017.

4.2.2 Quality check of MAG against UHGG reference

There were a total of 38 MAGs that failed to align to the UHGG database. Two plausible reasons could explain the MAGs that were not aligned. The first was that these MAGs might reflect specific strain diversity in this indigenous community that was not found nor included in the reference genome database. As intriguing as this may be, the other reason could have been, poor MAG quality may have impeded proper alignment. This finding can be confirmed by sequencing more Jehai gut metagenomes to determine if the MAGs were truly novel and possibly unique to the Jehai community or if that was an anomaly, probably attributed to sequencing artifacts.

There was one MAG that reached a quality score of 100, far exceeding the reference genome score of 70. These higher-quality MAGs can be considered as suitable replacements for the lower quality reference genomes within the UHGG catalogue to improve the quality and utility.

4.2.3 Limitations

This analysis could be further improved by expanding on the different *P. copri* strains present in the gut metagenome and studying it in association with Jehai diets. Functional annotations were underway as this thesis was in preparation, but the analysis was not completed on time to be included. A further way to improve this study was to compare Jehai metagenomes with other Western and non-Western metagenomes. One other limitation of this study that needs to be acknowledged is that despite the process to identify the MAGs have been put to the strictest quality control possible and sequence depth was done to 100x, the MAGs that were identified as *novel species have yet to be cultured. Other methods will be needed to confirm their novelty.

4.3 Cardio-metabolic trends in Orang Asli

4.3.1 Metabolic syndrome, Cardio-metabolic health and Framingham Risk Score

The overall prevalence of MetS among the OA reported in this study from 2017-2019 was approximately 44.63%, with a slightly higher prevalence among women compared to men (49.54% vs 36.76%). The prevalence of MetS was slightly higher among the Temiar compared to Jehai. A study with data collected from 2010-2016 reported MetS prevalence of almost 40% among mostly urbanised Proto-Malays (sub-groups include Temuan, Mah Meri and Orang Seletar) (Aghakhanian et al., 2019). The MetS prevalence observed in this current study is not directly comparable to what was reported by Aghakhanian et al. (2019) because of the vast difference in sample size (629 versus 215). Nonetheless, both studies demonstrate that MetS is increasing among OA, especially those living near towns and cities.

This current study reflects many previous reports where urbanised OA communities suffered from poorer cardio-metabolic health. The Negritos, such as the hunter-gatherers Jehai and Batek, were reportedly leaner but had a higher incidence of dyslipidaemia (Aghakhanian et al., 2019). My study reflected similar findings, where urban Temuans had higher general and abdominal obesity rates than rural Temiar and semi-nomadic Jehai. On the other hand, Jehai had twice the prevalence rate of dyslipidaemia and risk for cardiovascular disease compared to Temiar. However, studies have shown that despite Negritos displaying higher coronary risk factors, biomarkers for inflammation, endothelial activation and pro-thrombosis were low (Mokhtar et al., 2016, Mokhsin et al., 2018). These studies suggested genetic factors resulting in lower HDL levels among Negrito, which was also observed in this study, although the relationship remains to be elucidated.

Jehai had a slightly higher proportion of hypertensive individuals compared to urban Temuans. This finding was similar to previous studies where Negritos, of which Jehai is a sub-group, reported the highest rate of hypertension despite being one of the most remote OA communities studied (Phipps et al., 2015, Mokhsin et al., 2018, Aghakhanian et al., 2019). One study showed that genes promoting elevated blood pressure were positively selected in Negritos (Deng et al., 2014). CDH13 gene that influences metabolism and protects against atherogenesis were also positively selected among the Negrito group (Liu et al., 2015, Teng et al., 2015). This is a classic example of natural selection and exemplifies the importance of studying different populations living in different environments (Hoh et al., 2019).

Type 2 Diabetes (T2D) was generally low among Temiar and Jehai at 8.7% and 4.9% prevalence, but almost half of Temiar were pre-diabetic. However, the high numbers of pre-diabetics among the Temiar are especially concerning. Without timely intervention, these indigenous villagers are highly likely to progress to full diabetes, thus compounding the public health issue. Insulin resistance (IR) using a surrogate biomarker TG/HDL ratio indicated that 21.9% of Temiar and 28.4% of Jehai were presumed to have IR. T2D and IR was reportedly 0.30% and 4.40%, respectively, among OA in 1992 (Ali et al., 1993). Higher prevalence of T2D and IR were observed among other OA communities living near urban areas (Wong et al., 2015, Phipps et al., 2015, Tuan Abdul Aziz et al., 2016, Aghakhanian et al., 2019).

On another note, the use of a surrogate biomarker, TG/HDL, validated by Yeh et al. (2019) in an Asian population and previously used for an OA community (Yeo et al., 2019), is a very useful alternative to the more complicated oral glucose tolerance test (OGTT). Although OGTT was used previously to

measure IR among OA communities (Ali et al., 1993), limited time and past reports of OA not responding well to the test made me decide against OGTT (Aghakhanian et al., 2019). Venous blood was not collected for this study to reduce taking blood samples from OAs due to their discomfort. Tests that required blood, such as HbA1c and lipid tests, were obtained through a finger prick method. The OA responded quite well to the finger prick method, although a skilled nurse was required to quickly draw sufficient blood for both tests before the blood clotted. Nonetheless, the lack of tests done through conventional biochemistry methods needs to be acknowledged as a limitation.

FRS score was relatively low among the OA at 4.35%, which means they had a 4.35% chance of developing cardiovascular disease within the next five years. In contrast, other major ethnic groups in Malaysia had a risk score that ranged from 11-13.4% (Borhanuddin et al., 2018). OA men had a slightly higher estimated cardiovascular risk than women, probably because men had a slightly higher mean age (38 ± CI 3.42) than women (33 ± CI 1.92). Another observation that may have contributed to this higher risk score was that most men were self-reported smokers (60.26%). Only 15.38% of women reported being smokers.

The use of different anthropometrics, lipoprotein-ratio and risk scores in this study, supported by previous studies, indicated urban OA to be at the highest risk for poor cardio-metabolic health. The rural and semi-nomadic OA communities living in the rainforest who fared marginally better were undergoing epidemiological transition.

Smoking among the OA

Temuan may have had fewer self-reported smokers, but they chewed betel quid as I had observed, noting that a few Temuan women carried plastic bags filled with betel leaves to wrap the nuts and white powder-slaked lime. Betel quid chewing was observed among the Temiar from Pos Piah (Yeo et al., 2019). The Jehai, or at least the ones who participated in the study, said they did not practice this. However, through my communications with a Jehai friend, she mentioned that some villagers, including herself, often chew betel quid. Perhaps not all of them identified with the word “pinang” (the Malay word for betel nut) or chose not to acknowledge this. Visible tell-tale signs present in regular chewers would be a reddish discolouration of the lips, teeth, mouth, tongue, and a distinctive odour.

4.3.2 Limitation of cardio-metabolic health study

One of the limitations of this cardio-metabolic health study was that the sample size was too small for statistical analysis with biological relevance (Committee, 2011, Lovell, 2013). Furthermore, more women eventually participated in this study because most of the OA men left the village for sporadic work despite prior agreements to participate on the dates that had been fixed. This merely reflects the reality and challenges when working with OA communities that are unpredictable and unavoidable. Another limitation was that the OA FRS for cardiovascular diseases was compared to a Malaysian cohort data collected from 2006-2012, which was at least ten years ago. Regrettably, the mis-categorisation of OA as ‘Malay’ under the National Health Morbidity Survey 2019 (IPH Malaysia, 2020) made it impossible to precisely distinguish the latest cardio-metabolic health data of OA from Malays, the largest ethnic group in Malaysia. To date, and to the best of my knowledge, our small study provides the latest update.

4.4 Antibiotics resistance discovery

This pilot study is the first to investigate the intestinal carriage of ABR bacteria among a semi-nomadic Jehai community. Due to delays in shipment of the agar plates, the stool samples had to be frozen for two months. Coloured pigmentation was often observed on many of the chromogenic agar in the absence of colonies (Figure 3.26). This colour pigmentation could be from residual bacterial enzymes reacting with the chromogen in the agar. There is a possibility that the ABR bacteria, be it ESBL or VRE, were present, but they were not viable to form colonies. Thus, I believe that the actual prevalence of antibiotic resistance might be, in reality, higher among the Jehai community than reported here.

4.4.1 Vancomycin-resistant Enterococcus (VRE)

VRE is usually associated with nosocomial infection. Community-acquired VRE is relatively rare. Hence the findings in this study where VRE was identified in 23% of samples was rather unexpected. *Enterococcus faecium* was the dominant bacteria isolated here, which may be a cause of concern. *E. faecium* was reportedly more resistant to antibiotics among US clinical isolates (Hidron et al., 2008) (O'Driscoll and Crank, 2015). *E. faecium* (43%) and *E. faecalis* (41%) were the pre-dominant *Enterococcus spp*. isolated in a Malaysian hospital situated in Kuala Lumpur (Ngoi et al., 2021).

Historical studies revealed that *E. faecium* was persistent in pigs thirteen years after the ban of avoparcin in animal feed, where it was postulated that Vancomycin-resistant *E. faecium* were co-selected by tetracycline, also used in animal feed (Hammerum, 2012). An observation in this study is that a small number of Jehai had a high abundance of tetracycline-resistant genes in their resistome. However, the link is very premature and is stated as a mere observation until further studies are conducted. Community-acquired VRE was first reported in a patient with an infected wound, possibly from VRE contaminated herbal leaves, after receiving treatment from a “bomoh” or a shaman (Raja et al., 2005). It is currently unknown, aside from this study, the extent of VRE spread in OA communities.

4.4.2 Extended-spectrum Beta-Lactamase (ESBL) producers

Generally, there are relatively more studies regarding the community-spread of ESBL. However, my study may be the first study to describe the intestinal carriage of ESBL among the Jehai community, which was detected in 33% of stool samples. ESBL related infections are more commonly associated with nosocomial infection, especially urinary tract infections (UTI), but community-acquired ESBLs are increasingly reported. This includes sources such as poultry and pig farms, contaminated water, and asymptomatic carriers (Aliyu et al., 2016, Mughini-Gras et al., 2019). The detection of ESBLs, in one third of stool samples, could be contributed by the Jehai community lacking modern toilets. A recent study reported that 60% of community-acquired ESBL *E.coli* were attributed primarily to human-to-human transmission, especially through the faecal-oral route (Mughini-Gras et al., 2019). The Jehai were rather discreet about their natural ‘toilets’; thus, I did not press further. Nonetheless, the lack of modern sanitation and river water being their main and shared water source may have facilitated the spread of ESBL among the community, although not excluding the possibility of other routes of transmissions yet to be discovered.

Overall, to the best of my knowledge, there only find two studies investigating ABR among OA. One study specifically looked at *E.coli* and *Campylobacter jejuni* in wild birds, chicken, and the OA in Sungai Siput, Perak, the same area as the Temiar community that participated in our study. They denoted that *E.coli* isolates from OA were resistant to at least one of the ten antibiotics investigated, especially

erythromycin and tetracycline (Mohamed et al., 2019). Furthermore, all the wild birds trapped by the OA displayed multi-drug resistance, suggesting wild animals as a possible route of ABR spread in the OA community (Mohamed et al., 2019). The second study was a retrospective case review of patients who had been administered intravenous cephalosporins for urinary tract infections in Hospital Tapah, Perak, whereby 20% of the samples represented the OA. The study concluded that IV cephalosporins were mostly used to treat UTI and leptospirosis (Mohamed et al., 2021). Furthermore, antibiotic prescriptions were not done according to the recommended guidelines to combat antibiotic resistance (Mohamed et al., 2021). Unfortunately, samples were not analysed by ethnicity. Hence the prevalence of ESBL among OA specifically was not established.

4.4.3 Gut resistome

A total of twelve Jehai gut metagenomes were available to study the gut resistome. Exactly half of the samples harboured <15 ABR genes, while another half of the samples harboured 43-71 ABR genes.

This new data presented here suggests two things. Firstly, there may be undiscovered ABR genes in the Jehai gut metagenome since gene prediction is limited by the ABR reference database. Secondly, ABR genes may be low among gut metagenomes of Jehai, who presumably had little to no exposure to antibiotics therapies. Relative abundance of ABR genes reportedly increased with exposure to antibiotics (Pérez-Cobas et al., 2013, Buelow et al., 2014). However, it is also possible for the ABR genes to decrease in the resistome should a combination of antibiotics be used and the bacteria is susceptible to either antibiotic (Pérez-Cobas et al., 2013, van Schaik, 2015).

On the other hand, ABR genes are ubiquitous and present at low levels in the environment. The compounds produced by ABR genes may be involved in cell-to-cell communication via quorum sensing (Zhang et al., 2018), aside from their more tangible role in bacterial warfare. It is uncertain if the ABR genes identified at <15 ABR genes per person is the basal Jehai gut resistome level. Increasing the sample size might be able to elucidate this finding.

Subsequently, ABR genes were notably higher among Jehai, which could be contributed by exposure to antibiotics or acquired through a hospital stay. This is partially supported by a recent study where the gut resistome of patients exposed to antibiotics increased in resistome diversity and average ABR gene copy number (Li et al., 2019). Unfortunately, planned visits to further investigate this line of thought were disrupted by the Covid-19 pandemic. A study showed that the gut resistome of two distinct cohorts reacted differently to antibiotics and was presumably linked to individual basal level microbiomes (Willmann et al., 2019). It is premature to draw any conclusions from this pilot study, but it will definitely be interesting to investigate the basal level of Jehai gut resistome and its possible role in quorum sensing in the absence of antibiotics.

Tetracycline resistant genes are found in high abundance in the human gut resistome globally (Hu et al., 2013, van Schaik, 2015), which was also found in most of the Jehai gut. Besides the ease of horizontal transfer through transposons, Tet-like genes were found in the gut resistome of ancient humans from pre-Inca/Inca period (Santiago-Rodriguez et al., 2018). This suggests that *tet* genes or the predecessor of modern *tet* genes may have been ubiquitous in the human gut resistome, even prior to the clinical use of antibiotics. The interactions, if any, between the ABR genes and human host genes have yet to be studied. Can ABR genes exist in the human gut and yet be completely independent of each other?

Before the dynamics between the gut resistome and the human host is elucidated, talk of eliminating the gut resistome may be premature and too drastic (Tsigalou et al., 2020). Humans have co-existed and co-evolved with bacteria for numerous millennia and always harboured ABR genes as far as studies on ancient human remains have shown (D'Costa et al., 2011, Santiago-Rodriguez et al., 2018). da Silva et al. (2021) found that vegetables and minimally processed food were the main sources of ABR in the human gut microbiome. The Jehai consume a wide variety of plant material that grows in the fringe of the forest. Further studies are needed to ascertain this correlation.

4.4.4 Limitations and recommendations

Detection rates could be very much improved by culturing ABR bacteria with fresh stool. It was unfortunate that the specialized agar plates were not readily available in Malaysia and had to be shipped from the UK. The freight took a long time, and they eventually arrived with short expiry dates. The culture of ABR bacteria and study of the gut resistome was limited to a subset of the Jehai community, where a small sample size severely limits a more comprehensive analysis. Grant limitation focused my efforts on the intestinal carriage of VRE and ESBL for this preliminary study, which could be expanded to other ABR bacteria. A positive control strains for both VRE and ESBL should have been included to improve validity of this study. This was unfortunately not carried out due to budget constraints to purchase ATCC control strains. Another limitation of this study was that ABR isolates were presumptive ABR colonies. ABR isolates that grew in culture were purified and frozen in glycerol. The study can be improved by conducting whole genome sequencing of the ABR isolate to confirm its identity and investigate possible novel ABR genes carried by the isolate. We had no comparable data on ABR in rural areas in Malaysia. ABR is a constant threat and may well be the next pandemic that could happen in the near future. The mortality rate of ESBL *E.coli* sepsis was reportedly three times higher than for susceptible strains (Melzer and Petersen, 2007). Hence, it is important to conduct surveillance of community-acquired ABR, especially among the rural populations.

4.5 *Helicobacter pylori*

This is possibly one of the first studies to use a stool antigen kit to investigate *H. pylori* infection among the OA communities. The use of this stool antigen test kit as a preliminary test bypasses the more invasive methods such as biopsy specimens obtained with upper gastrointestinal endoscopy for culture, histology, and urease test and other testing methods that use serology and urea breath test. Rapid diagnostic tests such as the stool antigen kits used in this study are convenient when working in the field in rural areas, where laboratory equipment may be lacking. Furthermore, sample collection and administration of the test does not require special training. Local representatives can easily be trained to administer the stool antigen test, which may be a plausible method for preliminary surveillance of *H. pylori* infections among rural communities.

4.5.1 *Helicobacter pylori* infection

H. pylori infection prevalence in Malaysia was reported to range from 24.3% to 49%. The risk factors included older age, Chinese and Indian ethnicities, low educational level and poverty (Gunaletchumy et al., 2014). There is paucity in *H. pylori* studies among the OA communities. Nonetheless, one study reported the highest *H. pylori* infection rates among the Negritos (65.7%) (Thevakumar et al., 2016). The Negritos are traditionally hunter-gatherers, such as the Batek and Jehai. Their communities are mostly found near or within the rainforest. Another study reported tap water source and relocation

out from the forest area to a resettlement village as contributing factors associated with decreased *H. pylori* infection rates (Rahim et al., 2010). Our results, which showed the highest *H. pylori* infections among the Jehai (sub-group of Negrito group) community and lowest among the urban Temuan community, are in agreement with previous studies.

H. pylori was proposed to be mainly transmitted through the faecal-oral and oral-oral route (Bui et al., 2016). Two plausible explanations can be given for the high infection rates observed among this Jehai community. Firstly, person-to-person transmission, especially intra-familial transmission (Syam et al., 2015), could have been facilitated by Jehai living in a small, close-knit community. *H. pylori* transmission was suggested to occur most commonly during infancy, whereas infections from 1-8 years old were not reported (Okuda et al., 2015, Kato et al., 2019). Identical *H. pylori* strains among family members suggested mother-to-child and sibling-to-sibling transmission as highly possible routes for intra-familial transmission (Kivi et al., 2003, Osaki et al., 2013). Close contact and sharing of resources such as food, water and living space may expedite *H. pylori* transmission among family and community members.

It should be noted that the above studies were conducted in urban Japan, well known for their good sanitation and hygiene. Hence a plausible second route of transmission is through contaminated water. A study in Peru reported low similarity in *H. pylori* strains using DNA fingerprinting between infected children-mother pairs and other family members (Herrera et al., 2008), thereby suggesting an environmental source of *H. pylori* infection. The Jehai share a common, natural water source by the river where the women and children gather every morning to socialise while washing their clothes and cooking utensils and bathing (Figure 4.3).

The TemiarGM community had *H. pylori* infection rates of 16%, which was intermediate to the other two communities. They live in sub-urban Kelantan surrounded by mountains. The main rivers that run through the Temiar settlements have been heavily polluted from excessive logging activities (Figure 4.4). Some villages have access to clean tap water, while some obtain water from natural sources (personal communication with Tok Batin, who I call Tok Damak). In contrast, the Temuan community with the lowest *H.pylori* infection rate at 6% live in urban Selangor, where they have running tap water in each of their houses. From this observation, it seems plausible that water sources may play a transmissive role in *H. pylori* infection.

4.5.2 Limitations

This was a very small, preliminary study that requires further investigation to make any definitive claims. Aside from the limited sample size, adding a second confirmatory test for all the *H. pylori* positives could further improve this study. Teaming up with a gastroenterologist may allow for more detailed medical interviews on *H. pylori*-related symptoms such as bloatedness and acid reflux.

Figure 4.3 Jehai women and children gather at the riverside near their house to wash household items while socialising and bathing themselves and the children.

Figure 4.4 Motorcyle ride by a Temiar friend, Halimah, to speak to one of the Tok Batin (Headman) of TemiarGM community called Tok Damak. Many of the rivers running through the TemiarGM settlement had severe silt pollution from incessant logging in the area despite protest from the Temiar.

4.6 Fieldwork Reflections

Working with indigenous communities can be challenging yet very rewarding in terms of unexpected discoveries and individual growth. As a non-indigenous person coming in contact with indigenous culture, it may be inevitable that I would go in with preconceived notions only to have those challenged and reassessed. Hence it is vital to engage in a reflexive practice and challenge one's possibly prejudiced view (Russell-Mundine, 2012, Nilson, 2017).

4.6.1 Researcher's reflexivity

(i) Mutual benefits for all stakeholders

There is always a danger of measuring others by one's standards and beliefs, more so when encountering people whose culture is very different and foreign to me. During our fieldwork, we tried to construct, as best as we could, a mutually beneficially relationship with the OA communities. We may have brought them doctors, medicine, food, some compensatory money and small gifts. Yet one must ask, are we doing enough? I have embarked on a scientific career riding on my experiences gained and biological data generated from these communities. Yet looking out from the window of my ivory tower, what have I done for them? For without them, there may be no ivory tower for me.

I can let more people know about the Orang Asli, their health, culture and traditions that I have learnt along the way. I can teach people to be culturally sensitive and refrain from using derogatory terms such as "Jakun[13]" or "Sakai", historically used to describe Orang Asli as wild, primitive and stupid. Perhaps what I can do for them is to set the stage for the day their children may grow up side by side with my children, equally respected and safe in school.

(ii) Communication barrier

During my first fieldwork with the Jehai community in 2017, I observed that all their stools were loose with a lot of undigested plant materials. My experience with the more urbanized OA community was that they often suffered from constipation. Even for those who did provide me with their stool samples, they were often hard. Therefore, I was interested to know what sort of vegetables the Jehai community were consuming. I was shocked.

"Do you eat vegetables?" which in Bahasa Malaysia was "Ada makan sayur?". Everyone I asked answered with a resounding no. I was baffled because they seemed very certain, while some said sometimes they did. Their stool did not look like it came from people who ate vegetables sparingly. After noting that many people denied eating vegetables, I spoke to the village teacher. I asked what type of vegetables she ate. She said none as they were expensive. I was surprised. "Why do you have to buy them? Don't you eat vegetables from the forest?" She nodded, "Oh yes, every day." Then I realised "sayur" referred to vegetables they had to buy from Gerik town, which they rarely did. When I said "ulam", which I understood as salad consisting of local greens, commonly consumed by the Malays with a chilli dip, they admitted to eating forest greens every day. This made me wonder if any more such miscommunications were masked in our questionnaire where our choice of words meant something else to the OA.

[13] Jakun is actually the name of an Orang Asli tribe who are mostly live in the southern region of Peninsular Malaysia. Malaysians grew up jokingly using "Jakun" to describe stupidity. Most do not know it is actually the name of an Orang Asli tribe.

(iii) Respect

It is one of those lessons that you learn on the first day of a Bioethics lecture, respecting and seeking consent from community elders and leaders when working with indigenous peoples. When put to the test, how many will take the easy way out? Our engagement with the Temiar Gua Musang community was in collaboration with NADI (National Diabetes Institute), who had done most of the legwork of informing the community and elders. Nonetheless, I wanted to personally inform the Temiar elders or Tok Batin (headman) that we had come to their settlement and explain our project to them. My supervisor had also reminded me to do so prior to our departure. However, it was not easy to find someone who knew where the Temiar Tok Batin was.

The Temiar Gua Musang settlement was very big, made up of different villages. Finally, I approached a group of people standing timidly by the side, away from the chaos of NADI setting up tents. The Temiar shrunk away from me at first with their guards up until I explained that I wanted to get permission from the Temiar elders to start work. I think they appreciated that I had shown respect to their culture because a few men were nodding approvingly. Finally, they asked a Temiar girl, Halimah, to give me a ride to the Tok Batin's house. To be honest, I was scared of going in alone with a stranger. I tried memorising the route we took from the community hall to the Tok Batin's house, but all the trees started to look the same after a while. Anyway, we found the Tok Batin and got his blessings to carry on. When he came to visit us amid our sample collection, he gifted me with his hat made from leaves and flowers in front of a whole hall of Temiar waiting for their turn (Figure 4.5). Whether by imagination or not, we all felt that it was his way of telling the people that we were welcomed here, which helped us gain rapport with the participants. Thankfully, our fieldwork that started rocky ended smoothly. I learnt that day that my bioethics textbooks never mentioned the challenges and mental barriers that needed to be overcome by the researcher to achieve something so basic and simple as respecting the community.

(iv) Equality

During one of our courtesy visits to the Jehai, we told them that we would bring doctors, food and pay a small compensatory sum in return for their participation. One of the men negotiated, half-jokingly, to raise the compensatory sum to RM500 and include an iPhone. I mentioned this to an anthropologist, and he replied, "This is equality and empowerment. If a community is empowered, the people will possess negotiation power and tell you what they want, instead of simply accepting what you offered."

This thought has been stuck with me until today, and it is a very uncomfortable thought. Our OA communities are not exactly empowered to the point where they negotiate their terms with us. As hard as I have tried to meet their needs within my capabilities, our exchange is still unequal in my eyes. Hence my insistence to our fieldwork team that if at any point the OA feel uncomfortable and want to withdraw, they may.

Several times I have been consulted when a participant was unwilling to cooperate. I noted their looks of disbelief several times when I smiled and said it is okay and that they do not have to do whatever they were unwilling to do. My guess is that it is not often they get their way when they refuse, hence their surprise. Any statistician would know the inconvenience of having to deal with missing data. Yet, this is what little I can do to ensure that our OA participants were treated the same way I wish to be treated.

Figure 4.5 Traditional Temiar headdress that they weave out of pandanus leaves and decorate with fragrant bay leaves and small flowers.

4.6.2 Realities of working with Orang Asli communities

The importance of understanding the community's customs and taboos, such as for the Jehai, men and women have to maintain a respectful distance, especially for in-laws where the son-in-law must not walk in front of the mother-in-law. This would explain why as fieldwork was ongoing, both in 2017 and 2019, the wooden platform where the sampling and interviews were ongoing would either consist entirely of women or entirely men. This taboo was not practised in another hunter-gatherer community, Batek, as reported in 1972 (Endicott & Endicott 2007).

The Jehai were perhaps experiencing research fatigue when we visited them in 2019. They were visited by a parasitologist group the week before our arrival. We then bumped into a group of beaming corporate company representatives on their way out after completing what was presumably their CSR (corporate social responsibility).

Fieldwork was disrupted by the Covid-19 pandemic, where OA communities became fearful of this unknown virus and even took to barricading themselves. A lot of inference of the results in this study was dependent on understanding the OA diet, environment, and lifestyle practices.

Recommendation

The few teams of researchers working with OA communities are unfortunately working in silos which can be counter-productive. The ability to react to the epidemiological transition the OA are experiencing, whereby so many researchers have provided evidence, is hindered by the lack of surveillance.

My recommendation is to have a well-managed and long-termed central database for Orang Asli health. Researchers can submit their raw data of anthropometrics and cardio-metabolic health

parameters for each community collected. The database can be expanded to include social science and psychological assessments. Currently, these data are unfortunately scattered across the internet. With a centralised database, communities that are less visited by researchers and medical teams (perhaps due to inconvenience and inaccessibility) can be quickly identified.

Longitudinal changes in the community's health can easily be tracked longitudinally as different researchers collaborate with the communities. Knowledge is power. Surveillance will allow health policies and interventions to happen before outbreaks occur. Our current problem with limited biomedical data on the OA communities, highlighted by (Baer, 1999), makes it very difficult to determine any trends or changes to the community that resulted in their current epidemiological state. What we do have are sparing reports of OA health from various communities at various points in time.

Adela Baer, one of the earliest biomedical researchers on OA health, once said, “Health is an affirmation of life, which is culturally constructed.” (Baer, 2006). It is crucial to improve the health system to become more culturally sensitive and friendly towards OA. The Hospital Orang Asli in Gombak, Selangor is just one hospital. We often hear from the OA of fear and mistrust of going to the hospital due to plain disrespect and mistreatment by hospital staff (Nicholas and Baer, 2007, Wong et al., 2019b).

Field researchers, including doctors, anthropologists and scientists, are people who can provide preliminary information about what the communities need. The best people who can provide this information would be the OA themselves. Researchers can act as a bridge to getting the OA’s voices heard by the right people. WHO and OHCR (Office of the United Nations High Commissioner for Human Rights) stated, “the right to health is a fundamental human right” (WHO, 2008). Extraordinary things can happen when ordinary people try and do small, seemingly ordinary things for an extended time.

Chapter 5 Conclusion and Future Work

5.1 Conclusion

5.1.1 Orang Asli saliva and gut microbiomes

Human microbiome studies are of much interest, both to researchers and the general public because these findings can help inform the choices of our daily lives. Microbiome studies have helped to provide empirical evidence for benefits in increasing our dietary fibre intake (Zhao et al., 2018), breast-feeding our children (Mueller et al., 2015), practising caution when taking antibiotics (Langdon et al., 2016) and that perhaps a little dirt is good for our health (Bloomfield et al., 2016). Clinical studies have also shown that mediation of the human microbiome may provide alternative therapeutics to help improve post-surgical outcomes (Davies et al., 2019, Ferrie et al., 2021) and reduce the adverse effects of drug medications (Wu et al., 2017, Javdan et al., 2020).

Nonetheless, we are at the early stages of understanding the human microbiome. As Einstein once said, “Insanity is doing the same thing over and over again and expecting different results.” Hence, I turned my focus to understanding the microbiome of rural and traditional populations who lead lives that are rather distinct from urban people. The OA of Peninsular Malaysia are a very diverse group of people with unique lifestyles. This allowed us to explore and characterize the saliva and gut microbiomes of four OA communities ranging from urban, sub-urban, rural communities, and semi-nomadic hunter-gatherers. In the absence of available literature on the OA microbiomes when I started, the first objective was **to characterize their saliva and gut microbiomes using targeted 16S rRNA sequencing.**

I engaged with each OA community's elders and community leaders and invited villagers to participate in my study. Interviews, anthropometrics, saliva, and stool samples were collected during fieldwork. DNA was extracted from saliva and stool samples, followed by amplification and sequencing of the 16S rRNA hypervariable region. Sequencing data were analysed on the V4 region using bioinformatic tools such as QIIME2 and R packages. The saliva microbiomes of rural Temiar denoted the lowest alpha diversity among all the OA communities. The alpha diversity of semi-nomadic Jehai and urban Temuan saliva microbiomes were comparable and marginally higher in diversity when compared to Temiar and sub-urban TemiarGM. My findings suggest that saliva microbiomes of hunter-gatherers and urbanised OA have higher alpha diversity, whereas communities in transition may exhibit lower alpha diversity. The urban-like OA saliva microbiomes of Temuan and TemiarGM were enriched in bacteria commonly associated with poorer oral health, such as *Corynebacterium, Prevotella* and *Mogibacterium*.

The alpha diversity of OA gut microbiomes denoted a transition from most rural to urban where semi-nomadic Jehai exhibited the highest diversity, followed by Temiar and TemiarGM. Urban Temuan exhibited the lowest alpha diversity. *Sutterella*, a bacteria genus previously found in high abundance in gut microbiomes of other rural populations and obese individuals, was reduced in Jehai gut microbiomes. Jehai were leaner than most Temuans. *Odoribacter, Blautia, Lachnoclostridium, Parabacteroides, Bacteroides* and Ruminococcaceae UCG-013 were also found more abundant in urban Temuan gut microbiomes compared to Jehai.

With much interest in semi-nomadic communities and limited resources, I prioritized the Jehai group to investigate **the core microbiome and identify *novel bacterial species in their gut using shotgun metagenomics (objective 2).**

The Jehai gut metagenome mapped very well to UHGG reference database, where only 9% reads were not classified. There were 73 core species identified, with several core species mapping to relatively new and uncharacterized bacteria. *Prevotella copri* was present in high abundance. Several strains of *Faecalibacterium prausnitzii* were found at varying abundance as part of the core gut metagenome. *Blautia* of unknown species were found at low abundance levels. A total of 28 *novel species were discovered, many found at high abundance, which mapped to the genus *Collinsella.*

5.1.2 Orang Asli health

Cardio-metabolic health

Past research has indicated that the OA is currently undergoing epidemiological transition and are at the highest risk for poor health (Pell et al., 2016). There was no recent information about the health of these communities. Hence, I set out **to assess the current cardio-metabolic trend and metabolic syndrome among these four OA communities (objective 3).**

During fieldwork and sample collection, anthropometric measurements were collected. A finger-prick method was used to measure HbA1c and blood lipid levels. The lipoprotein ratio was calculated to report on dyslipidaemia and used as a surrogate marker for insulin resistance. Metabolic syndrome and Framingham Risk Score (FRS) were assessed for cardiovascular disease risk among the OA.

Overall, urban Temuan suffered from poorer cardio-metabolic health and the highest prevalence of general and abdominal obesity. Interestingly, semi-nomadic Jehai had double the prevalence rate of dyslipidaemia and had a marginally higher prevalence of hypertension despite being the community living in the most rural area. Type 2 Diabetes and insulin resistance, measured for Jehai and Temiar, was generally low, although there was a marked increase compared to data some thirty years ago. The overall prevalence of metabolic syndrome among the OA was 44.63% and seemed to more prevalent among women than men (49.54% vs 36.76%). MetS was more prevalent among Temiar than Jehai. FRS for cardiovascular risk in the next five years were overall low at 4.35%.

Antibiotic-resistant bacteria and Helicobacter pylori infections

Antibiotic resistance (ABR) is becoming a regional and global threat. ABR investigations in Malaysia have been conducted mainly in clinical settings, whereas ABR in the food-chain and community have primarily been carried out independently by research institutes. However, there was little information about ABR spread among the community, more specifically ESBL and VRE among the OA communities. Therefore, as a discovery-based preliminary study, **I gathered evidence of the intestinal carriage of ABR bacteria in the Jehai community by culturing ESBL and VRE on chromogenic agar (objective 4).** This is the first report of intestinal carriage of ESBL and VRE in Jehai, with reported rates of 33% and 23%, respectively. Contrary to the simple assumption that this semi-nomadic hunter-gatherer community would have less exposure to antibiotics and thus lower intestinal carriage of ABR bacteria, this finding was rather unexpected.

I furthered my search by identifying the gut resistome in Jehai to gain an overall perspective of ABR presence using whole genome sequencing (objective 5). Half the Jehai (n=6) harboured <15 ABR genes in their gut resistome, whereas the other half had 43-71 ABR genes. ABR genes that could be found in >80% of the gut resistome include *rpoB, CfxA6, dfrE* and *tetM, Q, W.* The major ABR gene

families such as *emr* and *mdt,* were identified only in Jehai gut resistomes that harboured higher levels of ABR genes. *K. pneumoniae* ABR genes were identified in <16% of the Jehai gut resistomes.

There has been very few reports on the spread of *H. pylori* infection among the OA communities. Therefore**, I determined the prevalence of *H.pylori* infections using a stool antigen kit (objective 6).** Past studies have mostly utilised more invasive methods. This may be the first study to use a stool antigen kit to investigate *H. pylori* infection among the OA. *H.pylori* infection was the highest among the Jehai at 78% prevalence, followed by TemiarGM at 16% and, lastly urban Temuan at 6%.

5.2 Future Work

5.2.1 Targeted 16S rRNA studies

(a) Characterising the microbiome using metabolic function

The core microbiome is usually described at the taxonomic level in most studies to date. An alternative and useful method would be to define the core microbiome using metabolic function. The functional core would offer a complementary understanding of the microbiome and host-microbe interaction by disregarding bacteria taxa to focus on essential biological functions that the microbiota performs in symbiosis with the host (Risely, 2020). One of the possible benefits would be the functions of rare, low abundant taxa could be characterised better as they have been suggested to be insightful biomarkers of dysbiosis (Davenport et al., 2014, Cena et al., 2021). Another approach that may yield interesting results would be by grouping bacteria into functional guilds. This would bypass the traditional taxonomic classification of bacteria and focus only on the functional whole of the microbiome (Wu et al., 2021).

(b) Enterotypes, diet and cardio-metabolic risks

My doctoral work was a relatively small study to elucidate and characterise the microbiomes of OA communities. Subsequent work may build on the data and information I have obtained to design a study that quantifies the diet of OA communities which may shed more light on the nutritional level of their diets in association with their health. Next, their microbiomes could be classified into enterotypes, explicitly looking for *Prevotella, Bacteroides* and *Ruminococcaceae* gut enterotypes. These enterotypes have been reported to reflect cardio-metabolic health indicators such as BMI, blood lipid levels and blood glucose (de Moraes et al., 2017). As a note of caution, studies looking at enterotypes seemed to reveal inconsistent findings when expanded to several populations across the globe (Gorvitovskaia et al., 2016). This is not unexpected given the dynamics and complexities of human microbiomes. I would advocate that microbiomes need to be studied in a localised context, taking into account the effects of diet, environment and lifestyle. Specifically for working with indigenous communities and elucidating the role diet plays in the microbiomes, I would strongly recommend engaging with a field anthropologist familiar with the community.

5.2.2 Shotgun metagenomics

(a) Strain-level diversity and analysis

From this preliminary study, it is evident that there are a few *novel species and different abundance of bacteria at the strain-level. *P. copri* is just one of many bacteria where different strains play different roles in the human microbiome (Metwaly and Haller, 2019, Claus, 2019). Other bacteria identified in this study such as *F. prausnitzii* and *B. obeum* have little available information at the strain-level. Hence,

there is much work to be done in investigating the contributions of these seemingly important bacteria to host health and diseases.

(b) Exploration of the oral metagenome and resistome

There is a bias in microbiome research towards gastrointestinal bacteria. Consequently, the oral metagenome and resistome are severely under-studied (Nath et al., 2021). A recent study denoted the oral resistome to harbour a high abundance of ABR genes, albeit relatively lower in diversity than the gut resistome (Carr et al., 2020). Furthermore, and perhaps not surprisingly, the oral resistome was rather distinct from the gut resistome (Carr et al., 2020). The oral metagenome and resistome, generally less studied, have yet to be characterized in an OA cohort.

(c) Culturomics

This study had only looked at the intestinal carriage of VRE and ESBL in the Jehai gut. Aside from expanding the study to investigate other ABR bacteria, it would be interesting to use culturomics, a mixture of culture and metagenomics, to study intestinal carriage of ABR. Culturomics might identify novel functions of ABR genes in the gut resistome that might have nothing to do with conferring antibiotic resistance (Raymond et al., 2019). Moreover, culturomics could characterise previously 'unculturable' species found in the OA guts, which may improve our overall understanding of microbial diversity with potential host health implications (Browne et al., 2016, Lagier et al., 2018, Cross et al., 2019).

5.2.3 *Helicobacter pylori*

(a) H. pylori genome as a mapping tool for prehistoric migration studies

The evolution patterns of *H. pylori* can be used to depict various prehistoric human migration events. An ancient *H. pylori* genome from the European Copper Age mummy, The "Iceman" from 5,300 years ago suggested a second wave of Out of Africa migration (Maixner et al., 2016). Southeast Asia and its *H. pylori* strain hpAsia2 have been used to trace the history of human migration, which may have impacted the incidence of gastric cancer in Asia, particularly China (Breurec et al., 2011). It may be useful to further study *H. pylori* strains isolated from OA, especially the Jehai, in the light of human population migration history.

(b) H. pylori and its relationship with the gut microbiome

H. pylori has been with its human host for a long time and has sometimes been referred to as a "bacterial parasite". However, its' recent absence in humans may have caused human physiological changes. Researchers have suggested its disappearance to be associated with the rise in asthma, allergies and oesophageal adenocarcinoma (Blaser, 2008, Blaser et al., 2008, Amedei et al., 2010, Arram et al., 2012). A few studies have hinted at crosstalk occurring between *H. pylori* and other bacterial residents in the gut, perhaps through metabolites or inflammation(He et al., 2016). The role of quorum sensing and biofilm formation have also been indicated in the success of *H. pylori* colonization (Challa and Neelapu, 2018). The role and effects of interactions between *H. pylori* and the gut microbiota are interesting questions to be answered. Studies have reported that eradicating *H. pylori* caused changes to gut microbiota (He et al., 2016, Yap et al., 2016). Whether this is due to "bykill" from the antibiotics or from the absence of quorum signals from *H. pylori* remains to be investigated. It should also be noted that active *H. pylori* infections seemed to contribute to gut dysbiosis (Frost et al., 2019, Wang et al., 2019, Yang et al., 2019).

5.2.4 Community outreach and engagement

More fieldwork and follow up visits that had been planned initially have been completely disrupted due to the Covid-19 pandemic. We are unable to return to the communities to share the findings of this research as Malaysia has been under movement control orders and frequent lockdowns since March 2020. All interstate travel has been prohibited, and the risk of outsiders transmitting Covid-19 to vulnerable OA communities remains high. In the meantime, I can organise online zoom meetings to share the findings and stories of working with the OA communities with my peers, university and high school students and OA interest groups. I can do my part by letting more people know about the OA, their current health situation and relevant research ongoing in Malaysia. These often marginalised and unrepresented communities have been unfairly excluded from our history textbooks in government schools. The more people know about them, the more they will be appreciated and the less likely they will be forgotten.

Appendix I

Saliva DNA extraction protocol – modified from (Quinque et al., 2006, Yeo et al., 2019) #updated and final modification 18th Feb 2020

1. Divide samples into 1ml aliquots and add equal volume of lysis buffer in 15ml tube *Lysis buffer: 50mM Tris, pH 8.0, 50mM EDTA, 50mM sucrose, 100mM NaCl, 1% SDS.
2. Add 40µl Proteinase K (20mg/ml) and 150 µl of 10% SDS to 2ml of saliva extraction buffer mix. Vortex ~2-secs. Incubate overnight at 53 degree celcius (~16 hrs) in shaking water bath.
3. Remove from water bath and add 1ul RNAse A (10mg/ml), incubate for 30mins in 37 degree celcius water bath.
4. Add 400µl of 5M NaCl and 1000ul of phenol:chloroform:isoamyl alcohol (25:24:1, v/v). Vortex ~25secs.
5. Centrifuge at highest speed (4.4 x1000 rpm) for 15mins at 10 degree celcius.
6. Extract the supernatant (~2ml, careful not to touch the phenol part at the bottom of the tube). Add in 2.5X volume (~5ml) of cold absolute ethanol. Vortex at lower speed to mix for ~15secs. Incubate for 2hours at -20 degree celcius.
7. Centrifuge for 15min at highest speed at 4 degree celcius. Discard supernatant.
8. Wash pellet with 1000ul of cold 70% ethanol. Centrifuge for 1 min at 4 degree celcius.
9. Discard the supernatant and dry the pellet overnight in laminar flow.
10. Resuspend pellet in 80ul TE buffer, stand at room temperature for 1 hour to fully dissolve pellet.
11. Measure quality and change 15ml to 1.5ml tube for storage.

IHMS fecal DNA extraction protocol Q -using Qiagen QIAamp DNA mini kit (Costea et al., 2017)

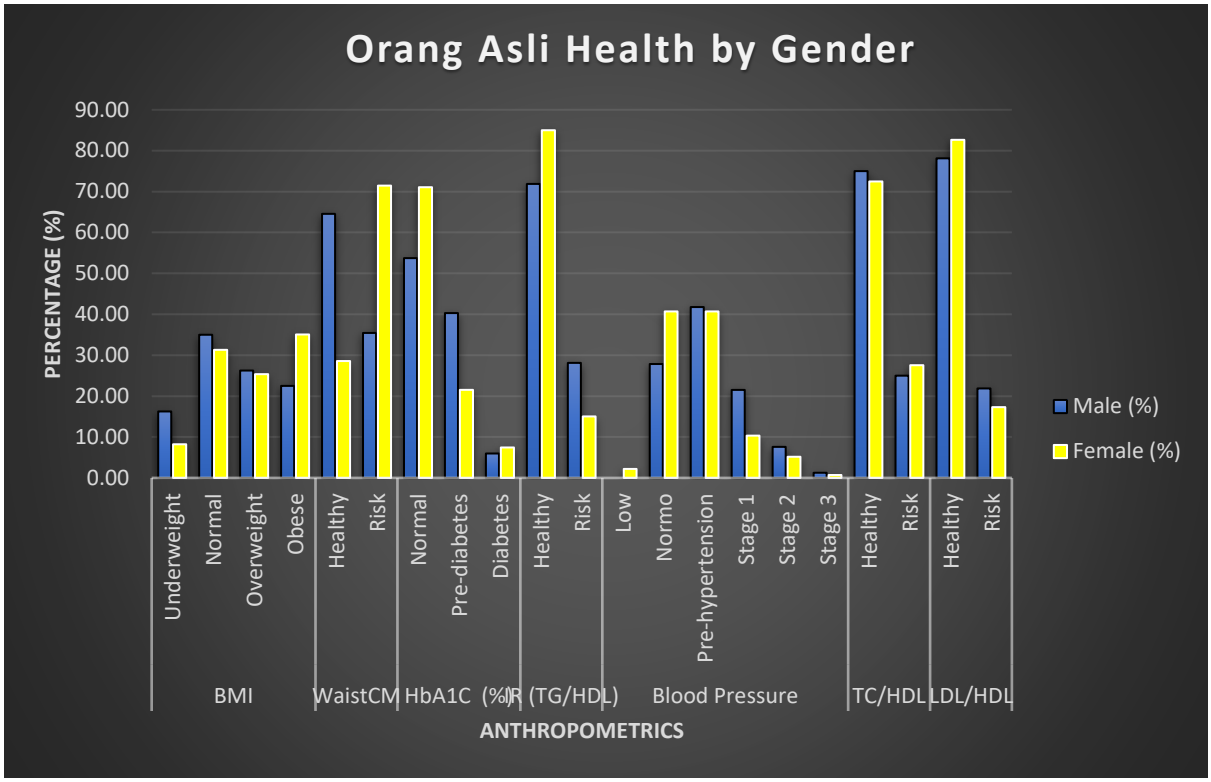
* modified by LF.Yeo due to material availability but overall flow remains unchanged.

1. Homogenize 150-200 mg (not more!) of frozen feces with 1 ml ASL lysis buffer by vortexing for 2 min in 2 ml tube containing 0.3 g of sterile zirconia beads (or glass beads 0.1 mm.
2. Incubate for 15 mins at 95℃
3. Mechanically lyse cells by running Fastprep Instrument for 8 min (series of beating 1 min and resting 5 mins)
4. Cool samples down (with/without ice). Centrifuge samples at 16000 xg, 4℃, 5 min. Supernatant is transferred to a new 2 ml tube
5. The pellet is mixed with 300 µl ASL lysis buffer, vortex and repeat step 2-6
6. Add 260 µl of 10M ammonium acetate to lysate tube, mix well and incubate on ice for 5 min.
7. Centrifuge at 16000xg, 4℃, 10 mins.
8. Transfer supernatant to two 1.5 ml Eppendorf tubes (~400 µl each) and 2X ice cold abs. EtOH (~1000 µl). Invert many times and incubate at -20℃ for 30 mins.
9. Centrifuge at 16000 xg 4℃, 15 min. Pour away supernatant and wash DNA pellet with 800 µl of 70% EtOH.
10. Centrifuge at 16000 xg, 4℃, 1min. Pour away supernatant and dry the pellet for 3 mins.
11. Dissolve DNA pellet in 100 µl of TE buffer, add 1 µl of RNAse A and incubate at 37℃ for 30 mins (until pellet dissolves, flicking the tube helps dislodge pellet and dissolve)
12. Pool samples together.
13. Add 15 µl proteinase K and 200 µl AL buffer to supernatant. Vortex for 15 sec, centrifuge tubes and incubate at 70℃ for 10 mins. [don't add proteinase K directly to AL)
14. Add 200 µl of abs EtOH to lysate. Mix by vortexing.
15. Transfer to QIAamp spin column and centrifuge at 16000 xg, RT, 1min
16. Discard flow, change collection tube, add 500 µl buffer AW1 and centrifuge at 16000 xg, RT, 1 min (must change collection tube! Prevent cross-contamination)
17. Discard flow, change collection tube, add 500 µl buffer AW2 and centrifuge at 16000xg, RT, 1 min
18. Dry the column
19. Add 80 µl of TE buffer, incubate for 3 mins at RT. Centrifuge for 1 min at 16000 xg to elute DNA.

Antibiotic resistance bacteria screening from frozen stool

1. Stool was collected in specimen containers in the field.
2. A sterile cotton swab was dipped inside the stool and inoculated in 15% glycerol +0.9M saline solution. Samples were kept in -80 degree celsius until agar plates arrived from UK.
3. Revivement of frozen sampels was done with Tryptic Soy Broth (TSB) in shaking incubator at 37 degree celsius for two days.
4. 200ul of broth was pipetted onto selective, chromogenic Brilliance VRE and ESBL (Oxoid, UK) agar plates. Incubate at 37 degree celsius for 24 and 48 hours
5. Coloured colonies according to manufacturer's instructions (presumptive VRE or ESBL) are purified on MacConkey agar.
6. Pure colonies are stored in 30% glycerol + TSB at -80 degree celcius for further analysis.

Appendix II

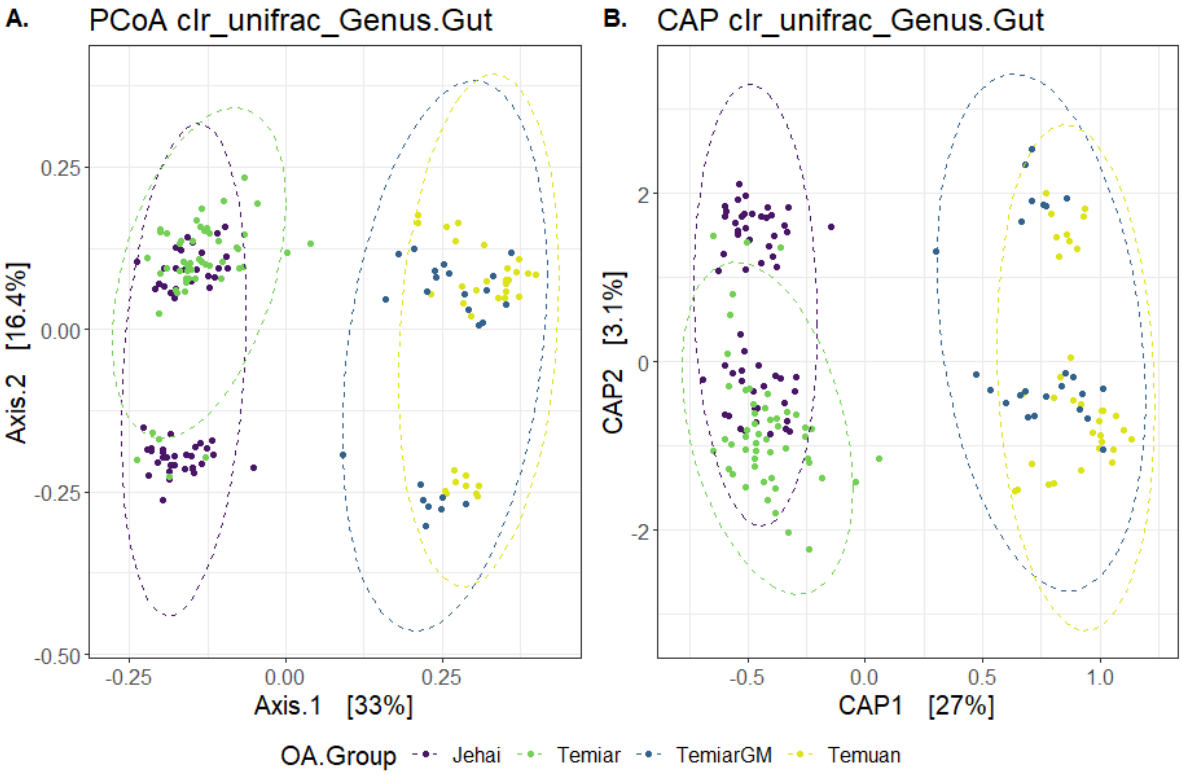


Supp Figure 1 Overall results of the Orang Asli cardio-metabolic health segregated by gender.

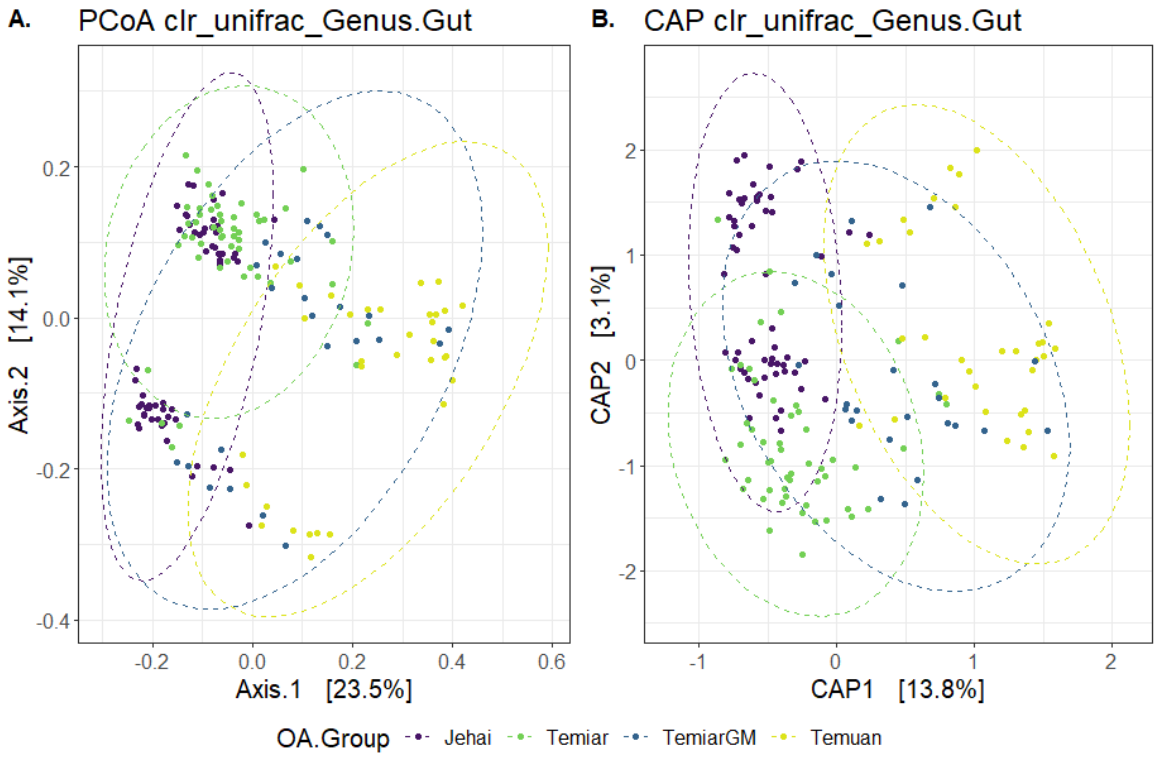
Supp Table 1 Number of Temiar from Pos Piah who chewed betel quids (Yeo, 2017 PgDip Thesis)

Gender	Betel nut consumption	Smoking status	Frequency
Men	Chew betel nuts	Smokers	15
		Never smoke	2
		Former smokers	3
	Don't chew betel nuts	Smokers	6
		Never smoke	4
		Former smokers	2
		Total	***32***
Women	Chew betel nuts	Smokers	6
		Never smoke	20
	Don't chew betel nuts	Smokers	1
		Never smoke	10
		Total	***37***

Testing suitability of unweighted UniFrac



Supp Figure 2 Unifrac at the Genus level using SEPP method shows a seemingly random separation horizontally across the plot.



Supp Figure 3 This plot was generated using data trimmed to V4 region. The samples seem to merge a bit better. However, the random separation cutting across the samples was still evident, especially in the PCoA plot. Hence, I decided that perhaps uniFrac was not suitable for the analysis, and chose Euclidean distance.

Supp Table 2 Complete list of ABR genes identified in Jehai gut resistome

		rifamycin	fluoroquinolone;macrolide;penam	cephamycin	lincosamide;macrolide;streptogramin	cephalosporin;fluoroquinolone;glycylcycline;penam;phenicol;rifamycin;tetracycline;triclosan	cephalosporin;penam	cephalosporin;penam	cephalosporin;penam	macrolide	benzalkonium_chloride;rhodamine;tetracycline	aminoglycoside;cephalosporin;macrolide;peptide;rifamycin;tetracycline	aminoglycoside;cephalosporin;macrolide;peptide;rifamycin;tetracycline
SampleID	Total ABR genes	**rpoB**	CRP	**CfxA6**	ErmF	E.coli_acrA	E.coli_amp	E.coli_amp	E.coli_amp	E.coli_emr	E.coli_mdf	K.pneumo	K.pneumo
JH19002	10	100	.	99.9	.	.	.	.	.	.	.	.	.
JH19004	11	100	.	99.9	.	.	.	.	.	.	.	.	.
JH19016	11	94.80;100.	.	99.9	.	.	.	.	.	.	.	.	.
JH19041	11	100	.	99.5	100	.	.	.	.	.	.	.	.
JH19010	14	100	.	99.5	100	.	.	.	.	99.7	.	.	.
JH19038	14	100	.	99.9	100	.	.	.	.	.	.	.	.
JH19024	43	100	100	99.5	100	.	100	.	.	.	.	.	100
JH19031	44	100	100	99.5	100	.	100	.	.	100	.	.	100
JH19018	52	100	100	99.9	100	100	100	.	100	100	100	.	.
JH19026	60	.	100	99.5	100	100	100	100	100	100	100	.	.
JH19004	67	99.89	100	.	.	100	100	100	100	100	100	.	.
JH19009	71	100	100	.	.	100	100	100	100	100.00;10	100	100	.

aminoglycoside;carbapenem;cephalosporin;fluoroquinolone;macrolide;penam;penem;peptide	aminoglycoside;carbapenem;cephalosporin;fluoroquinolone;macrolide;penam;penem;peptide	cephalosporin;fluoroquinolone;glycylcycline;penam;phenicol;rifamycin;tetracycline;triclosan	diaminopyrimidine	fluoroquinolone	fluoroquinolone	tetracycline	fluoroquinolone	aminocoumarin	aminocoumarin	aminocoumarin	fluoroquinolone;macrolide;penam	fluoroquinolone;macrolide;penam	fosfomycin
K.pneumo	K.pneumo	K.pneumo	**dfrF**	emrA	emrB	emrK	emrR	mdtA	mdtB	mdtC	mdtE	mdtF	mdtG
.	.	.	.	.	.	.	100	.	.	.	.	.	.
.	.	.	100	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	100	.	.	.	.	.	.	.	.	.	.
.	.	.	100	.	.	.	.	.	.	.	.	.	.
.	.	.	100	.	.	.	.	.	.	.	.	.	.
.	.	.	100	100	100	100	100	.	.	.	.	100	.
100	99.91	99.75	.	.	100	100	100	.	.	.	100	.	.
.	.	.	.	100	100	100	100	100	.	100	100	100	100
.	.	.	100	100	100	100	100	100	100	100	100	100	100
.	.	.	100	100	100	100	100	100	100	100	100	100	100
99.23	100	90.56	100	100	100	100	100	100	.	100	100	100	100

fluoroquinolone	acridine_dye;fluoroquinolone;lincosamide;nucleoside;phenicol	acridine_dye;nucleoside	acridine_dye;nucleoside	acridine_dye;nucleoside	Tetracyline	Tetracyline	Tetracyline	Tetracyline
mdtH	mdtM	mdtN	mdtO	mdtP	**tet32**	**tetM**	**tetQ**	tetW
.	.	.	.	.	.	.	100	100
.	.	.	.	.	100	100	100	.
.	.	.	.	.	100	100	97.92	.
.	.	.	.	.	.	100	98.43	97.34
.	.	.	.	.	100	100	100	100
.	.	.	.	.	100	.	100	100
.	98.3	100	100	.	100	100	100	100
100	.	100	100	.	100	100	99.34	98.59
100	.	100	100	100	100	100	99.34	100
100	100	100	100	100	100	100	100	100
100	100	100	100	100	100	100	.	100
100	100	100	100	100	.	100	.	98.75

Appendix III

F1000Research F1000Research 2019, 8:175 Last updated: 24 NOV 2019

Check for updates

RESEARCH ARTICLE

REVISED Health and saliva microbiomes of a semi-urbanized indigenous tribe in Peninsular Malaysia [version 3; peer review: 2 approved]

Li-Fang Yeo [1,2], Farhang F. Aghakhanian[1,2], James S. Y. Tan[3], Han Ming Gan [4], Maude E. Phipps[1]

[1]Jeffrey Cheah School of Medicine and Health Sciences, Monash University, Bandar Sunway, Selangor, 46150, Malaysia
[2]Tropical Medicine and Biology Platform, Monash University, Bandar Sunway, Selangor, 46150, Malaysia
[3]Imperial College London, South Kensington Campus, London, SW7 2AZ, UK
[4]School of Life & Env Sciences, Deakin University, Geelong Waurn Ponds Campus, Australia

v3 **First published:** 11 Feb 2019, 8:175 (https://doi.org/10.12688/f1000research.17706.1)
Second version: 09 May 2019, 8:175 (https://doi.org/10.12688/f1000research.17706.2)
Latest published: 28 May 2019, 8:175 (https://doi.org/10.12688/f1000research.17706.3)

Abstract

Background: The indigenous people of Peninsular Malaysia, also known as Orang Asli, have gradually been urbanized. A shift towards non-communicable diseases commonly associated with sedentary lifestyles have been reported in many tribes. This study engaged with a semi-urbanized Temiar tribe from Kampong Pos Piah, Perak, who are experiencing an epidemiological transition.
Methods: Weight, height, waist circumference, blood pressure, HbA1C and lipid levels were measured as indicators of cardio-metabolic health. DNA was extracted from saliva using salting-out method followed by PCR amplification of the V3-V4 region of the 16S rRNA gene and sequencing on Illumina MiSeq. Microbiome analysis was conducted on Qiime v1.9. Statistical analysis was conducted using Qiime v1.9 and R.
Results: The study revealed that 60.4% of the Temiar community were overweight/obese, with a higher prevalence among women. HbA1C levels showed that 45% of Temiar had pre-diabetes. Insulin resistance was identified in 21% of Temiar by using a surrogate marker, TG/HDL. In total, 56.5% of Temiar were pre-hypertensive, and the condition was prevalent across all age-groups. The saliva microbiome profiles of Temiar revealed significant differences by gender, BMI, abdominal obesity as well as smoking status. The relative abundance of the genus *Bifidobacterium* was increased in men whereas the genera *Prevotella, Capnocytophaga, Leptotrichia, Neisseria and Streptococcus* were increased in women. Proteobacteria was significantly depleted in smokers.
Conclusions: Temiar from Pos Piah had a high prevalence of cardio-metabolic risks, including general and abdominal obesity, pre-diabetes, prehypertension and hypertension. This phenomenon has not

Open Peer Review

Reviewer Status ✓ ✓

	Invited Reviewers 1	2
REVISED version 3 published 28 May 2019	✓ report	
REVISED version 2 published 09 May 2019	? report	✓ report
version 1 published 11 Feb 2019	? report	? report

1 **Andres Gomez**, University of Minnesota, Saint Paul, USA

2 **Siti Nursheena Mohd Zain**, University of Malaya, Kuala Lumpur, Malaysia

Any reports and responses or comments on the article can be found at the end of the article.

PUBLIC HEALTH 176 (2019) 106–113

Available online at www.sciencedirect.com

Public Health

journal homepage: www.elsevier.com/puhe

Themed Paper– Original Research

Metabolic syndrome and cardiometabolic risk factors among indigenous Malaysians

F. Aghakhanian [a,e], C. Wong [a], J.S.Y. Tan [b], L.F. Yeo [a], A. Ramadas [a], J. Edo [c], B.P. Hoh [d], B.A.K. Khalid [a], M.E. Phipps [a,*]

[a] Jeffrey Cheah School of Medicine and Health Sciences, Monash University Malaysia, Jalan Lagoon Selatan, 47500, Sunway City, Selangor, Malaysia
[b] Imperial College, South Kensington Campus, London, SW7 2AZ, United Kingdom
[c] Community Elder, Kampung Asli Erang, 36500, Ulu Bernam, Perak, Malaysia
[d] Faculty of Medicine and Health Sciences, UCSI University, Jalan Menara Gading, Taman Connaught, Cheras, 56000, Kuala Lumpur, Malaysia
[e] Tropical Medicine and Biology Platform, Monash University Malaysia, Jalan Lagoon Selatan, Sunway, 47500, Selangor, Malaysia

ARTICLE INFO

Article history:
Received 7 February 2018
Received in revised form
26 July 2018
Accepted 1 October 2018
Available online 1 December 2018

Keywords:
Metabolic syndrome
Cardiovascular disease risk
Indigenous communities

ABSTRACT

Objectives: This study was undertaken to investigate the occurrence of metabolic syndrome (MetS) and cardiovascular disease (CVD) risk in Orang Asli (OA), the indigenous people of Peninsular Malaysia. OA consist of Negrito, Proto-Malay, and Senoi groups who collectively comprise only 0.76% of the population of Peninsular Malaysia. Owing to the challenges in accessing their remote villages, these groups are often excluded in larger government health surveys. Although tropical diseases were scourges in the past, with rapid national development, many OA communities have been gradually urbanized. We believe an epidemiological transition is occurring and non-communicable diseases are on the rise.

Study design: A retrospective cross-sectional study.

Methods: Indigenous Malaysians (n = 629) from three major groups (Negrito, Proto-Malay, and Senoi) were recruited, after ethics approval and informed consent. Body mass index (BMI), body weight, height, waist circumference, and systolic and diastolic blood pressure were measured, and participants were examined for acanthosis nigricans. Venous blood samples were used for measurements of fasting blood sugar, triglycerides (TG), total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C) and low-density lipoprotein cholesterol (LDL-C). Insulin resistance was estimated using a surrogate measurement TG/HDL-C. The ratios of TC to HDL-C, and of LDL-C to HDL-C were determined. MetS was accessed according to the Joint Interim Statement of the IDF Task Force on Epidemiology and Prevention.

Results: MetS affected 29.57% of the OA population investigated and was significantly more prevalent ($P < 0.05$) in women than in men (35.25% vs 21.95%, $P < 0.001$). MetS prevalence was the highest among the Proto-Malays (39.56%), followed by Negritos (26.35%) and Senois (11.26%). The most prevalent risk factor among the Negritos with MetS was low HDL-C (95.35%), whereas central obesity was the most common risk factor among the Proto-Malays (82.91%). In contrast, hypertension was the commonest risk factor among the

* Corresponding author. Tel: +603 5514 5848, fax: +603 55146323
E-mail address: Maude.Phipps@monash.edu (M.E. Phipps).
https://doi.org/10.1016/j.puhe.2018.10.001

Appendix IV

Questionnaire and consent form

OA Study:STN1

SOSIODEMOGRAFI [Di isi oleh peg. kajiselidik]

1. Nama penuh responden:_____

2. No. kad pengenalan: _____

3. Tarikh lahir : ☐☐ ☐☐ ☐☐

4. Umur ☐☐ tahun

5. No. Talipon ☐☐☐ ☐☐☐☐☐☐☐

6. Alamat rumah:_____

_____

7. Jantina ☐ 1 Lelaki ☐ 2 Perempuan

8. Bangsa ☐ Nyatakan _____

9. Ahli keluarga

	Nama	Bangsa	Tempat lahir
Bapa			
Ibu			
Datuk			
Nenek			

10. Agama ☐ 1 Islam ☐ 2 Buddha ☐ 3 Kristian
☐ 4 Hindu ☐ 5 Sikh ☐ 6 Lain-lain Nyatakan_____

11. Pekerjaan

☐ 1 Penggubal undang-undang, pegawai kanan & pengurusan
☐ 2 Professional, juruteknik & professional bersekutu
☐ 3 Perkeranian
☐ 4 Pekerja perkhidmatan, kedai & jurujual
☐ 5 Pekerja mahir, pertanian & perikanan
☐ 6 Pekerja pertukangan & berkaitan
☐ 7 Operator loji, mesin & pemasang
☐ 8 Pekerjaan asas_____(nyatakan)

12. Tahap pendidikan tertinggi

☐ 1 Tiada persekolahan formal ☐ 2 Tidak tamat sekolah rendah
☐ 3 Tamat sekolah rendah ☐ 4 LCE/ SRP/ PMR
☐ 5 MCE/ SPM ☐ 6 STPM/Sijil/ Diploma
☐ 7 Ijazah ☐ 8 Ijazah lanjutan

1

13 Purata pendapatan bulanan isi rumah _____ (nyatakan)

☐ 1 RM 415 dan kurang ☐ 2 RM 416 – RM 690

☐ 3 RM 691 – RM 1000 ☐ 4 RM 1000 – RM 5000

☐ 5 RM 5000 – RM 10 000 ☐ 6 Lebih dari RM 10 000

13 Status perkahwinan ☐ 1 Kahwin ☐ 2 Bujang ☐ 3 Duda / Janda/ Balu

13.1 Jika berkahwin dan mempunyai anak, sila senaraikan nama anak-anak

_____

_____

_____

_____

_____

_____

13.2 Bilangan anak yang telah meninggal dunia (jika ada): _____ Umur: _____

13.3 Jika lelaki, nyatakan bilangan isteri: _____

13.4 Bagi ibu-ibu, nyatakan tempat anak dilahirkan:

☐ Rumah sendiri ☐ Rumah bidan ☐ Klinik/pusat perubatan

13.5 Siapakah yang membantu kelahiran anak anda?

☐ Bidan ☐ Jururawat/Doktor

14. Sejarah Pemakanan / Diet

14.1 Berapa kalikah anda makan dalam sehari? _____

14.2 Tandakan apa yang anda makan dan berapa kerap anda memakannya seminggu?

A. Nasi _____
B. Roti _____
C. Ikan/makanan laut _____
D. Sayur _____
E. Ubi _____
F. Daging _____
G. Jika lain lain, nyatakan _____

14.3 Kekerapan anda

	Tidak Pernah	Jarang-jarang	Kadang-kadang	Kerap-kali	Selalu
Makan sarapan					
Makan tengahari					
Makan malam					
Makan kuih muih diantara waktu makan					
Minum minuman bergas					
Makan makanan segera seperti burger					
Makan makanan bersantan					
Makan sayuran					
Makan buah-buahan					
Minum minuman beralkohol					
Makan semasa menonton TV					
Makan sireh					
Makan pianang					
Berjalan kaki					
Berkebun					
Berbasikal					

15. Adakah anda terdedah kepada matahari .....jam/sehari.

Terimakasih atas kerjasama anda.

Physical Examination

Height (m)	Weight (kg)	BMI

Waist Circumference (cm)	Pulse rate (bpm)

Systolic blood pressure **Diastolic blood pressure**

Time	Reading [mmHg]	Mean [mmHg]	Time	Reading [mmHg]	Mean [mmHg]
1			1		
2			2		
3			3		

HbA1c (mmol/mol)		Cholesterol (mg/dL)	
HbA1C (%)		LDL	
		HDL	
		TG	
		TC/HDL	

4

TO BE FILLED BY RESEARCH MEDICAL OFFICER / SPECIALIST

SEJARAH PERUBATAN [Di isi oleh peg.kaji selidik]

1. Penyakit yang telah dikenalpasti

☐ 1. Diabetes mellitus (jika ya nyatakan jangkamasa) _____

☐ 2. Darah tinggi (jika ya nyatakan jangkamasa) _____

☐ 3. Penyakit jantung koronori

☐ 4. Asma

☐ 5. TB

☐ 6. Lain-lain _____

2. Adakah anda pernah ke klinik /pusat kesihatan?

☐ Ya ☐ Tidak, jika tidak, siapakah yang anda jumpa jika sakit_____

3. Ubat-ubatan yang sedang diambil

1. .....
2. .....
3. .....
4. .....
5. .....
6. .....
7. .....
8. .....

4. Sejarah penyakit keluarga [ibu/bapa, adik beradik, anak]

☐ 1 Kencing manis ☐ 2 Darah tinggi ☐ 3 Sakit jantung koronari

☐ 4 Asma ☐ 5 TB ☐ 6 Lain-lain.....

5. Sejarah merokok

☐ 1 Masih merokok ☐ 2 Sudah berhenti ☐ 3 Tidak pernah merokok

☐ 4 Perokok pasif

5.1 Jika masih merokok a) umur mula merokok b) batang/hari

5.2 Jika telah berhenti - dari tahun / bulan ☐

6. Adakah anda meminum alkohol/arak? ☐ Ya ☐ Tidak

6.1 Jika ya, berapa kerap? _____

7. Lebih berat badan dan gemuk dalam keluarga [ibu/bapa, adik beradik, anak]

☐ 1 Ya Nyatakan ☐ a) ibu/bapa ☐ b) adik beradik ☐ c) anak

☐ 2 Tiada

Untuk wanita

8. Kitaran haid

☐ 1 Tetap ☐ 2 Tidak tetap ☐ 3 berhenti haid<6bln ☐ 4 (menopos)berhenti haid > 6bln

9. Sejarah amalan perancang keluarga

☐ 1 Tidak pernah ☐ 2 Pernah ☐ 3 Sedang mengamalkan

Jenis perancang keluarga yang sedang diamalkan atau pernah digunakan. Sila nyatakan tahun dan jangkamasa penggunaan jika pernah guna sebelum ini

1. .....
2. .....
3. .....
4. .....

10. Sejarah diabetes semasa mengandung

☐ 1 Ya Tahun diagnosa ..... ☐ 2 Tidak

DIABETES

- Is the subject on any medication for diabetes?
  ↳ ☐ No ☐ Diet only ☐ Glucose lowering agent[specify below]

☐ Insulin
☐ Meglitinides
☐ Biguanides
☐ Sulphonylureas
☐ Thiazolidinediones
☐ Incretin- based medications
☐ Others
☐ Don't know/Unspecified

- Has the subject been informed of any diabetes-related complication by his/her doctor?
  ↳ If Yes, specify

☐ Eye **specify** ☐ Laser treatment ☐ Blindness

☐ Kidney **specify** ☐ Dialysis

☐ Foot ulcer

☐ Amputation

☐ Nerve

☐ Erectile dysfunction

- Is the subject taking any traditional medication for diabetes?
  ☐ Yes , please specify_____ ☐ No

DYSLIPIDEMIA

- Has the patient been told to have abnormal lipid level? ☐ Yes ☐ No
- Is the subject on any lipid-lowering medication? ☐ Yes ☐ No
  ↳ If Yes, please specify:

☐ Statins
☐ Fibrates
☐ Ezetimibe
☐ Nicotinic Acid
☐ Resins
☐ Other
☐ Don't know/Unspecified

WEIGHT CONTROL

- Is the subject on any weight-reducing agent? ☐ Yes ☐ No
  ↳ If Yes, please specify below
- Has the subject ever joined any weight reduction program? ☐ Yes ☐ No
- Has the subject tried to reduce weight? ☐ Yes ☐ No

☐ Orlistat
☐ Sibutramine
☐ Traditional/Alternative medicine
☐ Other → Specify:
..... Medication (generic name)
..... Medication (generic name)
..... Medication (generic name)

TRADITIONAL MEDICINE

Is the subject taking any traditional medicine? ☐ Yes ☐ No ☐ Not applicable

If yes, please state reason

- ☐ Beauty Related
- ☐ Sexual Prowess
- ☐ Skin & Arthritis
- ☐ General Wellbeing

HEALTH SUPPLEMENT

- ☐ Vitamins
- ☐ Others _____

FAMILY PLANNING AND HRT

	Year of use	Duration of use
☐ Oral Contraceptive Pill		
☐ Depoprovera		
☐ Implant 'Norplant'		
☐ Hormone Replacement Therapy		
☐ Others **Specify** ----------		

[To be filled by medical officer – please tick √ if present]

- ☐ Xanthelasma
- ☐ Facial Hirsutism (for female only)
- ☐ Acanthosis nigricans

Name of examiner:

B. Taraf sosio-ekonomi, penjagaan kebersihan dan bekalan air

A. Sumber bekalan air

1 = Air paip kerajaan
2 = Sungai
3 = Perigi
4 = Air hujan
* Lebih daripada satu jawapan dibenarkan ☐

Adakah anda memasak air minuman?
1= Ya
2= Tidak ☐

B. Penjagaan kebersihan

Ada tandas dalam rumah kah?
1 = Ya
2 = Tidak ☐

Jenis kemudahan tandas
1 = Terdapat tandas bilas
2 = Tidak terdapat tandas bilas ☐

Di manakah anda membuang air besar?
1 = Tandas bilas/pump
2 = Semak/belukar
3 = Tandas curah
4 = sungai ☐

Di manakah anda mandi?
1= Sungai
2= Tandas
3= Lain-lain (Sila nyatakan) ☐

Di manakah anda cuci pakaian?
1= Sungai
2= Tandas
3= Lain-lain (Sila nyatakan) ☐

Jenis bekas simpanan air
1= Botol plastik
2= Baldi plastik
3= Tempayan atau pasu
4= Perigi
5= Tak ada
6= Lain-lain (Sila nyatakan) ☐

Adakah anda tutup bekas simpanan air?
1= Ya
2= Tidak ☐

C. Pengambilan makanan
1=Ya 2=Tidak

Adakah anda makan
1= Daging yang kurang masak
2= Hidangan laut yang kurang masak
3= Sayur-sayuran dan buah-buahan/ulam mentah ☐

D. Pembuangan sampah
1 = Tempat yang tertentu
2 = Merata-rata ☐

E. Binatang peliharaan / binatang ternakan
Adakah anda memiliki binatang peliharaan atau binatang ternakan?
1 = Ya
2 = Tidak ☐

Jika ada, jawab soalan yang berikut

Apakah binatang itu?
1= Anjing (Bilangan:)
2= Kucing
3= Ayam
4= Itik ☐
5= Lain-lain (Sila nyatakan)

Adakah anda rapat atau bermain dengan binatang peliharaan atau ternakan ini?
1 = Ya
2 = Tidak ☐

Adakah anda asingkan bekas makanan dan minuman binatang itu daripada bekas makanan keluarge?
1= Ya
2= Tidak ☐

Di manakah anda memelihara binatang tersebut?
1= Di dalam kandang
2= Bebas berkeliaran ☐

Adakah anda mandikan binatang pemeliharaan tersebut?
1= Ya
2= Tidak ☐

C. Tabiat peribadi dan penjagaan kebersihan

Pilihan jawapan untuk soalan (a) hingga (i) **KECUALI (g)**

1 = Ya
2 = Tidak

a) Adakah anda makan menggunakan tangan? ☐
b) Adakah anda mandi sekurang-kurangnya sekali sehari? ☐
c) Adakah anda menukar pakaian anda sekurang-kurangnya sekali sehari? ☐
d) Adakah anda memakai kasut / selipar semasa keluar rumah? ☐
e) Adakah anda mencuci tangan sebelum makan selepas bermain dengan tanah? ☐
f) Adakah anda mencuci tangan sebelum memasak? ☐
g) Adakah anda mencuci tangan selepas membuang air besar? ☐
h) Adakah anda mencuci tangan selepas bermain dengan binatang? ☐
i) Aktiviti kanak-kanak
   i) Berjalan tanpa kasut ☐
   ii) Bermain dengan tanah ☐

D. Status Kesihatan

Adakah anda mengambil sebarang antibiotic atau ubat dalam 6 bulan yang lalu?
1 = Ya
2 = Tidak ☐

Jika ya, nama ubat:_____

Adakah anda mengambil sebarang "iron supplement" dalam 12 bulan yang lalu?
1= Ya
2= Tidak ☐

Jika ya, dari mana?_____

Adakah anda mengambil sebarang ubat antihelmetik dalam 6 bulan yang lalu?
1= Ya
2= Tidak ☐
Jika ya, dari mana?_____

Adakah anda mengambil sebarang makanan probiotic, contohnya tapai (ubi kayu/pulut) , yoghurt, minuman yoghurt, susu kultur (vitagen atau yakult) dalam 6 bulan yang lalu?
1= Ya 2= Tida ☐

Example of consent form

JH19001

MONASH University

Borang Kelzinan

Projek ID: 11794, 18149 & 17859

Projek:
(i) Penerokaan dan pencirian Microbiomes dalam kalangan pemburu-pengumpul Orang Asli di Peninsula Malaysia;
(ii) Kajian Microbiome dan Genetik Olfaktori dalam kalangan Orang Asli di Malaysia;
(iii) Penyiasatan kepelbagaian genetik nematode berkenaan dengan usus dan gen tindak balas imun perumah dalam kalangan Orang Asli di Peninsula Malaysia

Ketua penyelidik: Profesor Dr. Maude E. Phipps & Profesor Madya Qasim Ayub

Saya telah dijemput untuk menyertai kajian penyelidikan Monash University Malaysia yang bertajuk seperti dinyatakan di atas. Saya telah membaca dan memahami tujuan, prosedur dan risiko yang mungkin dalam kajian ini. Saya, secara sukarela, bersetuju untuk menyertai kajian penyelidikan tersebut.

Saya bersetuju kepada yang berikut:	Ya	Tidak
1. Menyertai kajian penyelidikan dengan memberi maklumat dan sampel kajian: air liur, najis, darah dan swab hidung.	☐	☐
2. Maklumat yang dihasilkan dalam kajian ini boleh digunakan dalam kajian selanjutnya dan aktiviti penyelidikan.	☐	☐
3. Untuk diberikan maklumat lanjutan sekiranya terdapat sebarang penemuan yang memberi kesan kepada kesihatan saya	☐	☐
4. Foto dan video saya boleh digunakan untuk tujuan dokumentasi dan penyelidikan	☐	☐

Nama Peserta: _____

No. Kad Pengenalan Peserta: _____

Tandatangan/Cap tangan Peserta: _____

Tarikh : 27 July 2019.

Page 1 of 1

Appendix V

MUHREC Approval

Monash University Human Research Ethics Committee

Approval Certificate

This is to certify that the project below was considered by the Monash University Human Research Ethics Committee. The Committee was satisfied that the proposal meets the requirements of the *National Statement on Ethical Conduct in Human Research* and has granted approval.

Project Number: 11794
Project Title: Exploration and Characterization of the microbiomes within hunters gatherers and other indigenous communities in Malaysia
Chief Investigator: Professor Maude Phipps
Approval Date: 02/02/2018
Expiry Date: 02/02/2023

Terms of approval - failure to comply with the terms below is in breach of your approval and the *Australian Code for the Responsible Conduct of Research*.

1. The Chief Investigator is responsible for ensuring that permission letters are obtained, if relevant, before any data collection can occur at the specified organisation.
2. Approval is only valid whilst you hold a position at Monash University.
3. It is responsibility of the Chief Investigator to ensure that all investigators are aware of the terms of approval and to ensure the project is conducted as approved by MUHREC.
4. You should notify MUHREC immediately of any serious or unexpected adverse effects on participants or unforeseen events affecting the ethical acceptability of the project.
5. The Explanatory Statement must be on Monash letterhead and the Monash University complaints clause must include your project number.
6. Amendments to approved projects including changes to personnel must not commence without written approval from MHUREC.
7. Annual Report - continued approval of this project is dependent on the submission of an Annual Report.
8. Final Report - should be provided at the conclusion of the project. MUHREC should be notified if the project is discontinued before the expected completion date.
9. Monitoring - project may be subject to an audit or any other form of monitoring by MUHREC at any time.
10. Retention and storage of data - The Chief Investigator is responsible for the storage and retention of the original data pertaining to the project for a minimum period of five years.

Thank you for your assistance.

Professor Nip Thomson

Chair, MUHREC

CC: Assoc Professor umadevi palanisamy, Dr Amreeta Dhanoa, Dr Farhang Aghakhanian, Ms Li Yeo, Professor Khalid Kadir, Dr Ho Loon Shin

List of approved documents:

Document Type	File Name	Date	Version
Supporting Documentation	Exp_statement_MOA_121117muhrec	12/11/2017	01
Explanatory Statement	Exp_statement_MOA_121117muhrec	13/11/2017	1
Consent Form	consent_form_edit_MOA_121117muhrec	13/11/2017	1
Supporting Documentation	MicrobiomeQuestionnaire -ENG (1)	18/12/2017	1
Supporting Documentation	MicrobiomeQuestionnaire -ENG (1)	01/01/2018	1
Questionnaires / Surveys	MicrobiomeQuestionnaire -ENG (1)	01/01/2018	2

References

AB RAHMAN, N., TENG, C. L. & SIVASAMPU, S. 2016. Antibiotic prescribing in public and private practice: a cross-sectional study in primary care clinics in Malaysia. *BMC Infectious Diseases,* 16**,** 208.

ADLER, C. J., DOBNEY, K., WEYRICH, L. S., KAIDONIS, J., WALKER, A. W., HAAK, W., BRADSHAW, C. J., TOWNSEND, G., SOŁTYSIAK, A. & ALT, K. W. 2013. Sequencing ancient calcified dental plaque shows changes in oral microbiota with dietary shifts of the Neolithic and Industrial revolutions. *Nature genetics,* 45**,** 450.

AGHAKHANIAN, F., WONG, C., TAN, J., YEO, L., RAMADAS, A., EDO, J., HOH, B., KHALID, B. & PHIPPS, M. 2019. Metabolic syndrome and cardiometabolic risk factors among indigenous Malaysians. *Public health,* 176**,** 106-113.

AHMED, M. O., BAPTISTE, KEITH E. 2018. Vancomycin-Resistant Enterococci: A Review of Antimicrobial Resistance Mechanisms and Perspectives of Human and Animal Health. *Microbial Drug Resistance,* 24**,** 590-606.

ALAVI, S., MITCHELL, J. D., CHO, J. Y., LIU, R., MACBETH, J. C. & HSIAO, A. 2020. Interpersonal Gut Microbiome Variation Drives Susceptibility and Resistance to Cholera Infection. *Cell,* 181**,** 1533-1546.e13.

ALCOCK, B. P., RAPHENYA, A. R., LAU, T. T. Y., TSANG, K. K., BOUCHARD, M., EDALATMAND, A., HUYNH, W., NGUYEN, A.-L. V., CHENG, A. A., LIU, S., MIN, S. Y., MIROSHNICHENKO, A., TRAN, H.-K., WERFALLI, R. E., NASIR, J. A., OLONI, M., SPEICHER, D. J., FLORESCU, A., SINGH, B., FALTYN, M., HERNANDEZ-KOUTOUCHEVA, A., SHARMA, A. N., BORDELEAU, E., PAWLOWSKI, A. C., ZUBYK, H. L., DOOLEY, D., GRIFFITHS, E., MAGUIRE, F., WINSOR, G. L., BEIKO, R. G., BRINKMAN, F. S. L., HSIAO, W. W. L., DOMSELAAR, G. V. & MCARTHUR, A. G. 2019. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Research,* 48**,** D517-D525.

ALHAJ, N., MARIANA, N., RAHA, A. & ISHAK, Z. 2007. Prevalence of antibiotic resistance among Escherichia coli from different sources in Malaysia. *International Journal of Poultry Science,* 6**,** 293-297.

ALI, O., TAN, T. T., SAKINAH, O., KHALID, B. A. K., WU, L. L. & NG, M. L. 1993. Prevalence of NIDDM and Impaired Glucose Tolerance in Aborigines and Malays in Malaysia and Their Relationship to Sociodemographic, Health, and Nutritional Factors. *Diabetes Care,* 16**,** 68-75.

ALIYU, A. B., SALEHA, A. A., JALILA, A. & ZUNITA, Z. 2016. Risk factors and spatial distribution of extended spectrum β-lactamase-producing- Escherichia coli at retail poultry meat markets in Malaysia: a cross-sectional study. *BMC Public Health,* 16**,** 699.

ALMEIDA, A., MITCHELL, A. L., BOLAND, M., FORSTER, S. C., GLOOR, G. B., TARKOWSKA, A., LAWLEY, T. D. & FINN, R. D. 2019. A new genomic blueprint of the human gut microbiota. *Nature,* 568**,** 499-504.

ALMEIDA, A., NAYFACH, S., BOLAND, M., STROZZI, F., BERACOCHEA, M., SHI, Z. J., POLLARD, K. S., SAKHAROVA, E., PARKS, D. H., HUGENHOLTZ, P., SEGATA, N., KYRPIDES, N. C. & FINN, R. D. 2020. A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nature Biotechnology*.

ALMONACID, D. E., KRAAL, L., OSSANDON, F. J., BUDOVSKAYA, Y. V., CARDENAS, J. P., BIK, E. M., GODDARD, A. D., RICHMAN, J. & APTE, Z. S. 2017. 16S rRNA gene sequencing and healthy reference ranges for 28 clinically relevant microbial taxa from the human gut microbiome. *PLOS ONE,* 12**,** e0176555.

AMEDEI, A., CODOLO, G., DEL PRETE, G., DE BERNARD, M. & D'ELIOS, M. M. 2010. The effect of Helicobacter pylori on asthma and allergy. *Journal of asthma and allergy,* 3**,** 139-147.

AMMAM, F., MEZIANE-CHERIF, D., MENGIN-LECREULX, D., BLANOT, D., PATIN, D., BONECA, I. G., COURVALIN, P., LAMBERT, T. & CANDELA, T. 2013. The functional vanGCd cluster of Clostridium difficile does not confer vancomycin resistance. *Mol Microbiol,* 89**,** 612-25.

ANAND, R., DHINGRA, C., PRASAD, S. & MENON, I. 2014. Betel nut chewing and its deleterious effects on oral cavity. *Journal of cancer research and therapeutics,* 10**,** 499.

ANDERSON, M. J. 2017. Permutational Multivariate Analysis of Variance (PERMANOVA). *Wiley StatsRef: Statistics Reference Online.*

ANDERSON, M. J. & WILLIS, T. J. 2003. CANONICAL ANALYSIS OF PRINCIPAL COORDINATES: A USEFUL METHOD OF CONSTRAINED ORDINATION FOR ECOLOGY. *Ecology,* 84**,** 511-525.

ANUAR, T. S., SALLEH, F. M. & MOKTAR, N. 2014. Soil-transmitted helminth infections and associated risk factors in three Orang Asli tribes in Peninsular Malaysia. *Scientific reports,* 4**,** 1-7.

ANUURAD, E., SHIWAKU, K., NOGI, A., KITAJIMA, K., ENKHMAA, B., SHIMONO, K. & YAMANE, Y. 2003. The new BMI criteria for asians by the regional office for the western pacific region of WHO are suitable for screening of overweight to prevent metabolic syndrome in elder Japanese workers. *J Occup Health,* 45**,** 335-43.

ARIFFIN, H., NAVARATNAM, P., MOHAMED, M., ARASU, A., ARIFFIN ABDULLAH, W., LEE LEE, C. & HAI PENG, L. 2000. Ceftazidime-resistant Klebsiella pneumoniae bloodstream infection in children with febrile neutropenia. *International Journal of Infectious Diseases,* 4**,** 21-25.

ARRAM, E. O., SHAHIN, D. A. & SHERIF, M. M. 2012. Asthma is inversely associated with Helicobacter pylori status. *Egyptian Journal of Chest Diseases and Tuberculosis,* 61**,** 41-45.

BAER, A. 1999. *Health, Disease and Survival: A Biomedical and Genetic Analysis of the Orang Asli of Malaysia,* Subang Jaya, Malaysia, Center for Orang Asli Concerns.

BAER, A. E., KAREN; GIANNO, ROSEMARY; HOWELL, SIGNE; NOWAK, BARBARA S., VAN DER SLUYS, CORNELIA 2006. *Orang Asli Women of Malaysia: Perceptions, Situations and Aspirations,* Subang Jaya, Malaysia, Center for Orang Asli Concerns.

BAJAJ, J. S., BETRAPALLY, N. S., HYLEMON, P. B., HEUMAN, D. M., DAITA, K., WHITE, M. B., UNSER, A., THACKER, L. R., SANYAL, A. J. & KANG, D. J. 2015. Salivary microbiota reflects changes in gut microbiota in cirrhosis with hepatic encephalopathy. *Hepatology,* 62**,** 1260-1271.

BARTHOLOMEW, C. V., ZAINIR, M. I., ZALIPAH, M. N., HUSIN, M. H. & ABDULLAH, M. T. 2021. Wildlife Hunting Practices by the Indigenous People of Terengganu, Peninsular Malaysia. *Resource Use and Sustainability of Orang Asli: Indigenous Communities in Peninsular Malaysia***,** 137.

BENAHMED, A. G., GASMI, A., DOŞA, A., CHIRUMBOLO, S., MUJAWDIYA, P. K., AASETH, J., DADAR, M. & BJØRKLUND, G. 2020. Association between the gut and oral microbiome with obesity. *Anaerobe***,** 102248.

BENÍTEZ-PÁEZ, A., GÓMEZ DEL PUGAR, E. M., LÓPEZ-ALMELA, I., MOYA-PÉREZ, Á., CODOÑER-FRANCH, P. & SANZ, Y. 2020. Depletion of Blautia Species in the Microbiota of Obese Children Relates to Intestinal Inflammation and Metabolic Phenotype Worsening. *mSystems,* 5**,** e00857-19.

BENNO, Y., SUZUKI, K., SUZUKI, K., NARISAWA, K., BRUCE, W. R. & MITSUOKA, T. 1986. Comparison of the Fecal Microflora in Rural Japanese and Urban Canadians. *MICROBIOLOGY and IMMUNOLOGY,* 30**,** 521-532.

BEZABIH, Y. M., SABIITI, W., ALAMNEH, E., BEZABIH, A., PETERSON, G. M., BEZABHE, W. M. & ROUJEINIKOVA, A. 2020. The global prevalence and trend of human intestinal carriage of ESBL-producing Escherichia coli in the community. *Journal of Antimicrobial Chemotherapy,* 76**,** 22-29.

BIEDENBACH, D. J., LEWIS, M. T. & JONES, R. N. 1999. In vitro evaluation of cefepime and other broad-spectrum beta-lactams for isolates in Malaysia and Singapore medical centers. The Malaysia/Singapore Antimicrobial Resistance Study Group. *Diagn Microbiol Infect Dis,* 35**,** 277-83.

BLASER, M. J. 2008. Disappearing Microbiota: Helicobacter pylori Protection against Esophageal Adenocarcinoma. *Cancer Prevention Research,* 1**,** 308-311.

BLASER, M. J., CHEN, Y. & REIBMAN, J. 2008. Does Helicobacter pylori protect against asthma and allergy? *Gut,* 57**,** 561-567.

BLOOMFIELD, S. F., ROOK, G. A., SCOTT, E. A., SHANAHAN, F., STANWELL-SMITH, R. & TURNER, P. 2016. Time to abandon the hygiene hypothesis: new perspectives on allergic disease, the human

microbiome, infectious disease prevention and the role of targeted hygiene. *Perspectives in public health,* 136**,** 213-224.

BOLTON, J. M. 1972. Food taboos among the Orang Asli in West Malaysia: a potential nutritional hazard. *The American Journal of Clinical Nutrition,* 25**,** 789-799.

BOONYASIRI, A., TANGKOSKUL, T., SEENAMA, C., SAIYARIN, J., TIENGRIM, S. & THAMLIKITKUL, V. 2014. Prevalence of antibiotic resistant bacteria in healthy adults, foods, food animals, and the environment in selected areas in Thailand. *Pathogens and global health,* 108**,** 235-245.

BORHANUDDIN, B., MOHD NAWI, A., SHAH, S. A., ABDULLAH, N., SYED ZAKARIA, S. Z., KAMARUDDIN, M. A., VELU, C. S., ISMAIL, N., ABDULLAH, M. S., AHMAD KAMAT, S., AWANG, A., HAMID, M. A. & JAMAL, R. 2018. 10-Year Cardiovascular Disease Risk Estimation Based on Lipid Profile-Based and BMI-Based Framingham Risk Scores across Multiple Sociodemographic Characteristics: The Malaysian Cohort Project. *The Scientific World Journal,* 2018**,** 2979206.

BORRIELLO, P., DRASAR, B. & TOMKINS, A. 1978. Diet and faecal flora: a comparison of rural Northern Nigeria and London, UK. *Proc Nutr Soc,* 37**,** 40a.

BRENNAN, C. A. & GARRETT, W. S. 2019. Fusobacterium nucleatum — symbiont, opportunist and oncobacterium. *Nature Reviews Microbiology,* 17**,** 156-166.

BREUREC, S., GUILLARD, B., HEM, S., BRISSE, S., DIEYE, F. B., HUERRE, M., OUNG, C., RAYMOND, J., TAN, T. S., THIBERGE, J.-M., VONG, S., MONCHY, D. & LINZ, B. 2011. Evolutionary history of Helicobacter pylori sequences reflect past human migrations in Southeast Asia. *PloS one,* 6**,** e22058-e22058.

BROWNE, H. P., FORSTER, S. C., ANONYE, B. O., KUMAR, N., NEVILLE, B. A., STARES, M. D., GOULDING, D. & LAWLEY, T. D. 2016. Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature,* 533**,** 543-546.

BUELOW, E., GONZALEZ, T. B., VERSLUIS, D., OOSTDIJK, E. A., OGILVIE, L. A., VAN MOURIK, M. S., OOSTERINK, E., VAN PASSEL, M. W., SMIDT, H., D'ANDREA, M. M., DE BEEN, M., JONES, B. V., WILLEMS, R. J., BONTEN, M. J. & VAN SCHAIK, W. 2014. Effects of selective digestive decontamination (SDD) on the gut resistome. *J Antimicrob Chemother,* 69**,** 2215-23.

BUI, D., BROWN, H. E., HARRIS, R. B. & OREN, E. 2016. Serologic Evidence for Fecal-Oral Transmission of Helicobacter pylori. *The American journal of tropical medicine and hygiene,* 94**,** 82-88.

BUKIN, Y. S., GALACHYANTS, Y. P., MOROZOV, I. V., BUKIN, S. V., ZAKHARENKO, A. S. & ZEMSKAYA, T. I. 2019. The effect of 16S rRNA region choice on bacterial community metabarcoding results. *Scientific Data,* 6**,** 190007.

BUTLER, J., CRENGLE, S, CHEEK, JE., LEACH, AJ., LENNON, D., O'BRIEN, KL., ET AL 2001. Emerging Infectious Diseases Among Indigenous Peoples. *Emerging Infectious Diseases,* 7(7)**,** 554-555.

CALLAHAN, B. J., MCMURDIE, P. J., ROSEN, M. J., HAN, A. W., JOHNSON, A. J. A. & HOLMES, S. P. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods,* 13**,** 581-583.

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

CARR, V. R., WITHERDEN, E. A., LEE, S., SHOAIE, S., MULLANY, P., PROCTOR, G. B., GOMEZ-CABRERO, D. & MOYES, D. L. 2020. Abundance and diversity of resistomes differ between healthy human oral cavities and gut. *Nature Communications,* 11**,** 693.

CATALDI, M. & MONTAGNA, C. 1944. The intestinal microflora in infants fed on human milk. *Revista de la Asociacion Argentina de Dietologia,* 2**,** 36-40.

CENA, J. A. D., ZHANG, J., DENG, D., DAMÉ-TEIXEIRA, N. & DO, T. 2021. Low-Abundant Microorganisms: The Human Microbiome's Dark Matter, a Scoping Review. *Frontiers in Cellular and Infection Microbiology,* 11.

CHALLA, S. & NEELAPU, N. R. R. 2018. Quorum Sensing in Helicobacter pylori: Role of Biofilm and Its Implications for Antibiotic Resistance and Immune Evasion. *In:* PALLAVAL VEERA, B. (ed.)

Implication of Quorum Sensing System in Biofilm Formation and Virulence. Singapore: Springer Singapore.
CHATTOPADHYAY, I. & PANDA, M. 2019. Recent trends of saliva omics biomarkers for the diagnosis and treatment of oral cancer. *Journal of oral biosciences,* 61**,** 84-94.
CHAUMEIL, P. A., MUSSIG, A. J., HUGENHOLTZ, P. & PARKS, D. H. 2019. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics,* 36**,** 1925-7.
CHEN, C., HEMME, C., BELENO, J., SHI, Z. J., NING, D., QIN, Y., TU, Q., JORGENSEN, M., HE, Z., WU, L. & ZHOU, J. 2018. Oral microbiota of periodontal health and disease and their changes after nonsurgical periodontal therapy. *The ISME Journal,* 12**,** 1210-1224.
CHEN, W., JIANG, Q., YAN, G. & YANG, D. 2020. The oral microbiome and salivary proteins influence caries in children aged 6 to 8 years. *BMC Oral Health,* 20**,** 295.
CHEONG, Y. M., LIM, V. K., JEGATHESAN, M. & SULEIMAN, A. B. 1994. Antimicrobial resistance in 6 Malaysian general hospitals. *Med J Malaysia,* 49**,** 317-26.
CHING, O. O. & LEONG, T. F. 2011. Orang Asli and wildlife conservation in the Belum-Temengor forest complex, Malaysia. *TRAFFIC Bull,* 23**,** 94-104.
CHONG, C. W., AHMAD, A. F., LIM, Y. A. L., TEH, C. S. J., YAP, I. K. S., LEE, S. C., CHIN, Y. T. & CHUA, K. H. 2015. Effect of ethnicity and socioeconomic variation to the gut microbiota composition among pre-adolescent in Malaysia. *Scientific reports,* 5**,** 1-12.
CHUA, E. G., LOKE, M. F., GUNALETCHUMY, S. P., GAN, H. M., THEVAKUMAR, K., TAY, C. Y., YOUNG, S., AYE, T. T., MAW, W. W. & AYE, M. M. 2019. The influence of modernization and disease on the gastric microbiome of orang asli, myanmars and modern malaysians. *Microorganisms,* 7**,** 174.
CLAUS, S. P. 2019. The Strange Case of Prevotella copri: Dr. Jekyll or Mr. Hyde? *Cell Host & Microbe,* 26**,** 577-578.
CLEARY, D. W., MORRIS, D. E., ANDERSON, R. A., JONES, J., ALATTRAQCHI, A. G., RAHMAN, N. I. A., ISMAIL, S., RAZALI, M. S., AMIN, R. M. & ABD AZIZ, A. 2021. The upper respiratory tract microbiome of indigenous Orang Asli in north-eastern Peninsular Malaysia. *NPJ biofilms and microbiomes,* 7**,** 1-11.
COIL, D. A., JOSPIN, G., DARLING, A. E., WALLIS, C., DAVIS, I. J., HARRIS, S., EISEN, J. A., HOLCOMBE, L. J. & O'FLYNN, C. 2019. Genomes from bacteria associated with the canine oral cavity: a test case for automated genome-based taxonomic assignment. *PloS one,* 14**,** e0214354.
COMMITTEE, E. S. 2011. Statistical significance and biological relevance. *EFSA Journal,* 9**,** 2372.
COSSEAU, C., ROMANO-BERTRAND, S., DUPLAN, H., LUCAS, O., INGRASSIA, I., PIGASSE, C., ROQUES, C. & JUMAS-BILAK, E. 2016. Proteobacteria from the human skin microbiota: Species-level diversity and hypotheses. *One Health,* 2**,** 33-41.
COSTEA, P. I., ZELLER, G., SUNAGAWA, S., PELLETIER, E., ALBERTI, A., LEVENEZ, F., TRAMONTANO, M., DRIESSEN, M., HERCOG, R., JUNG, F.-E., KULTIMA, J. R., HAYWARD, M. R., COELHO, L. P., ALLEN-VERCOE, E., BERTRAND, L., BLAUT, M., BROWN, J. R. M., CARTON, T., COOLS-PORTIER, S., DAIGNEAULT, M., DERRIEN, M., DRUESNE, A., DE VOS, W. M., FINLAY, B. B., FLINT, H. J., GUARNER, F., HATTORI, M., HEILIG, H., LUNA, R. A., VAN HYLCKAMA VLIEG, J., JUNICK, J., KLYMIUK, I., LANGELLA, P., LE CHATELIER, E., MAI, V., MANICHANH, C., MARTIN, J. C., MERY, C., MORITA, H., O'TOOLE, P. W., ORVAIN, C., PATIL, K. R., PENDERS, J., PERSSON, S., PONS, N., POPOVA, M., SALONEN, A., SAULNIER, D., SCOTT, K. P., SINGH, B., SLEZAK, K., VEIGA, P., VERSALOVIC, J., ZHAO, L., ZOETENDAL, E. G., EHRLICH, S. D., DORE, J. & BORK, P. 2017. Towards standards for human fecal sample processing in metagenomic studies. *Nature Biotechnology,* 35**,** 1069-1076.
COURVALIN, P. 1994. Transfer of antibiotic resistance genes between gram-positive and gram-negative bacteria. *Antimicrob Agents Chemother,* 38**,** 1447-51.
CROSS, K. L., CAMPBELL, J. H., BALACHANDRAN, M., CAMPBELL, A. G., COOPER, S. J., GRIFFEN, A., HEATON, M., JOSHI, S., KLINGEMAN, D., LEYS, E., YANG, Z., PARKS, J. M. & PODAR, M. 2019.

Targeted isolation and cultivation of uncultivated bacteria by reverse genomics. *Nature Biotechnology,* 37**,** 1314-1321.

D'COSTA, V. M., KING, C. E., KALAN, L., MORAR, M., SUNG, W. W. L., SCHWARZ, C., FROESE, D., ZAZULA, G., CALMELS, F., DEBRUYNE, R., GOLDING, G. B., POINAR, H. N. & WRIGHT, G. D. 2011. Antibiotic resistance is ancient. *Nature,* 477**,** 457-461.

DA SILVA, S. F., REIS, I. B., MONTEIRO, M. G., DIAS, V. C., MACHADO, A. B. F., DA SILVA, V. L. & DINIZ, C. G. 2021. Influence of Human Eating Habits on Antimicrobial Resistance Phenomenon: Aspects of Clinical Resistome of Gut Microbiota in Omnivores, Ovolactovegetarians, and Strict Vegetarians. *Antibiotics (Basel, Switzerland),* 10**,** 276.

DANIEL, D. S., LEE, S. M., DYKES, G. A. & RAHMAN, S. 2015. Public health risks of multiple-drug-resistant Enterococcus spp. in Southeast Asia. *Applied and environmental microbiology,* 81**,** 6090-6097.

DANIEL, D. S., LEE, S. M., GAN, H. M., DYKES, G. A. & RAHMAN, S. 2017. Genetic diversity of Enterococcus faecalis isolated from environmental, animal and clinical sources in Malaysia. *Journal of Infection and Public Health,* 10**,** 617-623.

DAVENPORT, E. R., MIZRAHI-MAN, O., MICHELINI, K., BARREIRO, L. B., OBER, C. & GILAD, Y. 2014. Seasonal variation in human gut microbiome composition. *PloS one,* 9**,** e90731-e90731.

DAVID, L. A., MAURICE, C. F., CARMODY, R. N., GOOTENBERG, D. B., BUTTON, J. E., WOLFE, B. E., LING, A. V., DEVLIN, A. S., VARMA, Y. & FISCHBACH, M. A. 2014. Diet rapidly and reproducibly alters the human gut microbiome. *Nature,* 505**,** 559-563.

DAVIES, J. & DAVIES, D. 2010. Origins and evolution of antibiotic resistance. *Microbiol. Mol. Biol. Rev.,* 74**,** 417-433.

DAVIES, N. K., O'SULLIVAN, J. M., PLANK, L. D. & MURPHY, R. 2019. Altered gut microbiome after bariatric surgery and its association with metabolic benefits: A systematic review. *Surgery for Obesity and Related Diseases,* 15**,** 656-665.

DE ANGELIS, M., FERROCINO, I., CALABRESE, F. M., DE FILIPPIS, F., CAVALLO, N., SIRAGUSA, S., RAMPELLI, S., DI CAGNO, R., RANTSIOU, K., VANNINI, L., PELLEGRINI, N., LAZZI, C., TURRONI, S., LORUSSO, N., VENTURA, M., CHIEPPA, M., NEVIANI, E., BRIGIDI, P., O'TOOLE, P. W., ERCOLINI, D., GOBBETTI, M. & COCOLIN, L. 2020. Diet influences the functions of the human intestinal microbiome. *Scientific Reports,* 10**,** 4247.

DE FILIPPIS, F., PASOLLI, E., TETT, A., TARALLO, S., NACCARATI, A., DE ANGELIS, M., NEVIANI, E., COCOLIN, L., GOBBETTI, M., SEGATA, N. & ERCOLINI, D. 2019. Distinct Genetic and Functional Traits of Human Intestinal Prevotella copri Strains Are Associated with Different Habitual Diets. *Cell Host & Microbe,* 25**,** 444-453.e3.

DE FILIPPO, C., DI PAOLA, M., RAMAZZOTTI, M., ALBANESE, D., PIERACCINI, G., BANCI, E., MIGLIETTA, F., CAVALIERI, D. & LIONETTI, P. 2017. Diet, Environments, and Gut Microbiota. A Preliminary Investigation in Children Living in Rural and Urban Burkina Faso and Italy. *Front Microbiol,* 8**,** 1979.

DE MORAES, A. C. F., FERNANDES, G. R., DA SILVA, I. T., ALMEIDA-PITITTO, B., GOMES, E. P., PEREIRA, A. D. C. & FERREIRA, S. R. G. 2017. Enterotype May Drive the Dietary-Associated Cardiometabolic Risk Factors. *Frontiers in cellular and infection microbiology,* 7**,** 47-47.

DE VIENNE, D. M. 2016. Lifemap: Exploring the Entire Tree of Life. *PLoS biology,* 14**,** e2001624-e2001624.

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

DENG, L., HOH, B. P., LU, D., FU, R., PHIPPS, M. E., LI, S., NUR-SHAFAWATI, A. R., HATIN, W. I., ISMAIL, E., MOKHTAR, S. S., JIN, L., ZILFALIL, B. A., MARSHALL, C. R., SCHERER, S. W., AL-MULLA, F. & XU, S. 2014. The population genomic landscape of human genetic structure, admixture history and local adaptation in Peninsular Malaysia. *Hum Genet,* 133**,** 1169-85.

DEO, P. N. & DESHMUKH, R. 2019. Oral microbiome: Unveiling the fundamentals. *Journal of oral and maxillofacial pathology : JOMFP,* 23**,** 122-128.

DESAI, M. S., SEEKATZ, A. M., KOROPATKIN, N. M., KAMADA, N., HICKEY, C. A., WOLTER, M., PUDLO, N. A., KITAMOTO, S., TERRAPON, N., MULLER, A., YOUNG, V. B., HENRISSAT, B., WILMES, P., STAPPENBECK, T. S., NÚÑEZ, G. & MARTENS, E. C. 2016. A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. *Cell,* 167**,** 1339-1353.e21.

DEVARAJ, J. 1999. 'Between Myth and Reality: Why are Orang Asli More Prone to Illness?'. *World Conference of Primary Care Physicians, .* Kuching.

DEWHIRST, F. E., CHEN, T., IZARD, J., PASTER, B. J., TANNER, A. C., YU, W. H., LAKSHMANAN, A. & WADE, W. G. 2010. The human oral microbiome. *J Bacteriol,* 192**,** 5002-17.

DILL-MCFARLAND, K. A., TANG, Z.-Z., KEMIS, J. H., KERBY, R. L., CHEN, G., PALLONI, A., SORENSON, T., REY, F. E. & HERD, P. 2019. Close social relationships correlate with human gut microbiota composition. *Scientific Reports,* 9**,** 703.

DOYLE, R. M., O'SULLIVAN, D. M., ALLER, S. D., BRUCHMANN, S., CLARK, T., COELLO PELEGRIN, A., CORMICAN, M., DIEZ BENAVENTE, E., ELLINGTON, M. J., MCGRATH, E., MOTRO, Y., PHUONG THUY NGUYEN, T., PHELAN, J., SHAW, L. P., STABLER, R. A., VAN BELKUM, A., VAN DORP, L., WOODFORD, N., MORAN-GILAD, J., HUGGETT, J. F. & HARRIS, K. A. 2020. Discordant bioinformatic predictions of antimicrobial resistance from whole-genome sequencing data of bacterial isolates: an inter-laboratory study. *Microbial genomics,* 6**,** e000335.

DRAGO, L., ZUCCOTTI, G. V., ROMANÒ, C. L., GOSWAMI, K., VILLAFAÑE, J. H., MATTINA, R. & PARVIZI, J. 2019. Oral–Gut Microbiota and Arthritis: Is There an Evidence-Based Axis? *Journal of clinical medicine,* 8**,** 1753.

DRASAR, B. S., MONTGOMERY, F. & TOMKINS, A. M. 1986. Diet and faecal flora in three dietary groups in rural northern Nigeria. *J Hyg (Lond),* 96**,** 59-65.

DUBOIS, G., GIRARD, C., LAPOINTE, F.-J. & SHAPIRO, B. J. 2017. The Inuit gut microbiome is dynamic over time and shaped by traditional foods. *Microbiome,* 5**,** 151.

EBERHARD, D. M., GARY F.SIMONS, AND CHARLES D. FENNIG (EDS) 2021. Ethnologue: Languages of the World. Dallas, Texas: SIL International.

EDGAR, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics,* 26**,** 2460-2461.

EDMONDS, K. & WILLIAMS, L. 2017. The Role of the Negative Control in Microbiome Analyses. *The FASEB Journal,* 31**,** 940.3-940.3.

EISEMAN, B., SILEN, W., BASCOM, G. S. & KAUVAR, A. J. 1958. Fecal enema as an adjunct in the treatment of pseudomembranous enterocolitis. *Surgery,* 44**,** 854-9.

EISENHOFER, R., MINICH, J. J., MAROTZ, C., COOPER, A., KNIGHT, R. & WEYRICH, L. S. 2019. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends in Microbiology,* 27**,** 105-117.

ELLWANGER, J. H., KULMANN-LEAL, B., KAMINSKI, V. L., VALVERDE-VILLEGAS, J., VEIGA, A. B. G., SPILKI, F. R., FEARNSIDE, P. M., CAESAR, L., GIATTI, L. L. & WALLAU, G. L. 2020. Beyond diversity loss and climate change: Impacts of Amazon deforestation on infectious diseases and public health. *Anais da Academia Brasileira de Ciências,* 92.

ENDICOTT, K. M. E., KAREN L. 2012. *The Headman was a Woman: The Gender Egalitarian Batek of Malaysia,* Subang Jaya, Malaysia, Center for Orang Asli Concerns.

ESBERG, A., BARONE, A., ERIKSSON, L., LIF HOLGERSON, P., TENEBERG, S. & JOHANSSON, I. 2020. Corynebacterium matruchotii Demography and Adhesion Determinants in the Oral Cavity of Healthy Individuals. *Microorganisms,* 8**,** 1780.

ESCOBAR, C., BARRIOS, V., & PÉREZ DE ISLA, L. 2018. Niveles óptimos de colesterol en los pacientes con dislipemia. Revisión sistemática de la evidencia [Optimal cholesterol levels in patients in real-life. A systematic review]. *Semergen*, *44*(1), 42–49.

ESTAKI, M., JIANG, L., BOKULICH, N. A., MCDONALD, D., GONZÁLEZ, A., KOSCIOLEK, T., MARTINO, C., ZHU, Q., BIRMINGHAM, A., VÁZQUEZ-BAEZA, Y., DILLON, M. R., BOLYEN, E., CAPORASO, J. G. & KNIGHT, R. 2020. QIIME 2 Enables Comprehensive End-to-End Analysis of Diverse Microbiome Data and Comparative Studies with Publicly Available Data. *Current Protocols in Bioinformatics,* 70**,** e100.

FELLOWS YATES, J. A., VELSKO, I. M., ARON, F., POSTH, C., HOFMAN, C. A., AUSTIN, R. M., PARKER, C. E., MANN, A. E., NÄGELE, K., ARTHUR, K. W., ARTHUR, J. W., BAUER, C. C., CREVECOEUR, I., CUPILLARD, C., CURTIS, M. C., DALÉN, L., DÍAZ-ZORITA BONILLA, M., DÍEZ FERNÁNDEZ-LOMANA, J. C., DRUCKER, D. G., ESCRIBANO ESCRIVÁ, E., FRANCKEN, M., GIBBON, V. E., GONZÁLEZ MORALES, M. R., GRANDE MATEU, A., HARVATI, K., HENRY, A. G., HUMPHREY, L., MENÉNDEZ, M., MIHAILOVIĆ, D., PERESANI, M., RODRÍGUEZ MORODER, S., ROKSANDIC, M., ROUGIER, H., SÁZELOVÁ, S., STOCK, J. T., STRAUS, L. G., SVOBODA, J., TEßMANN, B., WALKER, M. J., POWER, R. C., LEWIS, C. M., SANKARANARAYANAN, K., GUSCHANSKI, K., WRANGHAM, R. W., DEWHIRST, F. E., SALAZAR-GARCÍA, D. C., KRAUSE, J., HERBIG, A. & WARINNER, C. 2021. The evolution and changing ecology of the African hominid oral microbiome. *Proceedings of the National Academy of Sciences,* 118**,** e2021655118.

FENG, Y., DUAN, Y., XU, Z., LYU, N., LIU, F., LIANG, S. & ZHU, B. 2019. An examination of data from the American Gut Project reveals that the dominance of the genus Bifidobacterium is associated with the diversity and robustness of the gut microbiota. *MicrobiologyOpen,* 8**,** e939-e939.

FERNANDES, A. D., REID, J. N. S., MACKLAIM, J. M., MCMURROUGH, T. A., EDGELL, D. R. & GLOOR, G. B. 2014. Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. *Microbiome,* 2**,** 15.

FERNANDES, A. D., VU, M. T., EDWARD, L.-M., MACKLAIM, J. M. & GLOOR, G. B. 2018. A reproducible effect size is more useful than an irreproducible hypothesis test to analyze high throughput sequencing datasets. *arXiv preprint arXiv:1809.02623*.

FERRIE, S., WEBSTER, A., WU, B., TAN, C. & CAREY, S. 2021. Gastrointestinal surgery and the gut microbiome: a systematic literature review. *European Journal of Clinical Nutrition,* 75**,** 12-25.

FERROCINO, I., DI CAGNO, R., DE ANGELIS, M., TURRONI, S., VANNINI, L., BANCALARI, E., RANTSIOU, K., CARDINALI, G., NEVIANI, E. & COCOLIN, L. 2015. Fecal microbiota in healthy subjects following omnivore, vegetarian and vegan diets: culturable populations and rRNA DGGE profiling. *PloS one,* 10**,** e0128669.

FINNICUM, C. T., BECK, J. J., DOLAN, C. V., DAVIS, C., WILLEMSEN, G., EHLI, E. A., BOOMSMA, D. I., DAVIES, G. E. & DE GEUS, E. J. C. 2019. Cohabitation is associated with a greater resemblance in gut microbiota which can impact cardiometabolic and inflammatory risk. *BMC Microbiology,* 19**,** 230.

FORSLUND, K., SUNAGAWA, S., KULTIMA, J. R., MENDE, D. R., ARUMUGAM, M., TYPAS, A. & BORK, P. 2013. Country-specific antibiotic use practices impact the human gut resistome. *Genome Research,* 23**,** 1163-1169.

FROST, F., KACPROWSKI, T., RÜHLEMANN, M., BANG, C., FRANKE, A., ZIMMERMANN, K., NAUCK, M., VÖLKER, U., VÖLZKE, H., BIFFAR, R., SCHULZ, C., MAYERLE, J., WEISS, F. U., HOMUTH, G. & LERCH, M. M. 2019. Helicobacter pylori infection associates with fecal microbiota composition and diversity. *Scientific Reports,* 9**,** 20100.

GASEVIC, D., FROHLICH, J., MANCINI, G.J. & LEAR, S.A., 2012. The association between triglyceride to high-density-lipoprotein cholesterol ratio and insulin resistance in a multiethnic primary prevention cohort. *Metabolism*, *61*(4), .583-589.

GETACHEW, Y., HASSAN, L., ZAKARIA, Z. & ABDUL AZIZ, S. 2013. Genetic variability of vancomycin-resistant Enterococcus faecium and Enterococcus faecalis isolates from humans, chickens, and pigs in Malaysia. *Applied and environmental microbiology,* 79**,** 4528-4533.

GHANI, W. M. N., RAZAK, I. A., YANG, Y.-H., TALIB, N. A., IKEDA, N., AXELL, T., GUPTA, P. C., HANDA, Y., ABDULLAH, N. & ZAIN, R. B. 2011. Factors affecting commencement and cessation of betel quid chewing behaviour in Malaysian adults. *BMC Public Health,* 11**,** 82.

GIBBONS, R. J., SOCRANSKY, S. S. & KAPSIMALIS, B. 1964. ESTABLISHMENT OF HUMAN INDIGENOUS BACTERIA IN GERM-FREE MICE. *J Bacteriol,* 88**,** 1316-23.

GLENDINNING, L., STEWART, R. D., PALLEN, M. J., WATSON, K. A. & WATSON, M. 2020. Assembly of hundreds of novel bacterial genomes from the chicken caecum. *Genome Biology,* 21**,** 34.

GLOOR, G. B., WU, J. R., PAWLOWSKY-GLAHN, V. & EGOZCUE, J. J. 2016. It's all relative: analyzing microbiome data as compositions. *Annals of Epidemiology,* 26**,** 322-329.

GOOGLE MAPS, INEGI. 2021. Gut microbiome of traditional populations. Available through https://www.google.com/maps/d/u/0/edit?mid=1qysP4YawBVjClUvJJLz20IbrbBgOzs2d&ll=17.501150046136143%2C5.712890625&z=3. Assessed Date: 13th June 2021

GOMEZ, A., PETRZELKOVA, K. J., BURNS, M. B., YEOMAN, C. J., AMATO, K. R., VLCKOVA, K., MODRY, D., TODD, A., ROBINSON, C. A. J., 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**,** 2142-2153.

GORVITOVSKAIA, A., HOLMES, S. P. & HUSE, S. M. 2016. Interpreting Prevotella and Bacteroides as biomarkers of diet and lifestyle. *Microbiome,* 4**,** 15.

GOULIOURIS, T., BLANE, B., BRODRICK, H. J., RAVEN, K. E., AMBRIDGE, K. E., KIDNEY, A. D., HADJIRIN, N. F., TÖRÖK, M. E., LIMMATHUROTSAKUL, D. & PEACOCK, S. J. 2016. Comparison of two chromogenic media for the detection of vancomycin-resistant enterococcal carriage by nursing home residents. *Diagnostic microbiology and infectious disease,* 85**,** 409-412.

GREEN, G. E., DODD, M. C. & INVERSO, H. S. 1957. Comparative microflora of developing dental plaques in caries-immune and susceptible individuals. *Journal of dental research,* 36**,** 331-337.

GRUNDY, S. M., CLEEMAN, J. I., DANIELS, S. R., DONATO, K. A., ECKEL, R. H., FRANKLIN, B. A., GORDON, D. J., KRAUSS, R. M., SAVAGE, P. J., SMITH, S. C., SPERTUS, J. A. & COSTA, F. 2005. Diagnosis and Management of the Metabolic Syndrome. *Circulation,* 112**,** e285-e290.

GUNALETCHUMY, S. P., SEEVASANT, I., TAN, M. H., CROFT, L. J., MITCHELL, H. M., GOH, K. L., LOKE, M. F. & VADIVELU, J. 2014. Helicobacter pylori genetic diversity and gastro-duodenal diseases in Malaysia. *Scientific reports,* 4**,** 7431.

HAMMERUM, A. M. 2012. Enterococci of animal origin and their significance for public health. *Clinical Microbiology and Infection,* 18**,** 619-625.

HAN, Y. W., REDLINE, R. W., LI, M., YIN, L., HILL, G. B. & MCCORMICK, T. S. 2004. Fusobacterium nucleatum induces premature and term stillbirths in pregnant mice: implication of oral bacteria in preterm birth. *Infection and immunity,* 72**,** 2272-2279.

HANAFIAH, A., BINMAEIL, H., ALI, R. A. R., ROSE, I. M. & LOPES, B. S. 2019. Molecular characterization and prevalence of antibiotic resistance in Helicobacter pylori isolates in Kuala Lumpur, Malaysia. *Infection and drug resistance,* 12**,** 3051.

HANDSLEY-DAVIS, M., KOWAL, E., RUSSELL, L. & WEYRICH, L. S. 2021. Researchers using environmental DNA must engage ethically with Indigenous communities. *Nature ecology & evolution,* 5**,** 146-148.

HANSEN, M. E. B., RUBEL, M. A., BAILEY, A. G., RANCIARO, A., THOMPSON, S. R., CAMPBELL, M. C., BEGGS, W., DAVE, J. R., MOKONE, G. G., MPOLOKA, S. W., NYAMBO, T., ABNET, C., CHANOCK, S. J., BUSHMAN, F. D. & TISHKOFF, S. A. 2019. Population structure of human gut bacteria in a diverse cohort from rural Tanzania and Botswana. *Genome Biology,* 20**,** 16.

HARAN, J. P., BHATTARAI, S. K., FOLEY, S. E., DUTTA, P., WARD, D. V., BUCCI, V. & MCCORMICK, B. A. 2019. Alzheimer's Disease Microbiome Is Associated with Dysregulation of the Anti-Inflammatory P-Glycoprotein Pathway. *mBio,* 10**,** e00632-19.

HARDHAM, J. M., KING, K. W., DREIER, K., WONG, J., STRIETZEL, C., EVERSOLE, R. R., SFINTESCU, C. & EVANS, R. T. 2008. Transfer of Bacteroides splanchnicus to Odoribacter gen. nov. as

Odoribacter splanchnicus comb. nov., and description of Odoribacter denticanis sp. nov., isolated from the crevicular spaces of canine periodontitis patients. *International Journal of Systematic and Evolutionary Microbiology,* 58**,** 103-109.

HE, C., YANG, Z. & LU, N. 2016. Imbalance of Gastrointestinal Microbiota in the Pathogenesis of helicobacter pylori‐associated Diseases. *Helicobacter,* 21**,** 337-348.

HERNANDEZ, B. Y., ZHU, X., GOODMAN, M. T., GATEWOOD, R., MENDIOLA, P., QUINATA, K. & PAULINO, Y. C. 2017. Betel nut chewing, oral premalignant lesions, and the oral microbiome. *PLoS One,* 12**,** e0172196.

HERRERA, P. M., MENDEZ, M., VELAPATIÕ, B., SANTIVAÃZ, L., BALQUI, J., FINGER, S. A., SHERMAN, J., ZIMIC, M., CABRERA, L., WATANABE, J., RODRÍGUEZ, C., GILMAN, R. H. & BERG, D. E. 2008. DNA-Level Diversity and Relatedness of Helicobacter pylori Strains in Shantytown Families in Peru and Transmission in a Developing-Country Setting. *Journal of Clinical Microbiology,* 46**,** 3912-3918.

HIDRON, A. I., EDWARDS, J. R., PATEL, J., HORAN, T. C., SIEVERT, D. M., POLLOCK, D. A., FRIDKIN, S. K., TEAM, N. H. S. N. & FACILITIES, P. N. H. S. N. 2008. Antimicrobial‐resistant pathogens associated with healthcare‐associated infections: annual summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2006–2007. *Infection control and hospital epidemiology,* 29**,** 996-1011.

HIIPPALA, K., BARRETO, G., BURRELLO, C., DIAZ-BASABE, A., SUUTARINEN, M., KAINULAINEN, V., BOWERS, J. R., LEMMER, D., ENGELTHALER, D. M., EKLUND, K. K., FACCIOTTI, F. & SATOKARI, R. 2020. Novel Odoribacter splanchnicus Strain and Its Outer Membrane Vesicles Exert Immunoregulatory Effects in vitro. *Frontiers in Microbiology,* 11.

HILL, C., SOARES, P., MORMINA, M., MACAULAY, V., MEEHAN, W., BLACKBURN, J., CLARKE, D., RAJA, J. M., ISMAIL, P., BULBECK, D., OPPENHEIMER, S. & RICHARDS, M. 2006. Phylogeography and Ethnogenesis of Aboriginal Southeast Asians. *Molecular Biology and Evolution,* 23**,** 2480-2491.

HOH, B.-P., ABDUL RAHMAN, T. & YUSOFF, K. 2019. Natural selection and local adaptation of blood pressure regulation and their perspectives on precision medicine in hypertension. *Hereditas,* 156**,** 1.

HOLLOWAY, K. A., KOTWANI, A., BATMANABANE, G., PURI, M. & TISOCKI, K. 2017. Antibiotic use in South East Asia and policies to promote appropriate use: reports from country situational analyses. *BMJ,* 358**,** j2291.

HOOI, J. K. Y., LAI, W. Y., NG, W. K., SUEN, M. M. Y., UNDERWOOD, F. E., TANYINGOH, D., MALFERTHEINER, P., GRAHAM, D. Y., WONG, V. W. S., WU, J. C. Y., CHAN, F. K. L., SUNG, J. J. Y., KAPLAN, G. G. & NG, S. C. 2017. Global Prevalence of Helicobacter pylori Infection: Systematic Review and Meta-Analysis. *Gastroenterology,* 153**,** 420-429.

HORNUNG, B. V. H., ZWITTINK, R. D. & KUIJPER, E. J. 2019. Issues and current standards of controls in microbiome research. *FEMS Microbiology Ecology,* 95.

HOSHIGA, M., KATO-KOGOE, N., SAKAGUCHI, S., KAMIYA, K., OMORI, M., NITTA, A., GU, Y., FUKUI, K., ITO, Y., NAKANO, T., TAMAKI, J. & UENO, T. 2019. Abstract 13277: Characterization of Salivary Microbiota in Patients With Atherosclerosis and Its Potential as a Biomarker. *Circulation,* 140**,** A13277-A13277.

HU, Y., YANG, X., QIN, J., LU, N., CHENG, G., WU, N., PAN, Y., LI, J., ZHU, L., WANG, X., MENG, Z., ZHAO, F., LIU, D., MA, J., QIN, N., XIANG, C., XIAO, Y., LI, L., YANG, H., WANG, J., YANG, R., GAO, G. F., WANG, J. & ZHU, B. 2013. Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nature Communications,* 4**,** 2151.

HUA, X., MCDERMOTT, R., LUNG, T., WENITONG, M., TRAN-DUY, A., LI, M. & CLARKE, P. 2017. Validation and recalibration of the Framingham cardiovascular disease risk models in an Australian Indigenous cohort. *European Journal of Preventive Cardiology,* 24**,** 1660-1669.

HUDSON, C. M., BENT, Z. W., MEAGHER, R. J. & WILLIAMS, K. P. 2014. Resistance determinants and mobile genetic elements of an NDM-1-encoding Klebsiella pneumoniae strain. *PloS one,* 9**,** e99209-e99209.

HUGHES, D. A., BACIGALUPE, R., WANG, J., RÜHLEMANN, M. C., TITO, R. Y., FALONY, G., JOOSSENS, M., VIEIRA-SILVA, S., HENCKAERTS, L., RYMENANS, L., VERSPECHT, C., RING, S., FRANKE, A., WADE, K. H., TIMPSON, N. J. & RAES, J. 2020. Genome-wide associations of human gut microbiome variation and implications for causal inference analyses. *Nature Microbiology,* 5**,** 1079-1087.

IARC, I. A. F. R. O. C. 2014. *Helicobacter pylori* Eradication as a Strategy for Preventing Gastric Cancer. Lyon, France: International Agency for Research on Cancer.

IBRAHIM, R., AHMAD, N. & AZIZ, M. N. 2013. P065: Vancomycin resistant enterococci among patient in Kuala Lumpur Hospital, Malaysia: the occurence and its associated risk factors. *Antimicrobial Resistance and Infection Control,* 2**,** P65-P65.

IPH MALAYSIA, I. F. P. H., NIH, NATIONAL INSTITUES OF HEALTH, MOH, MINISTRY OF HEALTH 2020. National Health and Morbidity Survery (NHMS) Vol 1: NCDs – Non-Communicable Diseases: Risk Factors and other Health Problems.

ISRAEL, D. A., SALAMA, N., ARNOLD, C. N., MOSS, S. F., ANDO, T., WIRTH, H.-P., THAM, K. T., CAMORLINGA, M., BLASER, M. J. & FALKOW, S. 2001. Helicobacter pylori strain-specific differences in genetic content, identified by microarray, influence host inflammatory responses. *The Journal of clinical investigation,* 107**,** 611-620.

IWAUCHI, M., HORIGOME, A., ISHIKAWA, K., MIKUNI, A., NAKANO, M., XIAO, J. Z., ODAMAKI, T. & HIRONAKA, S. 2019. Relationship between oral and gut microbiota in elderly people. *Immunity, inflammation and disease,* 7**,** 229-236.

IWGIA 2020. The Indigenous World 2021. *In:* MAMO, D. (ed.) 35 ed. Copenhagen, Denmark.

JAKOBSSON, H. E., JERNBERG, C., ANDERSSON, A. F., SJÖLUND-KARLSSON, M., JANSSON, J. K. & ENGSTRAND, L. 2010. Short-Term Antibiotic Treatment Has Differing Long-Term Impacts on the Human Throat and Gut Microbiome. *PLOS ONE,* 5**,** e9836.

JANSSEN, S., MCDONALD, D., GONZALEZ, A., NAVAS-MOLINA, J. A., JIANG, L., XU, Z. Z., WINKER, K., KADO, D. M., ORWOLL, E., MANARY, M., MIRARAB, S. & KNIGHT, R. 2018. Phylogenetic Placement of Exact Amplicon Sequences Improves Associations with Clinical Information. *mSystems,* 3**,** e00021-18.

JAVDAN, B., LOPEZ, J. G., CHANKHAMJON, P., LEE, Y.-C. J., HULL, R., WU, Q., WANG, X., CHATTERJEE, S. & DONIA, M. S. 2020. Personalized mapping of drug metabolism by the human gut microbiome. *Cell,* 181**,** 1661-1679. e22.

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**,** e2005396.

JIA, L., LI, D., FENG, N., SHAMOON, M., SUN, Z., DING, L., ZHANG, H., CHEN, W., SUN, J. & CHEN, Y. Q. 2017. Anti-diabetic Effects of Clostridium butyricum CGMCC0313.1 through Promoting the Growth of Gut Butyrate-producing Bacteria in Type 2 Diabetic Mice. *Scientific Reports,* 7**,** 7046.

KAAKOUSH, N. O. 2020. Sutterella Species, IgA-degrading Bacteria in Ulcerative Colitis. *Trends in Microbiology,* 28**,** 519-522.

KATAKAMI, N., KANETO, H., OSONOI, T., SAITOU, M., TAKAHARA, M., SAKAMOTO, F., YAMAMOTO, K., YASUDA, T., MATSUOKA, T.A., MATSUHISA, M. AND YAMASAKI, Y. 2011. Usefulness of lipoprotein ratios in assessing carotid atherosclerosis in Japanese type 2 diabetic patients. *Atherosclerosis*, *214*(2), pp.442-447.

KATO, M., OTA, H., OKUDA, M., KIKUCHI, S., SATOH, K., SHIMOYAMA, T., SUZUKI, H., HANDA, O., FURUTA, T., MABE, K., MURAKAMI, K., SUGIYAMA, T., UEMURA, N. & TAKAHASHI, S. I. 2019. Guidelines for the management of Helicobacter pylori infection in Japan: 2016 Revised Edition. *Helicobacter,* 24**,** e12597.

KATSIDZIRA, L., OCVIRK, S., WILSON, A., LI, J., MAHACHI, C. B., SONI, D., DELANY, J., NICHOLSON, J. K., ZOETENDAL, E. G. & O'KEEFE, S. J. D. 2019. Differences in Fecal Gut Microbiota, Short-Chain

Fatty Acids and Bile Acids Link Colorectal Cancer Risk to Dietary Changes Associated with Urbanization Among Zimbabweans. *Nutrition and Cancer,* 71**,** 1313-1324.

KHOR, C. S., HASSAN, H., MOHD-RAHIM, N. F., CHANDREN, J. R., NORE, S. S., JOHARI, J., LOONG, S. K., ABD-JAMIL, J., KHOO, J. J., LEE, H. Y., PIKE, B. L., LI-PING, W., LIM, Y. A. & ABUBAKAR, S. 2019. Seroprevalence of Borrelia burgdorferi among the indigenous people (Orang Asli) of Peninsular Malaysia. *J Infect Dev Ctries,* 13**,** 449-454.

KHOR, G. L. & ZALILAH, M. 2008. The ecology of health and nutrition of Orang Asli (Indigenous people) women and children in Peninsular Malaysia. *Tribes and Tribals,* 2**,** 66-77.

KIVI, M., TINDBERG, Y., SÖRBERG, M., CASSWALL, T. H., BEFRITS, R., HELLSTRÖM, P. M., BENGTSSON, C., ENGSTRAND, L. & GRANSTRÖM, M. 2003. Concordance of Helicobacter pylori Strains within Families. *Journal of Clinical Microbiology,* 41**,** 5604-5608.

KLARE, I., HEIER, H., CLAUS, H., REISSBRODT, R. & WITTE, W. V. 1995. vanA-mediated high-level glycopeptide resistance in Enterococcus faecium from animal husbandry. *FEMS Microbiology Letters,* 125**,** 165-171.

KLIEBE, C., NIES, B. A., MEYER, J. F., TOLXDORFF-NEUTZLING, R. M. & WIEDEMANN, B. 1985. Evolution of plasmid-coded resistance to broad-spectrum cephalosporins. *Antimicrobial Agents and Chemotherapy,* 28**,** 302-307.

KLINDWORTH, A., PRUESSE, E., SCHWEER, T., PEPLIES, J., QUAST, C., HORN, M. & GLÖCKNER, F. O. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic acids research,* 41**,** e1-e1.

KOVATCHEVA-DATCHARY, P., NILSSON, A., AKRAMI, R., LEE, Y. S., DE VADDER, F., ARORA, T., HALLEN, A., MARTENS, E., BJÖRCK, I. & BÄCKHED, F. 2015. Dietary fiber-induced improvement in glucose metabolism is associated with increased abundance of Prevotella. *Cell metabolism,* 22**,** 971-982.

KRUEGER, F. *Trim_galore* [Online]. [Accessed].

KUCHIKURA, Y. 1988. Food Use and Nutrition in a Hunting and Gethering Community in Transition, Peninuslar Malaysia. *Man and Culture in Oceania,* 4**,** 1-30.

LAGIER, J.-C., DUBOURG, G., MILLION, M., CADORET, F., BILEN, M., FENOLLAR, F., LEVASSEUR, A., ROLAIN, J.-M., FOURNIER, P.-E. & RAOULT, D. 2018. Culturing the human microbiota and culturomics. *Nature Reviews Microbiology,* 16**,** 540-550.

LAMARCHE, B., MOORJANI, S., LUPIEN, P.J., CANTIN, B., BERNARD, P.M., DAGENAIS, G.R. AND DESPRE´ S, J.P. 1996. Apolipoprotein AI and B Levels and the Risk of Ischemic Heart Disease During a Five-Year Follow-up of Men in the Que bec Cardiovascular Study. *Circulation*, *94*(3), pp.273-278.

LANGDON, A., CROOK, N. & DANTAS, G. 2016. The effects of antibiotics on the microbiome throughout development and alternative approaches for therapeutic modulation. *Genome medicine,* 8**,** 1-16.

LASSALLE, F., SPAGNOLETTI, M., FUMAGALLI, M., SHAW, L., DYBLE, M., WALKER, C., THOMAS, M. G., BAMBERG MIGLIANO, A. & BALLOUX, F. 2018. Oral microbiomes from hunter-gatherers and traditional farmers reveal shifts in commensal balance and pathogen load linked to diet. *Mol Ecol,* 27**,** 182-195.

LAW, L. S., NORHASMAH, S., GAN, W. Y., SITI NUR'ASYURA, A. & MOHD NASIR, M. T. 2018. The identification of the factors related to household food insecurity among indigenous people (Orang Asli) in peninsular Malaysia under traditional food systems. *Nutrients,* 10**,** 1455.

LECLERCQ, R., DERLOT, E., DUVAL, J. & COURVALIN, P. 1988. Plasmid-Mediated Resistance to Vancomycin and Teicoplanin in Enterococcus Faecium. *New England Journal of Medicine,* 319**,** 157-161.

LEE, K., KIM, D.-W., LEE, D.-H., KIM, Y.-S., BU, J.-H., CHA, J.-H., THAWNG, C. N., HWANG, E.-M., SEONG, H. J., SUL, W. J., WELLINGTON, E. M. H., QUINCE, C. & CHA, C.-J. 2020. Mobile resistome of human gut and pathogen drives anthropogenic bloom of antibiotic resistance. *Microbiome,* 8**,** 2.

LEE, S. C., TANG, M. S., EASTON, A. V., DEVLIN, J. C., CHUA, L. L., CHO, I., MOY, F. M., KHANG, T. F., LIM, Y. A. L. & LOKE, P. N. 2019. Linking the effects of helminth infection, diet and the gut microbiota with human whole-blood signatures. *PLOS Pathogens,* 15**,** e1008066.

LEE, S. C., TANG, M. S., LIM, Y. A. L., CHOY, S. H., KURTZ, Z. D., COX, L. M., GUNDRA, U. M., CHO, I., BONNEAU, R., BLASER, M. J., CHUA, K. H. & LOKE, P. N. 2014. Helminth Colonization Is Associated with Increased Diversity of the Gut Microbiota. *PLOS Neglected Tropical Diseases,* 8**,** e2880.

LEMIEUX, I., LAMARCHE, B., COUILLARD, C., PASCOT, A., CANTIN, B., BERGERON, J., DAGENAIS, G.R. AND DESPRÉS, J.P. 2001. Total cholesterol/HDL cholesterol ratio vs LDL cholesterol/HDL cholesterol ratio as indices of ischemic heart disease risk in men: the Quebec Cardiovascular Study. *Archives of internal medicine*, *161*(22), pp.2685-2692.

LETCHUMANAN, V., PUSPARAJAH, P., TAN, L. T.-H., YIN, W.-F., LEE, L.-H. & CHAN, K.-G. 2015. Occurrence and antibiotic resistance of Vibrio parahaemolyticus from shellfish in Selangor, Malaysia. *Frontiers in Microbiology,* 6**,** 1417.

LEWIS JR, C. M., OBREGÓN-TITO, A., TITO, R. Y., FOSTER, M. W. & SPICER, P. G. 2012. The Human Microbiome Project: lessons from human genomics. *Trends in microbiology,* 20**,** 1-4.

LI, H. & DURBIN, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics,* 25**,** 1754-60.

LI, J., QUINQUE, D., HORZ, H.-P., LI, M., RZHETSKAYA, M., RAFF, J. A., HAYES, M. G. & STONEKING, M. 2014. Comparative analysis of the human saliva microbiome from different climate zones: Alaska, Germany, and Africa. *BMC Microbiology,* 14**,** 316.

LI, J., RETTEDAL, E. A., VAN DER HELM, E., ELLABAAN, M., PANAGIOTOU, G. & SOMMER, M. O. A. 2019. Antibiotic Treatment Drives the Diversification of the Human Gut Resistome. *Genomics, Proteomics & Bioinformatics,* 17**,** 39-51.

LIM, H. M. & CHEE, H. L. 1998. Nutritional status and reproductive health of Orang Asli women in two villages in Kuantan, Pahang. *Malaysian journal of nutrition,* 4**,** 31-54.

LIM, V., CHEONG, Y. & SULEIMAN, A. B. 1994. Pattern of antibiotic usage in hospitals in Malaysia. *Singapore medical journal,* 34**,** 525-8.

LIM, Y., TOTSIKA, M., MORRISON, M. & PUNYADEERA, C. 2017. The saliva microbiome profiles are minimally affected by collection method or DNA extraction protocols. *Scientific Reports,* 7**,** 8523.

LIM, Y. A., ROMANO, N., COLIN, N., CHOW, S. C. & SMITH, H. V. 2009. Intestinal parasitic infections amongst Orang Asli (indigenous) in Malaysia: has socioeconomic development alleviated the problem? *Trop Biomed,* 26**,** 110-22.

LIU, F., LIANG, T., ZHANG, Z., LIU, L., LI, J., DONG, W., ZHANG, H., BAI, S., MA, L. & KANG, L. 2021a. Effects of altitude on human oral microbes. *AMB Express,* 11**,** 41.

LIU, J., CUI, L., YAN, X., ZHAO, X., CHENG, J., ZHOU, L., GAO, J., CAO, Z., YE, X. & HU, S. 2018. Analysis of Oral Microbiota Revealed High Abundance of Prevotella Intermedia in Gout Patients. *Cellular Physiology and Biochemistry,* 49**,** 1804-1812.

LIU, P.-Y., WU, W.-K., CHEN, C.-C., PANYOD, S., SHEEN, L.-Y. & WU, M.-S. 2020. Evaluation of Compatibility of 16S rRNA V3V4 and V4 Amplicon Libraries for Clinical Microbiome Profiling. *bioRxiv*.

LIU, X., MAO, B., GU, J., WU, J., CUI, S., WANG, G., ZHAO, J., ZHANG, H. & CHEN, W. 2021b. Blautia-a new functional genus with potential probiotic properties? *Gut Microbes,* 13**,** 1-21.

LIU, X., YUNUS, Y., LU, D., AGHAKHANIAN, F., SAW, W. Y., DENG, L., ALI, M., WANG, X., NOR, F. M., RAHMAN, T. A., SHAARI, S. A., SALLEH, M. Z., PHIPPS, M. E., ONG, R. T., XU, S., TEO, Y. Y. & HOH, B. P. 2015. Differential positive selection of malaria resistance genes in three indigenous populations of Peninsular Malaysia. *Hum Genet,* 134**,** 375-92.

LOKE, V. P., LIM, T. & CAMPOS-ARCEIZ, A. 2020. Hunting practices of the Jahai indigenous community in northern peninsular Malaysia. *Global Ecology and Conservation,* 21**,** e00815.

LOKMER, A., AFLALO, S., AMOUGOU, N., LAFOSSE, S., FROMENT, A., TABE, F. E., POYET, M., GROUSSIN, M., SAID-MOHAMED, R. & SÉGUREL, L. 2020. Response of the human gut and saliva microbiome to urbanization in Cameroon. *Scientific Reports,* 10**,** 2856.

LOVELL, D. P. 2013. Biological Importance and Statistical Significance. *Journal of Agricultural and Food Chemistry,* 61**,** 8340-8348.

LOZUPONE, C. A., STOMBAUGH, J. I., GORDON, J. I., JANSSON, J. K. & KNIGHT, R. 2012. Diversity, stability and resilience of the human gut microbiota. *Nature,* 489**,** 220-230.

LV, B.-M., QUAN, Y. & ZHANG, H.-Y. 2021. Causal Inference in Microbiome Medicine: Principles and Applications. *Trends in Microbiology*.

MAHENDRA RAJ, S., LEE, Y. Y., CHOO, K. E., NOORIZAN, A. M., ZULKIFLI, A., RADZI, M. & ANG, S. C. 2008. Further observations in an area with an exceptionally low prevalence of Helicobacter pylori infection. *Transactions of The Royal Society of Tropical Medicine and Hygiene,* 102**,** 1163-1164.

MAIXNER, F., KRAUSE-KYORA, B., TURAEV, D., HERBIG, A., HOOPMANN, M. R., HALLOWS, J. L., KUSEBAUCH, U., VIGL, E. E., MALFERTHEINER, P., MEGRAUD, F., O'SULLIVAN, N., CIPOLLINI, G., COIA, V., SAMADELLI, M., ENGSTRAND, L., LINZ, B., MORITZ, R. L., GRIMM, R., KRAUSE, J., NEBEL, A., MOODLEY, Y., RATTEI, T. & ZINK, A. 2016. The 5300-year-old Helicobacter pylori genome of the Iceman. *Science,* 351**,** 162-165.

MAJID, A. & BURENHULT, N. 2014. Odors are expressible in language, as long as you speak the right language. *Cognition,* 130**,** 266-270.

MALHI, R. S. & BADER, A. 2015. ENGAGING NATIVE AMERICANS IN GENOMICS RESEARCH. *American anthropologist,* 117**,** 743-744.

MARAN, S., LEE, Y. Y., XU, S. H., RAJ, M. S., ABDUL MAJID, N., CHOO, K. E., ZILFALIL, B. A. & GRAHAM, D. Y. 2013. Towards understanding the low prevalence of Helicobacter pylori in Malays: Genetic variants among Helicobacter pylori-negative ethnic Malays in the north-eastern region of Peninsular Malaysia and Han Chinese and South Indians. *Journal of Digestive Diseases,* 14**,** 196-202.

MARQUARDT, R. R. & LI, S. 2018. Antimicrobial resistance in livestock: advances and alternatives to antibiotics. *Animal Frontiers,* 8**,** 30-37.

MARTIN, E. H. R. (ed.) 2015. *A Dictionary of Biology (7 ed.)*: Oxford University Press.

MARTIN, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal,* 17**,** 10-12.

MARTINA, E., CAMPANATI, A., DIOTALLEVI, F. & OFFIDANI, A. 2020. Saliva and Oral Diseases. *Journal of Clinical Medicine,* 9**,** 466.

MARTÍNEZ, I., MULLER, C. E. & WALTER, J. 2013. Long-Term Temporal Analysis of the Human Fecal Microbiota Revealed a Stable Core of Dominant Bacterial Species. *PLOS ONE,* 8**,** e69621.

MATA, L. J., CARRILLO, C. & VILLATORO, E. 1969. Fecal microflora in health persons in a preindustrial region. *Appl Microbiol,* 17**,** 596-602.

MATHEWS, S.C., MALLIDI, J., KULKARNI, K., TOTH, P.P. & JONES, S.R. 2012. Achieving secondary prevention low-density lipoprotein particle concentration goals using lipoprotein cholesterol-based data. *PLoS One*, *7*(3), p.e33692.

MCKEOWN R. E. 2009. The Epidemiologic Transition: Changing Patterns of Mortality and Population Dynamics. *American journal of lifestyle medicine*, *3*(1 Suppl), 19S–26S. https://doi.org/10.1177/1559827609335350

MCMURDIE, P. J. & HOLMES, S. 2013. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE,* 8**,** e61217.

MELZER, M. & PETERSEN, I. 2007. Mortality following bacteraemic infection caused by extended spectrum beta-lactamase (ESBL) producing E. coli compared to non-ESBL producing E. coli. *Journal of Infection,* 55**,** 254-259.

METWALY, A. & HALLER, D. 2019. Strain-Level Diversity in the Gut: The P. copri Case. *Cell Host & Microbe,* 25**,** 349-350.

MILLÁN, J., PINTÓ, X., MUÑOZ, A., ZÚÑIGA, M., RUBIÉS-PRAT, J., PALLARDO, L. F., MASANA, L., MANGAS, A., HERNÁNDEZ-MIJARES, A., GONZÁLEZ-SANTOS, P., ASCASO, J. F. & PEDRO-BOTET, J. 2009. Lipoprotein ratios: Physiological significance and clinical usefulness in cardiovascular prevention. *Vascular health and risk management,* 5**,** 757-765.

MINISTRY OF HEALTH MALAYSIA, A. O. M. O. M., MALAYSIAN SOCIETY OF INFECTIOUS DISEASES AND CHEMOTHERAPY 2001. Consensus Guidelines For The Management of Infections by ESBL-producing Bacteria. *In:* N PARASAKHTI, H. A. (ed.).

MOHAMED, M., ABU, J., ABDUL AZIZ, S., ZAKARIA, Z., KHAN, A. & MOHAMED, E. 2019. Occurrence of Antibiotic Resistant C. jejuni and E. coli in Wild Birds, Chickens, Environment and Humans from Orang Asli Villages in Sungai Siput, Perak, Malaysia. *American Journal of Animal and Veterinary Sciences,* 14.

MOHAMED, N., HUSSIN, H., CHANG, K. M., HASHIM, R. & AHMAD, N. 2015. Vancomycin resistant Enterococcus (VRE): Prevalence and characteristics in a tertiary hospital in Malaysia. 11**,** 241-246.

MOHAMED, N. I., ABDULLAH, N. H., DAWAM, N. N. M. & SUI, C. F. 2021. Drug Utilization Evaluation of Intravenous (IV) Cephalosporins in Hospital Tapah: A 6-Month Retrospective Case Review Study. *Perak Pharmacy Journal,* 3.

MOHAMUD, W. N. & SURAIAMI, M. 2010. Prevalence of diabetes, impaired fasting glucose and metabolic syndrome among female Orang Asli community in Peninsular Malaysia. *International Journal of Diabetes in Developing Countries,* 30**,** 118.

MOKHSIN, A., MOKHTAR, S. S., MOHD ISMAIL, A., M NOR, F., SHAARI, S. A., NAWAWI, H., YUSOFF, K., ABDUL RAHMAN, T. & HOH, B. P. 2018. Observational study of the status of coronary risk biomarkers among Negritos with metabolic syndrome in the east coast of Malaysia. *BMJ open,* 8**,** e021580-e021580.

MOKHTAR, S., MOKHSIN, N. M., NOR, F. M., SHAARI, S., NAWAWI, H., HOH, B. & RAHMAN, T. 2016. Status of inflammation, endothelial activation and prothrombogenesis among Negritos and Malays with metabolic syndrome. *Atherosclerosis,* 252**,** e127.

MOORE, A. M., PATEL, S., FORSBERG, K. J., WANG, B., BENTLEY, G., RAZIA, Y., QIN, X., TARR, P. I. & DANTAS, G. 2013. Pediatric Fecal Microbiota Harbor Diverse and Novel Antibiotic Resistance Genes. *PLOS ONE,* 8**,** e78822.

MORGAN, X. C., TICKLE, T. L., SOKOL, H., GEVERS, D., DEVANEY, K. L., WARD, D. V., REYES, J. A., SHAH, S. A., LELEIKO, N., SNAPPER, S. B., BOUSVAROS, A., KORZENIK, J., SANDS, B. E., XAVIER, R. J. & HUTTENHOWER, C. 2012. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biology,* 13**,** R79.

MOSTAFA, S. A., DAVIES, M. J., MORRIS, D. H., YATES, T., SRINIVASAN, B. T., WEBB, D., BRADY, E. & KHUNTI, K. 2012. The association of the triglyceride-to-HDL cholesterol ratio with insulin resistance in White European and South Asian men and women. *PloS one,* 7**,** e50931-e50931.

MUELLER, N. T., BAKACS, E., COMBELLICK, J., GRIGORYAN, Z. & DOMINGUEZ-BELLO, M. G. 2015. The infant microbiome development: mom matters. *Trends in molecular medicine,* 21**,** 109-117.

MUGHINI-GRAS, L., DORADO-GARCÍA, A., VAN DUIJKEREN, E., VAN DEN BUNT, G., DIERIKX, C. M., BONTEN, M. J. M., BOOTSMA, M. C. J., SCHMITT, H., HALD, T., EVERS, E. G., DE KOEIJER, A., VAN PELT, W., FRANZ, E., MEVIUS, D. J. & HEEDERIK, D. J. J. 2019. Attributable sources of community-acquired carriage of Escherichia coli containing β-lactam antibiotic resistance genes: a population-based modelling study. *The Lancet Planetary Health,* 3**,** e357-e369.

MURRAY, T. S. & CASSESE, T. 2016. 2 - Bacteriology of the Head and Neck Regions. *In:* HUPP, J. R. & FERNEINI, E. M. (eds.) *Head, Neck, and Orofacial Infections.* St. Louis: Elsevier.

MUSA, A. F., YUNOS, M. N. M., RAHMAN, S. A. & NORDIN, R. B. 2014. The seroprevalence and eradication success of Helicobacter pylori in indigenous people of Seletar in Southern Malaysia. *Journal of Advances in Medicine and Medical Research***,** 1854-1863.

NATH, S., HANDSLEY-DAVIS, M., WEYRICH, L. S. & JAMIESON, L. M. 2021. Diversity and bias in oral microbiome research: A commentary. *EClinicalMedicine,* 36.

NAYFACH, S., SHI, Z. J., SESHADRI, R., POLLARD, K. S. & KYRPIDES, N. C. 2019. New insights from uncultivated genomes of the global human gut microbiome. *Nature,* 568**,** 505-510.

NEPAL, G. & BHATTA, S. 2018. Self-medication with antibiotics in WHO Southeast Asian Region: a systematic review. *Cureus,* 10.

NEVES, P. R., MCCULLOCH, J. A., MAMIZUKA, E. M. & LINCOPAN, N. 2014. PSEUDOMONAS | Pseudomonas aeruginosa. *In:* BATT, C. A. & TORTORELLO, M. L. (eds.) *Encyclopedia of Food Microbiology (Second Edition).* Oxford: Academic Press.

NGOI, S. T., CHONG, C., PONNAMPALAVANAR, S., TANG, S., IDRIS, N., ABDUL JABAR, K., GREGORY, M., HUSAIN, T. & TEH, C. 2021. Genetic mechanisms and correlated risk factors of antimicrobial-resistant ESKAPEE pathogens isolated in a tertiary hospital in Malaysia. *Antimicrobial Resistance & Infection Control,* 10.

NICHOLAS, C. 2000. The Orang Asli and the contest for resources: Indigenous politics, development and identity in Peninsular Malaysia. *IWGIA document*.

NICHOLAS, C. & BAER, A. 2007. Health care for the Orang Asli: Consequences of Paternalism and Colonialism. *Health care in Malaysia***,** 119-136.

NILSON, C. 2017. A Journey Toward Cultural Competence: The Role of Researcher Reflexivity in Indigenous Research. *J Transcult Nurs,* 28**,** 119-127.

NURFAHILIN, T. & NORHASMAH, S. 2015. Factors and coping strategies related to food insecurity and nutritional status among Orang Asli women in Malaysia. *International Journal of Public Health and Clinical Sciences,* 2**,** 55-66.

NURK, S., MELESHKO, D., KOROBEYNIKOV, A. & PEVZNER, P. A. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res,* 27**,** 824-834.

O'DRISCOLL, T. & CRANK, C. W. 2015. Vancomycin-resistant enterococcal infections: epidemiology, clinical manifestations, and optimal management. *Infection and drug resistance,* 8**,** 217-230.

OBREGON-TITO, A. J., TITO, R. Y., METCALF, J., SANKARANARAYANAN, K., CLEMENTE, J. C., URSELL, L. K., ZECH XU, Z., VAN TREUREN, W., KNIGHT, R., GAFFNEY, P. M., SPICER, P., LAWSON, P., MARIN-REYES, L., TRUJILLO-VILLARROEL, O., FOSTER, M., GUIJA-POMA, E., TRONCOSO-CORZO, L., WARINNER, C., OZGA, A. T. & LEWIS, C. M. 2015. Subsistence strategies in traditional societies distinguish gut microbiomes. *Nature Communications,* 6**,** 6505.

ODUARAN, O. H., TAMBURINI, F. B., SAHIBDEEN, V., BREWSTER, R., GÓMEZ-OLIVÉ, F. X., KAHN, K., NORRIS, S. A., TOLLMAN, S. M., TWINE, R., WADE, A. N., WAGNER, R. G., LOMBARD, Z., BHATT, A. S. & HAZELHURST, S. 2020. Gut microbiome profiling of a rural and urban South African cohort reveals biomarkers of a population in lifestyle transition. *BMC Microbiology,* 20**,** 330.

OKUDA, M., OSAKI, T., LIN, Y., YONEZAWA, H., MAEKAWA, K., KAMIYA, S., FUKUDA, Y. & KIKUCHI, S. 2015. Low Prevalence and Incidence of Helicobacter pylori Infection in Children: A Population-Based Study in Japan. *Helicobacter,* 20**,** 133-138.

OLIVEIRA, A., OLIVEIRA, L. C., ABURJAILE, F., BENEVIDES, L., TIWARI, S., JAMAL, S. B., SILVA, A., FIGUEIREDO, H. C. P., GHOSH, P., PORTELA, R. W., DE CARVALHO AZEVEDO, V. A. & WATTAM, A. R. 2017. Insight of Genus Corynebacterium: Ascertaining the Role of Pathogenic and Non-pathogenic Species. *Frontiers in Microbiology,* 8.

OLM, M. R., BROWN, C. T., BROOKS, B. & BANFIELD, J. F. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *Isme j,* 11**,** 2864-2868.

ORGANIZATION, W. H. 2011. Waist circumference and waist-hip ratio: report of a WHO expert consultation, Geneva, 8-11 December 2008.

OSAKI, T., OKUDA, M., UEDA, J., KONNO, M., YONEZAWA, H., HOJO, F., YAGYU, K., LIN, Y., FUKUDA, Y. & KIKUCHI, S. 2013. Multilocus sequence typing of DNA from faecal specimens for the analysis of intra-familial transmission of Helicobacter pylori. *Journal of medical microbiology,* 62**,** 761-765.

OZATO, N., SAITO, S., YAMAGUCHI, T., KATASHIMA, M., TOKUDA, I., SAWADA, K., KATSURAGI, Y., KAKUTA, M., IMOTO, S. & IHARA, K. 2019. Blautia genus associated with visceral fat accumulation in adults 20–76 years of age. *NPJ biofilms and microbiomes,* 5**,** 1-9.

PANA, Z. D. & ZAOUTIS, T. 2018. Treatment of extended-spectrum β-lactamase-producing Enterobacteriaceae (ESBLs) infections: what have we learned until now? *F1000Research,* 7.

PARK, S. H., YOON, J. S., WON, K. C. & LEE, H. W. 2012. Usefulness of glycated hemoglobin as diagnostic criteria for metabolic syndrome. *Journal of Korean medical science,* 27**,** 1057-1061.

PARKS, D. H., IMELFORT, M., SKENNERTON, C. T., HUGENHOLTZ, P. & TYSON, G. W. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome research,* 25**,** 1043-1055.

PARKS, D. H., RINKE, C., CHUVOCHINA, M., CHAUMEIL, P.-A., WOODCROFT, B. J., EVANS, P. N., HUGENHOLTZ, P. & TYSON, G. W. 2017. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nature Microbiology,* 2**,** 1533-1542.

PÄRNÄNEN, K., KARKMAN, A., HULTMAN, J., LYRA, C., BENGTSSON-PALME, J., LARSSON, D. G. J., RAUTAVA, S., ISOLAURI, E., SALMINEN, S., KUMAR, H., SATOKARI, R. & VIRTA, M. 2018. Maternal gut and breast milk microbiota affect infant gut antibiotic resistome and mobile genetic elements. *Nature Communications,* 9**,** 3891.

PASOLLI, E., ASNICAR, F., MANARA, S., ZOLFO, M., KARCHER, N., ARMANINI, F., BEGHINI, F., MANGHI, P., TETT, A., GHENSI, P., COLLADO, M. C., RICE, B. L., DULONG, C., MORGAN, X. C., GOLDEN, C. D., QUINCE, C., HUTTENHOWER, C. & SEGATA, N. 2019. Extensive Unexplored Human Microbiome Diversity Revealed by Over 150,000 Genomes from Metagenomes Spanning Age, Geography, and Lifestyle. *Cell,* 176**,** 649-662.e20.

PATRICK, S. 2015. Chapter 51 - Bacteroides. *In:* TANG, Y.-W., SUSSMAN, M., LIU, D., POXTON, I. & SCHWARTZMAN, J. (eds.) *Molecular Medical Microbiology (Second Edition).* Boston: Academic Press.

PEDRET, A., VALLS, R. M., CALDERÓN-PÉREZ, L., LLAURADÓ, E., COMPANYS, J., PLA-PAGÀ, L., MORAGAS, A., MARTÍN-LUJÁN, F., ORTEGA, Y. & GIRALT, M. 2019. Effects of daily consumption of the probiotic Bifidobacterium animalis subsp. lactis CECT 8145 on anthropometric adiposity biomarkers in abdominally obese subjects: a randomized controlled trial. *International Journal of Obesity,* 43**,** 1863-1868.

PEI, C. S., APPANNAH, G. & SULAIMAN, N. 2018. Household food insecurity, diet quality, and weight status among indigenous women (Mah Meri) in Peninsular Malaysia. *Nutrition research and practice,* 12**,** 135-142.

PELL, C., ALLOTEY, P., EVANS, N., HARDON, A., IMELDA, J. D., SOYIRI, I., REIDPATH, D. D. & THE, S. T. 2016. Coming of age, becoming obese: a cross-sectional analysis of obesity among adolescents and young adults in Malaysia. *BMC Public Health,* 16**,** 1082.

PENDERS, J., STOBBERINGH, E. E., SAVELKOUL, P. H. & WOLFFS, P. F. 2013. The human microbiome as a reservoir of antimicrobial resistance. *Front Microbiol,* 4**,** 87.

PÉREZ-COBAS, A. E., ARTACHO, A., KNECHT, H., FERRÚS, M. L., FRIEDRICHS, A., OTT, S. J., MOYA, A., LATORRE, A. & GOSALBES, M. J. 2013. Differential effects of antibiotic therapy on the structure and function of human gut microbiota. *PLoS One,* 8**,** e80201.

PHIPPS, M. E., CHAN, K. K., NAIDU, R., MOHAMAD, N. W., HOH, B.-P., QUEK, K.-F., AHMAD, B., HARNIDA, S. M., ZAIN, A. Z. & KADIR, K. A. 2015. Cardio-metabolic health risks in indigenous populations of Southeast Asia and the influence of urbanization. *BMC public health,* 15**,** 47.

PYRON, M. 2010. *Characterizing Communities* [Online]. [Accessed].

QUINQUE, D., KITTLER, R., KAYSER, M., STONEKING, M. & NASIDZE, I. 2006. Evaluation of saliva as a source of human DNA for population and association studies. *Analytical Biochemistry,* 353**,** 272.

QUISPE, R., ELSHAZLY, M.B., ZHAO, D., TOTH, P.P., PURI, R., VIRANI, S.S., BLUMENTHAL, R.S., MARTIN, S.S., JONES, S.R. & MICHOS, E.D. 2020. Total cholesterol/HDL-cholesterol ratio discordance with LDL-cholesterol and non-HDL-cholesterol and incidence of atherosclerotic cardiovascular

disease in primary prevention: The ARIC study. *European journal of preventive cardiology*, *27*(15), pp.1597-1605.
R CORE TEAM 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
RAHIM, A. A., LEE, Y. Y., MAJID, N. A., CHOO, K. E., RAJ, S. M., DERAKHSHAN, M. H. & GRAHAM, D. Y. 2010. Helicobacter pylori infection among Aborigines (the Orang Asli) in the northeastern region of Peninsular Malaysia. *The American journal of tropical medicine and hygiene,* 83**,** 1119-1122.
RAJA, N. S., KARUNAKARAN, R., NGEOW, Y. F. & AWANG, R. 2005. Community-acquired vancomycin-resistant Enterococcus faecium: a case report from Malaysia. *Journal of Medical Microbiology,* 54**,** 901-903.
RASHID, S. S., MHM, N., BAHARUDIN, R. & RAMLI, A. N. M. 2019. A clinical update on Antibiotic Resistance Gram-negative bacteria in Malaysia-a review. *arXiv preprint arXiv:1903.03486*.
RAYMOND, F., BOISSINOT, M., OUAMEUR, A. A., DÉRASPE, M., PLANTE, P.-L., KPANOU, S. R., BÉRUBÉ, È., HULETSKY, A., ROY, P. H., OUELLETTE, M., BERGERON, M. G. & CORBEIL, J. 2019. Culture-enriched human gut microbiomes reveal core and accessory resistance genes. *Microbiome,* 7**,** 56.
RIESENFELD, C. S., GOODMAN, R. M. & HANDELSMAN, J. 2004. Uncultured soil bacteria are a reservoir of new antibiotic resistance genes. *Environmental Microbiology,* 6**,** 981-989.
RILEY, P., PARASAKTHI, N. & TEH, K. H. 1996. *Enterococcus faecium with high level vancomycin resistance isolated from the blood culture of a bone marrow transplant patient in Malaysia*.
RISELY, A. 2020. Applying the core microbiome to understand host–microbe systems. *Journal of Animal Ecology,* 89**,** 1549-1558.
ROGERS, G., WARD, J., BROWN, A. & WESSELINGH, S. 2019. Inclusivity and equity in human microbiome research. *The Lancet,* 393**,** 728-729.
ROSENBAUM, J., USYK, M., CHEN, Z., ZOLNIK, C. P., JONES, H. E., WALDRON, L., DOWD, J. B., THORPE, L. E. & BURK, R. D. 2019. Evaluation of Oral Cavity DNA Extraction Methods on Bacterial and Fungal Microbiota. *Scientific Reports,* 9**,** 1531.
ROTHSCHILD, D., WEISSBROD, O., BARKAN, E., KURILSHIKOV, A., KOREM, T., ZEEVI, D., COSTEA, P. I., GODNEVA, A., KALKA, I. N. & BAR, N. 2018. Environment dominates over host genetics in shaping human gut microbiota. *Nature,* 555**,** 210.
RUGGLES, K. V., WANG, J., VOLKOVA, A., CONTRERAS, M., NOYA-ALARCON, O., LANDER, O., CABALLERO, H. & DOMINGUEZ-BELLO, M. G. 2018. Changes in the Gut Microbiota of Urban Subjects during an Immersion in the Traditional Diet and Lifestyle of a Rainforest Village. *mSphere,* 3.
RUSSELL-MUNDINE, G. 2012. Reflexivity in Indigenous Research: Reframing and Decolonising Research? *Journal of Hospitality and Tourism Management,* 19**,** 85-90.
SAFETY, W. N. A. F. 2011. Waist circumference and waist-hip ratio: report of a WHO expert consultation. *In:* ORGANIZATION, W. H. (ed.).
SAIBUL, N., SHARIFF, Z. M., LIN, K. G., KANDIAH, M., NAWALYAH, A. G. & HEJAR, A. R. 2009. *Food Variety Score Is Associated with Dual Burden of Malnutrition in "Orang Asli" (Malaysian Indigenous Peoples) Households: Implications for Health Promotion*, HEC Press.
SALONEN, A., SALOJÄRVI, J., LAHTI, L. & DE VOS, W. M. 2012. The adult intestinal core microbiota is determined by analysis depth and health status. *Clinical Microbiology and Infection,* 18**,** 16-20.
SAMPAIO-MAIA, B., CALDAS, I. M., PEREIRA, M. L., PÉREZ-MONGIOVI, D. & ARAUJO, R. 2016. Chapter Four - The Oral Microbiome in Health and Its Implication in Oral and Systemic Diseases. *In:* SARIASLANI, S. & MICHAEL GADD, G. (eds.) *Advances in Applied Microbiology.* Academic Press.
SANNA, S., VAN ZUYDAM, N. R., MAHAJAN, A., KURILSHIKOV, A., VICH VILA, A., VÕSA, U., MUJAGIC, Z., MASCLEE, A. A. M., JONKERS, D. M. A. E., 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**,** 600-605.

SANTIAGO-RODRIGUEZ, T. M., FORNACIARI, G., LUCIANI, S., TORANZOS, G. A., MAROTA, I., GIUFFRA, V., SANGWAN, N. & CANO, R. J. 2018. Tetracycline-like resistome of ancient human guts. *Human Microbiome Journal,* 10**,** 21-26.

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.

SEEMANN, T. 2020. Abricate. Github.

SHAH-MAJID, M., AZLINA, A. M., ANA MARIA, A. R., ZAHARAH, B. & NORHALIZA, A. H. 2004. Occurrence of vancomycin-resistant enterococci in ducks in Malaysia. *Vet Rec,* 155**,** 680-1.

SHAQ KOYOK., D. A. 2020. A Future Rooted in our Forests. *In:* DEBORAH AUGUSTIN, S. K. (ed.) *Southeast Asia Dispatches.* New Naratif.

SHAW, L., RIBEIRO, A. L. R., LEVINE, A. P., PONTIKOS, N., BALLOUX, F., SEGAL, A. W., ROBERTS, A. P. & SMITH, A. M. 2017. The Human Salivary Microbiome Is Shaped by Shared Environment Rather than Genetics: Evidence from a Large Family of Closely Related Individuals. *mBio,* 8.

SHERWANI, S. I., KHAN, H. A., EKHZAIMY, A., MASOOD, A., & SAKHARKAR, M. K. 2016. Significance of HbA1c Test in Diagnosis and Prognosis of Diabetic Patients. *Biomarker insights*, *11*, 95–104. https://doi.org/10.4137/BMI.S38440

SHI, X.-R., CHEN, B.-Y., LIN, W.-Z., LI, Y.-L., WANG, Y.-L., LIU, Y., HUANG, J.-J., ZHANG, W.-W., MA, X.-X., SHAO, S., LI, R.-G. & DUAN, S.-Z. 2021. Microbiota in Gut, Oral Cavity, and Mitral Valves Are Associated With Rheumatic Heart Disease. *Frontiers in Cellular and Infection Microbiology,* 11.

SINGH, R. K., CHANG, H.-W., YAN, D., LEE, K. M., UCMAK, D., WONG, K., ABROUK, M., FARAHNIK, B., NAKAMURA, M. & ZHU, T. H. 2017. Influence of diet on the gut microbiome and implications for human health. *Journal of translational medicine,* 15**,** 1-17.

SINHA, R., ABU-ALI, G., VOGTMANN, E., FODOR, A. A., REN, B., AMIR, A., SCHWAGER, E., CRABTREE, J., MA, S., THE MICROBIOME QUALITY CONTROL PROJECT, C., ABNET, C. C., KNIGHT, R., WHITE, O. & HUTTENHOWER, C. 2017. Assessment of variation in microbial community amplicon sequencing by the Microbiome Quality Control (MBQC) project consortium. *Nature Biotechnology,* 35**,** 1077.

SOKOL, H., PIGNEUR, B., WATTERLOT, L., LAKHDARI, O., BERMÚDEZ-HUMARÁN, L. G., GRATADOUX, J.-J., BLUGEON, S., BRIDONNEAU, C., FURET, J.-P., CORTHIER, G., GRANGETTE, C., VASQUEZ, N., POCHART, P., TRUGNAN, G., THOMAS, G., BLOTTIÈRE, H. M., DORÉ, J., MARTEAU, P., SEKSIK, P. & LANGELLA, P. 2008. Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proceedings of the National Academy of Sciences,* 105**,** 16731-16736.

SOMMER, A. J., PETERS, A., ROMMEL, M., CYRYS, J., GRALLERT, H., HALLER, D., MÜLLER, C. L. & BIND, M.-A. C. 2021. A randomization-based causal inference framework for uncovering environmental exposure effects on human gut microbiota. *bioRxiv*.

SOMMER, M. O. A., DANTAS, G. & CHURCH, G. M. 2009. Functional characterization of the antibiotic resistance reservoir in the human microflora. *Science,* 325**,** 1128-1131.

SPERR, E. 2016. *PubMed by Year* [Online]. [Accessed].

STOLL, M. L. 2020. Genetics, Prevotella, and the pathogenesis of rheumatoid arthritis. *The Lancet Rheumatology,* 2**,** e375-e376.

SYAM, A. F., MIFTAHUSSURUR, M., MAKMUN, D., NUSI, I. A., ZAIN, L. H., AKIL, F., USWAN, W. B., SIMANJUNTAK, D., UCHIDA, T. & ADI, P. 2015. Risk factors and prevalence of Helicobacter pylori in five largest islands of Indonesia: A preliminary study. *PloS one,* 10**,** e0140186.

TAN, D., WEE, S. K., JAGANATHAN KING, S. & SINGH, J. 2006. *SCREENING AND CHARACTERIZATION OF VANCOMYCIN-RESISTANT ENTEROCOCCI (VRE) ISOLATED FROM POULTRY IN MALAYSIA*.

TAN, S. C., CHONG, C. W., TEH, C. S. J., OOI, P. T. & THONG, K. L. 2018. Occurrence of virulent multidrug-resistant Enterococcus faecalis and Enterococcus faecium in the pigs, farmers and farm environments in Malaysia. *PeerJ,* 6**,** e5353-e5353.

TENG, C. L., TONG, S. F., KHOO, E. M., LEE, V., ZAILINAWATI, A. H., MIMI, O., CHEN, W. S. & NORDIN, S. 2011. Antibiotics for URTI and UTI: prescribing in Malaysian primary care settings. *Australian family physician,* 40**,** 325.

TENG, M.-S., HSU, L.-A., WU, S., SUN, Y.-C., JUAN, S.-H. & KO, Y.-L. 2015. Association of CDH13 Genotypes/Haplotypes with Circulating Adiponectin Levels, Metabolic Syndrome, and Related Metabolic Phenotypes: The Role of the Suppression Effect. *PLOS ONE,* 10**,** e0122664.

TETT, A., HUANG, K. D., ASNICAR, F., FEHLNER-PEACH, H., PASOLLI, E., KARCHER, N., ARMANINI, F., MANGHI, P., BONHAM, K., ZOLFO, M., DE FILIPPIS, F., MAGNABOSCO, C., BONNEAU, R., LUSINGU, J., AMUASI, J., REINHARD, K., RATTEI, T., BOULUND, F., ENGSTRAND, L., ZINK, A., COLLADO, M. C., LITTMAN, D. R., EIBACH, D., ERCOLINI, D., ROTA-STABELLI, O., HUTTENHOWER, C., MAIXNER, F. & SEGATA, N. 2019. The Prevotella copri Complex Comprises Four Distinct Clades Underrepresented in Westernized Populations. *Cell Host & Microbe,* 26**,** 666-679.e7.

THEVAKUMAR, K., CHANDREN, J. R., PEREZ-PEREZ, G. I., CHUA, E. G., TEH, L. K., SALLEH, M. Z., TAN, J. A., LEOW, A. H., GOH, K. L., TAY, A. C., MARSHALL, B. J., VADIVELU, J., LOKE, M. F. & WONG, L. P. 2016. Assessment of Risk and Sero-Prevalence of Helicobacter pylori Colonization among Remote Orang Asli Tribes in Peninsula Malaysia. *PLoS One,* 11**,** e0159830.

TOKUDOME, S., TRININGSIH, F. E., ANANTA, I., SUZUKI, S., KURIKI, K., AKASAKA, S., KOSAKA, H., ISHIKAWA, H., AZUMA, T. & MOORE, M. A. 2005. Rare Helicobacter pylori infection as a factor for the very low stomach cancer incidence in Yogyakarta, Indonesia. *Cancer letters,* 219**,** 57-61.

TOMOVA, A., BUKOVSKY, I., REMBERT, E., YONAS, W., ALWARITH, J., BARNARD, N. D. & KAHLEOVA, H. 2019. The Effects of Vegetarian and Vegan Diets on Gut Microbiota. *Frontiers in Nutrition,* 6.

TONG, X., XU, J., LIAN, F., YU, X., ZHAO, Y., XU, L., ZHANG, M., ZHAO, X., SHEN, J., WU, S., PANG, X., TIAN, J., ZHANG, C., ZHOU, Q., WANG, L., PANG, B., CHEN, F., PENG, Z., WANG, J., ZHEN, Z., FANG, C., LI, M., CHEN, L., ZHAO, L., BIK, E. M., BLASER, M. J., WANG, J. & XIAO, X. 2018. Structural Alteration of Gut Microbiota during the Amelioration of Human Type 2 Diabetes with Hyperlipidemia by Metformin and a Traditional Chinese Herbal Formula: a Multicenter, Randomized, Open Label Clinical Trial. *mBio,* 9**,** e02392-17.

TOOSA, H., RADU, S., RUSUL, G., LATIF, A. R., RAHIM, R. A., AHMAD, N. & LING, O. W. 2001. Detection of Vancomycin-Resistant Enterococcus Spp. (VRE) from Poultry. *The Malaysian journal of medical sciences : MJMS,* 8**,** 53-58.

TREERAT, P., REDANZ, U., REDANZ, S., GIACAMAN, R. A., MERRITT, J. & KRETH, J. 2020. Synergism between Corynebacterium and Streptococcus sanguinis reveals new interactions between oral commensals. *The ISME Journal,* 14**,** 1154-1169.

TSIGALOU, C., KONSTANTINIDIS, T., STAVROPOULOU, E., BEZIRTZOGLOU, E. E. & TSAKRIS, A. 2020. Potential Elimination of Human Gut Resistome by Exploiting the Benefits of Functional Foods. *Frontiers in Microbiology,* 11.

TUAN ABDUL AZIZ, T. A., TEH, L. K., MD IDRIS, M. H., BANNUR, Z., ASHARI, L. S., ISMAIL, A. I., AHMAD, A., ISA, K. M., NOR, F. M., RAHMAN, T. H. A., SHAARI, S. A. B., JAN MOHAMED, H. J., MOHAMAD, N. & SALLEH, M. Z. 2016. Increased risks of cardiovascular diseases and insulin resistance among the Orang Asli in Peninsular Malaysia. *BMC Public Health,* 16**,** 284.

TULLY, B. J., GRAHAM, E. D. & HEIDELBERG, J. F. 2018. The reconstruction of 2,631 draft metagenome-assembled genomes from the global oceans. *Scientific Data,* 5**,** 170203.

UNITED NATIONS 2019. World Population Prospects 2019: Volum I: Comprehensive Tables. *In:* DEPARMENT OF ECONOMIC AND SOCIAL AFFAIRS, P. D. (ed.).

URITSKIY, G. V., DIRUGGIERO, J. & TAYLOR, J. 2018. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome,* 6**,** 158.

UTTLEY, A. C., COLLINS, C. H., NAIDOO, J. & GEORGE, R. C. 1988. VANCOMYCIN-RESISTANT ENTEROCOCCI. *The Lancet,* 331**,** 57-58.

VAN SCHAIK, W. 2015. The human gut resistome. *Philos Trans R Soc Lond B Biol Sci,* 370**,** 20140087.

VANHATALO, A., L'HEUREUX, J. E., KELLY, J., BLACKWELL, J. R., WYLIE, L. J., FULFORD, J., WINYARD, P. G., WILLIAMS, D. W., VAN DER GIEZEN, M. & JONES, A. M. 2021. Network analysis of nitrate-sensitive oral microbiome reveals interactions with cognitive function and cardiovascular health across dietary interventions. *Redox Biology,* 41**,** 101933.

VERHOOG, S., TANERI, P. E., ROA DÍAZ, Z. M., MARQUES-VIDAL, P., TROUP, J. P., BALLY, L., FRANCO, O. H., GLISIC, M. & MUKA, T. 2019. Dietary Factors and Modulation of Bacteria Strains of Akkermansia muciniphila and Faecalibacterium prausnitzii: A Systematic Review. *Nutrients,* 11**,** 1565.

VIDENSKA, P., SMERKOVA, K., ZWINSOVA, B., POPOVICI, V., MICENKOVA, L., SEDLAR, K. & BUDINSKA, E. 2019. Stool sampling and DNA isolation kits affect DNA quality and bacterial composition following 16S rRNA gene sequencing using MiSeq Illumina platform. *Scientific Reports,* 9**,** 13837.

WADA, Y., HARUN, A., YEAN, C. & ZAIDAH, A. 2019. Vancomycin-Resistant Enterococcus: Issues in Human Health, Animal Health, Resistant Mechanisms and the Malaysian Paradox. *Advances in Animal and Veterinary Sciences,* 7.

WADE, K. H. & HALL, L. J. 2019. Improving causality in microbiome research: can human genetic epidemiology help? *Wellcome Open Research,* 4.

WADE, W. G. 2021. Resilience of the oral microbiome. *Periodontology 2000,* 86**,** 113-122.

WALTER, J., ARMET, A. M., FINLAY, B. B. & SHANAHAN, F. 2020. Establishing or Exaggerating Causality for the Gut Microbiome: Lessons from Human Microbiota-Associated Rodents. *Cell,* 180**,** 221-232.

WANG, D., LI, Y., ZHONG, H., DING, Q., LIN, Y., TANG, S., ZONG, Y., WANG, Q., ZHANG, X., YANG, H., WANG, R. & LIU, X. 2019. Alterations in the human gut microbiome associated with Helicobacter pylori infection. *FEBS open bio,* 9**,** 1552-1560.

WELLS, J. C. K., MARPHATIA, A. A., COLE, T. J. & MCCOY, D. 2012. Associations of economic and gender inequality with global obesity prevalence: Understanding the female excess. *Social Science & Medicine,* 75**,** 482-490.

WEN, J., ZHONG, Y., KUANG, C., LIAO, J., CHEN, Z. & YANG, Q. 2017. Lipoprotein ratios are better than conventional lipid parameters in predicting arterial stiffness in young men. *The Journal of Clinical Hypertension*, *19*(8), 771-776.

WHO 2011. World Health Organization: Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus. Abbreviated report of a WHO consultation 2011.

WHO, O. 2008. The Right to Health: Fact Sheet No 31. United Nations, Geneva: Office of the United Nations High Commissioner for Human Rights.

WILLMANN, M., VEHRESCHILD, M., BIEHL, L. M., VOGEL, W., DÖRFEL, D., HAMPRECHT, A., SEIFERT, H., AUTENRIETH, I. B. & PETER, S. 2019. Distinct impact of antibiotics on the gut microbiome and resistome: a longitudinal multicenter cohort study. *BMC Biol,* 17**,** 76.

WOERTHER, P.-L., BURDET, C., CHACHATY, E. & ANDREMONT, A. 2013. Trends in human fecal carriage of extended-spectrum β-lactamases in the community: toward the globalization of CTX-M. *Clinical microbiology reviews,* 26**,** 744-758.

WONG, C. Y., ZALILAH, M. S., CHUA, E. Y., NORHASMAH, S., CHIN, Y. S. & SITI NUR'ASYURA, A. 2015. Double-burden of malnutrition among the indigenous peoples (Orang Asli) of Peninsular Malaysia. *BMC Public Health,* 15**,** 680.

WONG, W.-K., MOHD-NOR, N., NOORDIN, R., FOO, P.-C., MOHAMED, Z., HAQ, J. A., ACOSTA, A., SARMIENTO, M. E., SUBRAMANIAM, P. & DONY, J. F. 2019a. Parasitic infections in Malaysian aborigines with pulmonary tuberculosis: a comparative cross-sectional study. *Parasitology research,* 118**,** 2635-2642.

WONG, Y. S., ALLOTEY, P. & REIDPATH, D. D. 2019b. Why we run when the doctor comes: Orang Asli responses to health systems in transition in Malaysia. *Critical Public Health,* 29**,** 192-204.
WOOD, D. E., LU, J. & LANGMEAD, B. 2019. Improved metagenomic analysis with Kraken 2. *Genome Biology,* 20**,** 257.
WU, D., JIAO, N., ZHU, R., ZHANG, Y., GAO, W., FANG, S., LI, Y., CHENG, S., TIAN, C., LAN, P., LOOMBA, R. & ZHU, L. 2020. Identification of the keystone species in non-alcoholic fatty liver disease by causal inference and dynamic intervention modeling. *bioRxiv***,** 2020.08.06.240655.
WU, G., ZHAO, N., ZHANG, C., LAM, Y. Y. & ZHAO, L. 2021. Guild-based analysis for understanding gut microbiome in human health and diseases. *Genome Medicine,* 13**,** 22.
WU, G. D., CHEN, J., HOFFMANN, C., BITTINGER, K., CHEN, Y.-Y., KEILBAUGH, S. A., BEWTRA, M., KNIGHTS, D., WALTERS, W. A. & KNIGHT, R. 2011. Linking long-term dietary patterns with gut microbial enterotypes. *Science,* 334**,** 105-108.
WU, H., ESTEVE, E., TREMAROLI, V., KHAN, M. T., CAESAR, R., MANNERÅS-HOLM, L., STÅHLMAN, M., OLSSON, L. M., SERINO, M., PLANAS-FÈLIX, M., XIFRA, G., MERCADER, J. M., TORRENTS, D., BURCELIN, R., RICART, W., PERKINS, R., FERNÀNDEZ-REAL, J. M. & BÄCKHED, F. 2017. Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug. *Nature Medicine,* 23**,** 850.
YAN, Y., NGUYEN, L. H., FRANZOSA, E. A. & HUTTENHOWER, C. 2020. Strain-level epidemiology of microbial communities and the human microbiome. *Genome Medicine,* 12**,** 71.
YANG, L., ZHANG, J., XU, J., WEI, X., YANG, J., LIU, Y., LI, H., ZHAO, C., WANG, Y., ZHANG, L. & GAI, Z. 2019. Helicobacter pylori Infection Aggravates Dysbiosis of Gut Microbiome in Children With Gastritis. *Frontiers in Cellular and Infection Microbiology,* 9.
YAP, T. W.-C., GAN, H.-M., LEE, Y.-P., LEOW, A. H.-R., AZMI, A. N., FRANCOIS, F., PEREZ-PEREZ, G. I., LOKE, M.-F., GOH, K.-L. & VADIVELU, J. 2016. Helicobacter pylori eradication causes perturbation of the human gut microbiome in young adults. *PLoS One,* 11**,** e0151893.
YEH, W.-C., TSAO, Y.-C., LI, W.-C., TZENG, I. S., CHEN, L.-S. & CHEN, J.-Y. 2019. Elevated triglyceride-to-HDL cholesterol ratio is an indicator for insulin resistance in middle-aged and elderly Taiwanese population: a cross-sectional study. *Lipids in Health and Disease,* 18**,** 176.
YEO, L.-F., AGHAKHANIAN, F. F., TAN, J. S., GAN, H. M. & PHIPPS, M. E. 2019. Health and saliva microbiomes of a semi-urbanized indigenous tribe in Peninsular Malaysia. *F1000Research,* 8.
YEO, L-F. 2017. Investigation of anthropometrics, cardio-metabolic health and the oral microbiome among Temiar [Postgraduate Diploma Thesis, Monash University Malaysia]. Figshare 10.26180/14909160
YEOH, Y. K., CHAN, M. H., CHEN, Z., LAM, E. W. H., WONG, P. Y., NGAI, C. M., CHAN, P. K. S. & HUI, M. 2019. The human oral cavity microbiota composition during acute tonsillitis: a cross-sectional survey. *BMC Oral Health,* 19**,** 275.
YONG, D., LU, D., LI, N., TING, S., XIAN, S., LIQUN, W., SHULAN, H., ZHIZHONG, W., ZHENHUI, L., FENG, L., XIAOFEI, G., YANHONG, Z., JIN, T., BO, W., SIFAN, Z., CHAO, Y. & XIAO, L. 2021. Prevalence of Chronic Diseases and Alterations of Gut Microbiome in Ningxia China During Urbanization: An Epidemiological Survey. *Research Square*.
ZAFAR, H. & SAIER, M. H. 2021. Gut Bacteroides species in health and disease. *Gut Microbes,* 13**,** 1-20.
ZHANG, J., GUO, Z., LIM, A. A. Q., ZHENG, Y., KOH, E. Y., HO, D., QIAO, J., HUO, D., HOU, Q. & HUANG, W. 2014. Mongolians core gut microbiota and its correlation with seasonal dietary changes. *Scientific reports,* 4**,** 1-11.
ZHANG, L., LIU, Y., ZHENG, H. J. & ZHANG, C. P. 2020. The Oral Microbiota May Have Influence on Oral Cancer. *Frontiers in Cellular and Infection Microbiology,* 9.
ZHAO, L., ZHANG, F., DING, X., WU, G., LAM, Y. Y., WANG, X., FU, H., XUE, X., LU, C., MA, J., YU, L., XU, C., REN, Z., XU, Y., XU, S., SHEN, H., ZHU, X., SHI, Y., SHEN, Q., DONG, W., LIU, R., LING, Y., ZENG, Y., WANG, X., ZHANG, Q., WANG, J., WANG, L., WU, Y., ZENG, B., WEI, H., ZHANG, M., PENG, Y. & ZHANG, C. 2018. Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes. *Science,* 359**,** 1151-1156.

ZHU, L., LU, Z., ZHU, L., OUYANG, X., YANG, Y., HE, W., FENG, Y., YI, F., & SONG, Y. 2015. Lipoprotein ratios are better than conventional lipid parameters in predicting coronary heart disease in Chinese Han people. *Kardiologia polska*, *73*(10), 931–938.

ZHU, X., QIN, J., TAN, C. & NING, K. 2021. The seasonal changes of the gut microbiome of the population living in traditional lifestyles are represented by characteristic species-level and functional-level SNP enrichment patterns. *BMC Genomics,* 22**,** 83.