
Microbiomes of Long-Tailed Macaques in Southeast Asia: Prokaryotic and Eukaryotic Interactions**Chissa-Louise Rivaldi****Publication Date**

10-07-2022

License

This work is made available under a CC BY 4.0 license and should only be used in accordance with that license.

Citation for this work (American Psychological Association 7th edition)

Rivaldi, C.-L. (2022). *Microbiomes of Long-Tailed Macaques in Southeast Asia: Prokaryotic and Eukaryotic Interactions* (Version 1). University of Notre Dame. <https://doi.org/10.7274/7d278s48v09>

This work was downloaded from CurateND, the University of Notre Dame's institutional repository.

For more information about this work, to report or an issue, or to preserve and share your original work, please contact the CurateND team for assistance at curate@nd.edu.

MICROBIOMES
OF LONG-TAILED MACAQUES IN SOUTHEAST ASIA:
PROKARYOTIC AND EUKARYOTIC INTERACTIONS

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

Chissa-Louise Rivaldi

Hope Hollocher, Director

Graduate Program in Biological Sciences

Notre Dame, Indiana

July 2022

© Copyright by
Chissa-Louise Rivaldi
2022
CC-BY-4.0

MICROBIOMES
OF LONG-TAILED MACAQUES IN SOUTHEAST ASIA:
PROKARYOTIC AND EUKARYOTIC INTERACTIONS

Abstract

by

Chissa-Louise Rivaldi

The interior and exterior margins of vertebrate bodies are blurred by the microbes occupying these spaces. Microbial assemblages associated with a host are thus built and shaped by a multitude of sources which originate from both the host and the environment. Host functions and external environment exhibit *inside-out* control over residential and introduced bacteria, leading to the fluctuation or depletion of a host-associated bacterial community composition. Bacteria exhibit *outside-in* behavior in turn, presenting products and functions of their own immune system to the host. In addition, host behaviors such as foraging, eating, and physical contact with other hosts provide nearly constant influxes of microbes which act to replenish populations and introduce competition to existing communities. Eukaryotic microbes, such as protozoans and helminths, add to the already incredibly dynamic mixture of organisms associated with a mammalian host, inducing control over cohabiting microbes as well as the host.

This dissertation aims to analyze host-associated microbiome dynamics on varying scales and increase our understanding of the underlying interactions by expanding our scope to include eukaryotes. First, I examine site-specific bacterial communities of long-tailed macaques in Singapore, concentrating on the oral and

gut microbiomes and how they vary through geographical space. Revealed in this work is evidence showing that, though oral microbes have greater diversity than gut microbiomes, they show greater similarity across geographic space, which reflects their role as a barrier between hosts and the environment in which they live. Next, I expand this work to Bali, Indonesia, and examine the differences of gut bacterial microbiomes across landscapes through host diet. I find macaque diets vary with location, and that this variation is associated with changes in the gut bacterial community. I then show how macaque consumption of plants with medicinal properties is tied to parasite presence by incorporating a medicinal database as a framework to evaluate diet-parasite interactions. Finally, I find variation of gut parasites of macaques to be driven by the protozoan *Blastocystis* and extend this finding to analyze the prokaryotic microbiome, the diversity and composition of which shifts in the presence of *Blastocystis*. Collectively, this work shows that without considering eukaryotic context, aims to understand mechanisms driving host-microbiome relationships fall short.

For Tinx.

CONTENTS

| | |
|--|----|
| Figures | v |
| Tables | ix |
| Acknowledgments | xi |
| Chapter 1: Introduction | 1 |
| 1.1 Overview of Literature | 1 |
| 1.2 Variation in Prokaryotic Microbiomes | 2 |
| 1.3 Parasite interactions with the microbiome | 5 |
| 1.4 Project Overview | 7 |
| Chapter 2: Environmental and Host Effects on Oral and Gut Bacterial Com- munities in Free-ranging Long-tailed Macaques in Singapore | 12 |
| 2.1 Abstract | 12 |
| 2.2 Background | 13 |
| 2.3 Methods | 16 |
| 2.3.1 Study Site, Sampling, and DNA extraction | 16 |
| 2.3.2 Target Amplification and High-Throughput Sequencing | 17 |
| 2.3.3 Bioinformatic Workflow and Statistical Analysis | 19 |
| 2.4 Results | 21 |
| 2.5 Discussion | 37 |
| 2.6 Conclusions | 42 |
| Chapter 3: Diet Presents Opportunities To Study The Role Of Human Inter- actions In Shaping Gut Microbiome In Urban Environments | 43 |
| 3.1 Abstract | 43 |
| 3.2 Background | 44 |
| 3.3 Methods | 45 |
| 3.3.1 Sampling and DNA Extraction | 45 |
| 3.3.2 16S and 18S rRNA Read Processing | 46 |
| 3.3.3 Statistical Analysis | 47 |
| 3.4 Results | 48 |
| 3.5 Discussion | 55 |
| 3.6 Conclusions | 58 |

| | |
|--|-----|
| Chapter 4: Evidence for Self-Medication in Long-Tailed Macaques in South-east Asia | 60 |
| 4.1 Abstract | 60 |
| 4.2 Background | 61 |
| 4.3 Methods | 63 |
| 4.3.1 Sample Collection | 63 |
| 4.3.2 DNA extraction/Sequencing | 64 |
| 4.3.3 16S and 18S Read Processing | 64 |
| 4.3.4 Medicinal Plant Database Construction | 65 |
| 4.3.5 Statistical analysis | 65 |
| 4.4 Results | 66 |
| 4.4.1 Medicinal plant categorization | 66 |
| 4.4.2 Protozoan and Helminth Associations with Medicinal Plants | 69 |
| 4.5 Discussion | 79 |
| 4.6 Conclusions | 82 |
| Chapter 5: <i>Blastocystis</i> is Associated with Shifts in Prokaryotic Diversity and Composition in Free-Ranging Long-Tailed Macaques | 83 |
| 5.1 Abstract | 83 |
| 5.2 Background | 84 |
| 5.3 Methods | 86 |
| 5.3.1 Sample Collection and DNA Processing | 86 |
| 5.3.2 Read Processing | 87 |
| 5.3.3 Statistical Analysis | 87 |
| 5.4 Results | 89 |
| 5.5 Discussion | 103 |
| 5.6 Conclusions | 105 |
| Appendix A: Appendix | 106 |
| A.1 Chapter 2 Supplementary Materials | 106 |
| A.2 Chapter 3 Supplementary Materials | 110 |
| A.2.1 16S 18S Supplementary Methods | 110 |
| A.2.1.1 Library prep for Illumina sequencing | 110 |
| A.2.1.2 18S Ribosomal RNA Gene Information | 110 |
| A.2.1.3 16S Ribosomal RNA Gene Information | 112 |
| A.3 Chapter 3 Supplementary Data | 116 |
| A.4 Chapter 4 Supplementary Materials | 119 |
| Bibliography | 122 |

FIGURES

| | | |
|-----|--|----|
| 2.1 | Sampling sites indicate populations from which samples were obtained. Samples were collected throughout accessible macaque habitat. Geographic coordinates and number of samples for each site are as follows: AP - Admiralty Park: 1° 27' 2.9844" N, 103° 46' 38.9964" E (N = 12); BB - Bukit Batok: 1° 22' 36.2900" N, 103° 45' 30.6242" E (N = 3); BSI - Big Sister's Island: 1° 12' 52.4916" N, 103° 50' 6.3055" E (N = 8); BT - Bukit Timah: 1° 22' 47.0205" N, 103° 46' 25.42636" E (N = 11); MR - MacRitchie Reserve 1° 20' 28.0572" N, 103° 49' 48.5616" E (N = 17); OUT - Old Upper Thomson Road: 1° 22' 49.7233" N, 103° 49' 6.3001" E (N = 15); SI - Sentosa Island: 1° 15' 29.3885" N, 103° 48' 33.5866" E (N = 4); UP - Upper Pierce: 1° 22' 32.2896" N, 103° 48' 17.3988" E (N = 6); US - Upper Seletar: 1° 23' 55.5432" N, 103° 48' 23.2812 E; WW - Woodlands Waterfront: 1° 27' 6.6652" N, 103° 46' 52.2617" E (N = 11). | 18 |
| 2.2 | Family richness (A), dissimilarity (B), and evenness (C) of sites containing both saliva and fecal sampling - Admiralty Park (AP), Bukit Timah (BT), MacRitchie Nature Reserve (MR), and Woodlands Waterfront (WW). Saliva samples display greater richness, similarity, and evenness throughout sites. | 22 |
| 2.3 | Relative abundance of bacterial phyla in oral and gut microbiomes. Samples are ordered left to right by decreasing F/B ratio. | 25 |
| 2.4 | LEfSe analysis of family level taxa in oral and gut microbiome. Taxa shown scored at least 2.0 in Linear Discriminant Analysis (57 total). Bars are colored by phylum corresponding to Figure 2.3. For taxa classified as undefined or uncultured at the family level, class name is provided. | 27 |
| 2.5 | Non-metric Multidimensional Scaling of ASVs colored by body site (stress = 0.07). Saliva samples group together more tightly, while fecal samples show greater variation. | 28 |
| 2.6 | Distribution of F/B ratio in oral and gut microbiomes (A). Oral and gut samples ordered by log(F/B ratio). Families which showed significant differences between Bacteroidetes-rich and Firmicutes-rich bacterial communities in the gut environment are shown in (B). Taxa are described with family and phylum-level. Specific p-values for each taxon listed in Table 2.2 | 29 |

| | | |
|-----|---|----|
| 3.1 | K-Means Clustering of Combined Samples | 49 |
| 3.2 | Linear Relationship Between Chloroplasts and Non-Crop Plant Reads. Relationship is positive and significant ($\beta = 3.51e^{-03}$, $R^2 = 0.156$, $p < 0.001$). Chloroplast counts obtained from 16S rRNA reads and non-crop counts are calculated from 18S rRNA reads | 51 |
| 3.3 | Crop and Non-Crop Plant Reads Between Clusters. Samples in Cluster 1 (red) have fewer relative crop plants than samples in Cluster 3 (green; $t = -2.6305$, $df = 20.124$, $p = 0.016$). Miscellaneous plants make up more of the diet in samples from Cluster 1 than samples in Clusters 2 and 3 (brown; 1 and 2: $t = -5.1049$, $df = 66.85$, $p = 2.95 \times 10^{-06}$; 1 and 3: $t = 2.495$, $df = 23.517$, $p = 0.020$). . | 52 |
| 3.4 | Linear relationships between bacterial ASV richness and crop richness (left) and dietary genera richness (right) in Singapore and Bali samples. Crop richness and dietary genera richness both show a linear relationship with bacterial richness (crop: $R^2 = 0.109$, $p = 0.001$, dietary genera $R^2 = 0.190$, $p = 1.67 \times 10^{-05}$). | 53 |
| 3.5 | Dietary genera richness (left) and crop genera richness (right) by bacteria cluster. Only samples in Singapore and Bali are tested and plotted. Cluster 1 has higher dietary and crop richness than Clusters 2 or 3 (dietary richness: 1-2 $p=0.004$, 1-3 $p = 0.007$; crop richness: 1-2 $p = 0.04$, 1-3 $p=0.04$). | 54 |
| 4.1 | Intersections of plants in medicinal categories. The medicinal plant database categorizes plants by their documented medicinal properties (Anti-helminthic, Anti-protozoan, Anti-fungal, Anti-bacterial, and human medicinal use). The far-right column contains plant taxa which have combined medicinal properties, also referred to as “Anti-All” throughout analyses. | 68 |
| 4.2 | General medicinal plant abundance in <i>Plasmodium</i> -positive and <i>Plasmodium</i> -negative samples. (A-C) Medicinal plants were more abundant when <i>Plasmodium</i> was present (Welch’s $t = -4.654$, $p = 1.101 \times 10^{-5}$), Singapore (Welch’s $t = -4.190$, $p = 4.055 \times 10^{-4}$), and Bali (Welch’s $t = -3.449$, $p = 1.028 \times 10^{-3}$). (D-F) General medicinal plant abundance is not associated with <i>Plasmodium</i> abundance. | 75 |
| 4.3 | Anti-all plant abundance in <i>Plasmodium</i> -positive and <i>Plasmodium</i> -negative samples. (A-C) Anti-all plants are more abundant in <i>Plasmodium</i> -positive across samples (combined samples: Welch’s $t = -5.709$, $p = 8.64 \times 10^{-6}$; Singapore: Welch’s $t = -5.274$, $p = 2.02 \times 10^{-6}$; Bali: Welch’s $t = -2.955$, $p = 4.63 \times 10^{-3}$). (D-F) Anti-all plants are positively associated with <i>Plasmodium</i> abundance in all samples and Singapore (combined samples: $R^2 = 0.304$, $p = 1.78 \times 10^{-11}$; Singapore: $R^2 = 0.305$, $p = 2.75 \times 10^{-6}$, but not Bali. | 76 |

| | | |
|-----|--|----|
| 4.4 | General Medicinal plant richness in <i>Plasmodium</i> -positive and <i>Plasmodium</i> -negative samples. (A-C) Medicinal plant richness is higher in <i>Plasmodium</i> -positive samples (all samples: Welch's $t = -6.750$, $p = 3.63 \times 10^{-6}$; Singapore: Welch's $t = -3.914$, $p = 9.71 \times 10^{-3}$; Bali: Welch's $t = -2.846$, $p = 0.010$). (D-F) <i>Plasmodium</i> abundance is positively related to medicinal plant richness in combined samples and Bali, but not Singapore (all samples: $R^2 = 0.09774$, $p = 3.455 \times 10^{-4}$; Bali: $R^2 = 0.146$, $p = 0.002$). | 77 |
| 4.5 | Anti-all plant observed richness in <i>Plasmodium</i> -positive and <i>Plasmodium</i> -negative samples. (A-C) Anti-all plant richness is higher in <i>Plasmodium</i> -positive samples across data (all samples: Welch's $t = -9.527$, $p = 5.54 \times 10^{-14}$; Singapore: Welch's $t = -5.429$, $p = 0.001$; Bali: Welch's $t = -4.750$, $p = 4.22 \times 10^{-5}$). (D-F) <i>Plasmodium</i> abundance is positively related to Anti-all plant richness across samples (combined samples: $R^2 = 0.115$, $p = 9.37 \times 10^{-5}$; Singapore: $R^2 = 0.100$, $p = 0.01$; Bali $R^2 = 0.098$, $p = 0.01$). | 78 |
| 5.1 | Top correlating taxa (loading scores $> \pm 0.1$). <i>Blastocystis</i> is a leading taxon for PC1 in all samples combined, and Singapore and Bali individually. Taxa were Hellinger transformed before PCA. | 90 |
| 5.2 | Boxplots of alpha diversity metrics compared to <i>Blastocystis</i> presence (P) and absence (A). Prokaryotic alpha diversity is significantly higher in the presence of <i>Blastocystis</i> using Shannon Diversity (A) and richness (C) as metrics. Evenness (B) does not differ in the presence of <i>Blastocystis</i> . P-values shown in title represents test from all samples combined. In Singapore, Shannon Diversity and richness are significantly higher when <i>Blastocystis</i> is present. In Bali, Shannon Diversity and richness are not different when <i>Blastocystis</i> is present, while evenness is higher. Asterisks denote significance from Wilcoxon-Rank Sum tests controlled for location. | 93 |
| 5.3 | Prokaryotic alpha diversity metric relationships with <i>Blastocystis</i> abundance. A) Shannon diversity of prokaryotes increases as <i>Blastocystis</i> abundance increases in all samples and with Singapore alone. B) Pielou evenness of prokaryotes is significantly associated with <i>Blastocystis</i> abundance in all samples and in Singapore alone. C) Richness, or observed taxa, is significantly associated with <i>Blastocystis</i> abundance in Singapore but not in all samples or Bali. <i>Blastocystis</i> displayed on a logarithmic scale for visibility of samples. . . . | 96 |
| 5.4 | Multidimensional scaling using Euclidean distances. Ellipses group samples by <i>Blastocystis</i> presence (represented by colors) and location (grey dashed lines) | 97 |

| | | |
|-----|---|-----|
| A.1 | Rarefaction curves for saliva samples (top) and fecal samples (bottom). Curves were created after filtering non-prokaryotic ASVs from reads, all samples show plateau before reaching read depth. . . . | 107 |
|-----|---|-----|

TABLES

| | | |
|-----|---|-----|
| 2.1 | Body site and sampling location factors included in PERMANOVA phyla and families of ASVs. Significant variables and values in bold for clarity | 23 |
| 2.2 | Families with significant changes between Bacteroidetes-rich and Firmicutes-rich gut microbiomes. Significance tested with Mann-Whitney U tests, α of 0.05 | 30 |
| 2.3 | Predicted MetaCyC pathways with an effect size of > 1.0 as estimated by ALDEx2 – comparison for oral or gut samples. | 31 |
| 2.4 | Predicted MetaCyC pathways with an effect size of > 1.0 as estimated by ALDEx2 – comparison for Firmicutes- or Bacteroidetes-rich gut samples. | 35 |
| 3.1 | Sample Breakdown By Cluster | 49 |
| 3.2 | Highly correlating bacterial genera with clusters | 50 |
| 4.1 | Ten most abundant plants in each category used for analysis. | 67 |
| 4.2 | Unique Plant Items in Medicinal Categories | 69 |
| 4.3 | Selected helminth and protozoan genera recovered in 18S rRNA data sorted by abundance | 70 |
| 4.4 | Medicinal and Non-Medicinal Plant Correlations to Parasites | 72 |
| 5.1 | Component variation and loading scores of correlating protozoa | 90 |
| 5.2 | PERMANOVA: Variation in Community Composition | 94 |
| 5.3 | Differentially abundant bacterial taxa in <i>Blastocystis</i> -positive samples. | 98 |
| A.1 | Mean Alpha Diversity metrics among all sites | 108 |
| A.2 | 18S Ribosomal RNA Gene Primer Information | 111 |
| A.3 | 18S Ribosomal RNA Gene PCR reaction for both 1380F and 1389F with 1510R | 112 |
| A.4 | 18S ribosomal RNA Gene PCR Thermocycler Conditions for primers 1380F and 1510R | 113 |
| A.5 | 18S ribosomal RNA Gene PCR Thermocycler Conditions for primers 1389F and 1510R | 113 |

| | | |
|------|---|-----|
| A.6 | 16S Ribosomal RNA Gene Primer Information | 114 |
| A.7 | 16S Ribosomal RNA Gene PCR reaction for primers PrvkV4-F and PrvkV4-R | 115 |
| A.8 | 16S ribosomal RNA Gene PCR Thermocycler Conditions for primers PrkV4-F and PvkV4-R | 115 |
| A.9 | Genera of Crops Detected in Bali and Singapore | 116 |
| A.10 | Sample Sites, location, sampling frequency | 120 |

ACKNOWLEDGMENTS

I'd like to thank my advisor, Dr. Hope Hollocher, for her guidance and instruction in completing the work of this dissertation. I owe much gratitude to my committee: Drs. Elizabeth Archie, MaryAnn McDowell, and Mike Pfrender. Your perspectives, insight, and expertise guided me throughout the years; this work would be incomplete without your collective input.

I owe a great deal of the work in this dissertation to past and present members of the Hollocher Lab. Justin and Ben, you've been such necessary resources throughout the years and I would be lost without you. Even though we've never worked together, my sincerest gratitude goes to the past graduate students - Amy Klegarth, Lisa Jones-Engel, and Kelly Lane-deGraaf - who built the foundation of macaque work on which I've been able to build. I've been fortunate to work with an amazing group of undergraduate researchers in the lab and I can't wait to see where they go in the future. Thank you and I can't wait to see how you change the world.

People I've worked with at Notre Dame who have helped me along the way are collectively the backbone of the work I've been able to accomplish here. Jackie Lopez and Melissa Stephens deserve everything good for the amount of time they spent at the sequencing core helping my projects go from a notebook sketch to tangible data. Without Jennifer Robichaud, semesters I spent teaching would have been immeasurably less fun and instrumental in developing what pedagogical skills I have gained here. I also want to thank the entire staff and faculty at the Center for Digital Scholarship at Hesburgh, who work tirelessly to help graduate

students learn the computational skills they need to learn to conduct research.

One of the first things I learned when I began graduate school was the importance of social support. For me and many other graduate scholars, the experience of graduate school is peppered with the joy of finding new results and sharing them with the world. These spots shine all the brighter against the backdrop of frustration that comes along with research (not to mention a pandemic). The frustrating bits - stubborn code, confusing results, and failed experiments - are inevitable in graduate school. To come through on the other side, a social network is what gets you through. I consider myself incredibly lucky to have built a group of friends here and I'd like to acknowledge a few individually. Rachel and Ann, I don't think I've ever laughed as hard as I have in your company. To Tyson, your optimism is truly staggering. Chelsea, in addition to being such a great friend, you pointed out once all the loose skin puppies have around their neck and now I always look for that and think of you. Emily, you are a constant inspiration - stay radical. Laura, you quite literally taught me how to survive in the midwest - how could I ever thank you enough? And Mauna...I struggle putting into words how valuable our friendship has been to me through the years. I've learned so much about life, friendship, and the world from you and with you.

To my family, you've stuck by my side and have never flinched in your support for my work and my career - it means more than you'll ever know.

And finally, to Ricardo, my rock. Thank you for listening to my rants, for being my sounding board for ideas (science-related and not), and for sharing your life with me. Here's to the rest of it being filled to the brim with sunshine, lollipops, and rainbows. I love you.

CHAPTER 1

INTRODUCTION

1.1 Overview of Literature

The interior and exterior margins of animal anatomies are connected through millions of microbes inhabiting even the most elusive spaces. Microbial assemblages associated with a host are thus built and shaped by a multitude of sources which originate from both the host and the environment. Host secretions, such as mucus and anti-microbial peptides, exhibit *inside-out* control over epithelial bacteria, leading to the fluctuation or depletion of bacterial phyla (Hooper et al., 2012; Sommer et al., 2014; Tailford et al., 2015). Bacteria exhibit *outside-in* behavior in turn, introducing lipopolysaccharides and flagellins, which promote the expression of the bacteriocidal elements and influence both anti- and pro-inflammatory T-cells in the gut mucosa (Hooper et al., 2012). In addition, host behavior — such as foraging or physical contact — provides nearly constant influxes of environmental microbes that replenish populations and provide competition to existing colonies. The extent to which host-associated microbiomes influence the host have been the target of a rich body of research aiming to explain possible associations with health issues like obesity (Turnbaugh et al., 2006), nervous system function (De Vadder et al., 2018), and gastrointestinal distress (Sartor, 2008).

Influences from diet and parasites on these prokaryotic assemblages is frequently the subject of scrutiny, as these factors contribute greatly to the composition and fluctuation of host-associated microbiomes (Clayton et al., 2016; David

et al., 2014b; Laforest-Lapointe and Arrieta, 2018). Upon close examination, diet is an environmental factor that plays an important role in intestinal microbiome composition, although the extent to which this role manifests varies greatly. For example, diet is a major determinant of gut microbiome structure and function in humans, and has important implications for human health (Rothschild et al., 2018; Singh et al., 2017; Zmora et al., 2019). On the other hand, diets designed to restore 'wild' microbiomes in captive primates did not significantly alter prokaryotic taxa abundances or the microbiome resemblance to wild counterparts (Cabana et al., 2019).

The literature suggests strong potential relationships, although the interactions between prokaryotes and parasites in mammalian hosts have not been well characterized (Avelo and Medlar, 2018; Leung et al., 2018). *Giardia* infections can alter microbiomes to favor the rise of specific bacterial phyla (Šlapeta et al., 2015), soil-transmitted helminths exerted long-lasting changes in human prokaryotic assemblages (Rosa et al., 2018), and presence of the protozoan *Blastocystis* in rural human populations may drive community composition and diversity (Nieves-Ramírez et al., 2018). Infection of *Blastocystis* is often asymptomatic, but this genus consists of subtypes which may vary in degree of pathogenicity (although subtype is not reliably correlated with pathogenicity)(Roberts et al., 2013; Tan et al., 2010). The extent to which prokaryotic organisms prevent or enhance parasitic infections depends on a great deal of factors, which is where the focus of my research lies.

1.2 Variation in Prokaryotic Microbiomes

A multitude of factors influence the composition and fluctuation of mammal microbiomes, such as diet (Muegge et al., 2010), social interactions (Kolodny et al.,

2019; Tung et al., 2015), host immune interactions (Hooper et al., 2012; McKenna et al., 2008), and host genetic determinants (Blekhman et al., 2015; Grieneisen et al., 2021). Environmental factors related to direct exposure account for a large portion of variation occurring within the prokaryotic microbiome (Rothschild et al., 2018), and humans in urban dwellings share microbes with each other and their homes, but tend to deposit microbes into their generally abiotic surroundings more than they receive them (Lax et al., 2014).

Compared to the gut, the oral microbiome has an increased direct exposure to the environment and responds to several different microhabitats within the oral cavity, for example areas under the tongue and between the teeth (Strużycka, 2014; Zaura et al., 2009). Taxa present in the oral microbiomes of humans are altered by oral hygiene practices and influence health through periodontal diseases, such as caries (Mashima et al., 2017), and the functional profiles of the taxa present differ in humans with little exposure to dental hygiene practices (Clemente et al., 2015). In a population of semi-captive rhesus macaques (i.e., not caged, but cared for by humans), several taxa identified from saliva swabs were similar to those collected from the Human Microbiome Project (hum, 2012), supporting the suitability of macaques to model health effects (Chen et al., 2018).

The gut microbiome in humans is affected by reproductive health and status (Lindheim et al., 2017; Mallott and Amato, 2018), and general environmental surroundings, e.g. household sharing (Lax et al., 2014; Menke et al., 2017; Rothschild et al., 2018). Seasonality (Orkin et al., 2019a) and reproductive status (Mallott and Amato, 2018) have shown responses in microbial variation are driven by host energy needs.

Plasticity is key in microbiome research. Environmental influences such as diet and geography are important factors in shaping gut communities (Greene et al., 2018), but these interactions are difficult to pinpoint. Sympatry is associated with

convergence in microbiomes in wild chimps, bonobos, and gorillas (Moeller and Ochman, 2013), although host physiology also has an important role in constraining the variation found in the microbiome, even when dietary niches greatly differ (Amato et al., 2019).

Microbiome community structure is subject to fluctuation due to constant influxes of different nutrients associated with varying diets, their sensitivity to the health condition of the host, and exposure to other infectious organisms in the gut, yet there are numerous examples of host populations that have configurations of taxa appearing more frequently than others. Repeated patterns of changes in the structure of gut communities, termed enterotypes, are theorized to result possibly from shared ecological relationships such as diet types (Arumugam et al., 2011; Moeller et al., 2013). Structuring and categorization of microbiome landscapes is an attractive concept for researchers searching to identify biomarkers and apply them to new populations, but identifying how exactly stratifications are useful in predicting trends has proven challenging (Hosoda et al., 2020; Knights et al., 2014). Standardization in analysis throughout the field and inclusion of expert knowledge are needed to advance application, but enterotype analysis can still be useful in advancing microbiome research (Costea et al., 2017).

In humans, the dominant genera which generally define different enterotypes are *Prevotella*, *Bacteroides* and *Ruminococcaceae*, but the phyla Firmicutes more generally define clusters (Costea et al., 2017; Falony et al., 2016). The ratio of the phyla Firmicutes to Bacteroidetes has been used as a potential indicator of the effect of the environment and diet on the microbiome in humans (De Filippo et al., 2010; Ley et al., 2006; Turnbaugh et al., 2006), but this has not been fully explored in wild animal systems (Amato, 2013). In great apes, the phyla Firmicutes, Bacteroidetes, and Proteobacteria, and Actinobacteria are often highly abundant and enterotype-defining taxa in gut prokaryotic communities, but genera such as

Akkermansia and *Methanobrevibacter* also define in these environments (Falony et al., 2016). Some of these taxa have been proposed to help fill host functional roles, e.g. *Prevotella* in dietary fiber digestion (David et al., 2014b).

Diet especially has a profound effect on the structure of prokaryotic microbiota (David et al., 2014b; Greene et al., 2018; Wu et al., 2011), particularly in wild animals (Amato et al., 2013). Dominant taxa of primates in captivity converge with those of humans, likely as a result of altered diets and shared environment (Clayton et al., 2016). In addition to the microbes which are present on food items, digestion provides changes in carbohydrates and nutrients to act as resources, which shape the microbial community (Tailford et al., 2015). Compounds in food (e.g. thiosulfinates in garlic) can have antimicrobial effects by inhibiting RNA synthesis and nitric oxide production (Dugasani et al., 2010; Feldberg et al., 1988), affecting prokaryotes which may fill a functional role in the gut (Muegge et al., 2010).

1.3 Parasite interactions with the microbiome

Protozoans, helminths, and fungi have been studied in the context of ecological drivers in gut environments (Laforest-Lapointe and Arrieta, 2018). Helminths are known for their immunoregulatory effects, and this relationship has recently been expanded to incorporate the cascading effects of helminths on host immunity and bacterial communities (Montero et al., 2021). In Wilcox and Hollocher (Wilcox and Hollocher, 2018), eukaryotic microbes were categorized into functional guilds, such as grazers, predators, and intracellular protozoa, and followed top-down and bottom-up trophic food web dynamics (Wilcox and Hollocher, 2018). Protozoans associated with a host are often categorized as pathogenic, but some often act as mediators of bacterial communities by keeping bacterial populations from becom-

ing too abundant and causing an imbalance in a host (Mann et al., 2020; Parfrey et al., 2014).

Gut eukaryotes are generally protozoans, heminths, or fungi, and may be parasitic, commensal, or beneficial (Parfrey et al., 2014). These eukaryotic microbes affect the host and the prokaryotic microbes in ways that are not thoroughly understood, but interactions may be helpful in determining the propensity of a parasite to act in a way that is harmful or helpful to the host, by disrupting or maintaining homeostasis. The presence of eukaryotes is associated with changes in the structure of the prokaryotic community, such as richness (Audebert et al., 2016; Barash and Maloney, 2017) and abundance of prokaryotic taxa (Beghini et al., 2017; Nieves-Ramírez et al., 2018). Mucus produced by the host in response to parasitic infections contains glycans, which are consumed by some bacteria in the gut (Li et al., 2012). The overproduction of mucus, which lines the gastrointestinal tract and produces glycans consumed by bacteria, is associated with inflammation and may be a response to the infection of some eukaryotic parasites (Li et al., 2012; Tailford et al., 2015). Other host immune responses affecting bacteria include the release of alpha-defensins and the stimulation of IgA production, which result in phagocytosis of bacteria in the mucus layer (Costello et al., 2011; Hooper et al., 2012).

Members of the prokaryotic community may act against colonization or infection of eukaryotes. The bacteria *Pseudomonas aeruginosa* can alter and control the yeast *Candida albicans* (Hogan et al., 2004; Morales et al., 2013), as can the production of short-chain fatty acids by common carbohydrate metabolizing prokaryotic taxa (Noverr and Huffnagle, 2004). *Candida albicans* is able to control *P. aeruginosa* as well, through signaling and swarming inhibition (Brand et al., 2008; McAlester et al., 2008). The protozoan *Blastocystis* is found widely in wild animals and which may or may not cause illness and contribute to physical symp-

toms in the host (Deng et al., 2021; Nieves-Ramírez et al., 2018). Members of the protozoan genus *Blastocystis* (Stramenopile, phylum Heterokontophyta) are commonly found in the gut of primates and human populations in developing countries can have a prevalence rate of up to 70% (Alfellani et al., 2013). The role of *Blastocystis* in health and disease is not understood and research is conflicting concerning factors affecting pathogenicity. Subtypes of *Blastocystis* have been tested as possible explanation for pathogenicity, but these results are inconsistent, showing other factors are at play (Beghini et al., 2017; Clark et al., 2013; Jiménez et al., 2019; Lepczyńska et al., 2017). Gut microbiomes of *Blastocystis*-colonized individuals are associated higher diversity (Audebert et al., 2016; Tito et al., 2018), which suggests the protozoan's role is that of a potential driver in the gut (Laforest-Lapointe and Arrieta, 2018; Nieves-Ramírez et al., 2018).

1.4 Project Overview

My dissertation work uses two populations of free-ranging long-tailed macaques (*Macaca fascicularis*) to explore host-associated prokaryotic and eukaryotic dynamics. The long-tailed macaque has an estimated population of approximately 3,000,000 in insular and peninsular Southeast Asia, making it the second most successful primate in this region of the world, after humans (Fooden, 2006). The Hollocher lab has been researching and collecting data from these populations since 2006, providing ample history of work on which this dissertation builds. Our work focuses on Singapore and Bali, Indonesia, which are home to populations of long-tailed macaques that are exposed to anthropogenic environments to varying degrees.

Singapore is an island of approximately 720 km² with a population approaching six million people. In 2015, a nationwide census estimated the population of *M.*

fascicularis to be approximately 2000 individuals dispersed among 92 groups (Riley et al., 2015). The history of Singapore is rife with deforestation and subsequent strategies to rebuild and prioritize nature, including nature reserves that are habitat to many macaques along with humans who utilize these spaces for tourism, recreation and living purposes (Corlett, 1992). The highly urbanized habitat surrounding nature reserves creates frequent opportunities for human-macaque interactions – some of which is based in conflict and has led to culling practices (Fuentes et al., 2008; Riley et al., 2015; Sha et al., 2009b,c). Aside from conflict, this proximity to humans impacts the macaques in numerous ways, including foraging behavior, access to water, and dispersal.

To contrast, Bali, Indonesia has a much larger area of 5,780 km², a human population of approximately four million, and population estimates of macaques place their number 10-12,000, most of which are associated with religious temples (Wheatley, 1999). Macaque populations which are commonly associated with temples throughout the landscape contribute to the tourist industry, in which provisioning can be part of the religious experience and leads to high degrees of interaction between macaques and humans (Brotcorne, 2014; Fuentes, 2011). Conflict is also present in Bali, stemming from interactions such as crop-raiding, hunting, and trade (Lane-deGraaf et al., 2014b; Loudon et al., 2006).

Past work in this study system has aimed to derive and elucidate aspects of how the anthropogenic environment is reflected in macaques through analyzing dispersal, genetic structure, and parasite distribution. Movement detected with the use of GPS collars reflect higher variability in ranging patterns in macaques heavily influenced by provisioning in urban environments (Klegarth et al., 2017). Genetic structure is detected in Bali through microsatellite heterozygosity and genetic distance and reflects dispersion patterns affected by human-macaque interactions and the landscape (Lane-deGraaf et al., 2014b). In both Singapore and

Bali, macaques experience population structure changes that can be detected in mitochondrial DNA haplotypes following a “continental” and “insular” divide (Kle-garth et al., 2017).

Anthropogenic influence is also reflected in Bali through the lens of parasites. Human behavior presents opportunities for parasite transmission, and sites characterized as more anthropogenic mitigate parasite prevalence and intensity (Lane et al., 2011; Lane-deGraaf et al., 2014b). Looking further into community-level parasite interactions in macaques revealed multiple infections of helminths were common and explained variation in microparasite shedding (Wilcox et al., 2014). The lab’s first use of barcoding amplicon sequences revealed eukaryotic diversity previously unseen in macaques, which is governed by trophic web interactions between eukaryotic symbionts (Wilcox and Hollocher, 2018). The role of prokaryotic impact on this system is a missing piece in macaque host-microbe interactions, and my dissertation research is addressed at elucidating this role. My work is first aimed at uncovering the prokaryotic interactions occurring within the macaque, and then incorporates diet and parasite information to take a comprehensive approach in viewing how the bacterial microbiome reflects shifts from these factors.

In Chapter 2, I compare and contrast the oral and gut prokaryotic communities across sampling sites to ask how these populations differ and shift through geospatial distance in the highly urban environment of Singapore. I find that oral communities consistently exhibit higher diversity within and similarity between sampling in comparison to gut communities, suggesting greater host control over gut microbiomes and a stronger environmental influence on the oral microbiome. I also show that fecal microbiomes have greater dispersion than oral microbiomes. Community composition differs through phylum-level changes, and I explore taxa distinguishing these environments at the family level. I then incorporate K-means clustering methods to validate these results and find that samples are recovered

entirely according to body site and that correlating taxa match distinguishing taxa identified through differential abundance. Collectively, I show how differences in the oral and gut communities reflect environmental influence and host control in ways which help us to understand their roles in host-associated microbial ecology.

For Chapter 3, I aim to find dietary changes driving microbial variation in the gut prokaryotic communities of Bali and Singapore macaques. I employ K-means clustering and bring in data from a study of Japanese macaques to detect patterns of bacterial variation. I then incorporate dietary information obtained through eukaryote-targeted amplicon sequencing to investigate this variation and find differences in dietary richness and crop consumption between these clusters. Prokaryotic taxa correlating to clusters reveals further information relating to dietary function and leads to further insight of how host gut microbiomes might aid host adaptability to environmental differences.

Chapter 4 strays from prokaryotic focus to take a deeper look into eukaryotic interactions. Here I evaluate macaque consumption of medicinal plants in comparison to parasites present to investigate a potential of mitigative effects. This work uses a barcoding approach to build upon hypotheses of self-medication derived from field observations of wild primates. Correlations reveal connections between medicinal plants and protozoans driven by the genus *Plasmodium*. I further analyzed these data and found that abundance and richness of medicinal plants is higher when *Plasmodium* is higher, and these effects are stronger when plants are filtered to those with the most medicinal properties. These relationships provide support for self-medication in macaques and emphasize the importance of maintaining biodiversity from host health perspectives.

Finally, Chapter 5 brings us back to the prokaryotic microbiome to view variation through the lens of gut protozoans. I use eukaryotic and prokaryotic sequencing data to construct principal components of analysis and find *Blastocystis* to drive

variation within gut protozoans. I then use the presence and abundance of *Blastocystis* to detect differences in prokaryotic ecology metrics and find *Blastocystis* is associated with an increase in alpha diversity and a explains significant variation in community composition. Differential abundances of bacterial taxa reveal nine taxa are more abundant in the presence of *Blastocystis*, which are similar to those identified in studies using data collected from humans. This work builds upon existing evidence to explore the role of *Blastocystis* as a driver of host gut ecology.

Collectively, my dissertation expands knowledge of primate-associated microbial ecology by incorporating new information from macaques in Bali and Singapore. A large portion of my work focuses on incorporating the prokaryotic microbiome of macaques into the greater genetic, and anthropogenic, and parasite-related context known in this system, thereby increasing the scope of future lab work. Microbiome research is dominated by work focusing exclusively on prokaryotes, which has revealed a wealth of information in explaining host-microbial relationships, yet is incomplete without understanding how these prokaryotic assemblages interact with other symbionts. Including symbiotic eukaryotes and parasites is essential in understanding the underlying dynamics within the microbiome which holds meaning for host health and disease transmission.

CHAPTER 2

ENVIRONMENTAL AND HOST EFFECTS ON ORAL AND GUT BACTERIAL COMMUNITIES IN FREE-RANGING LONG-TAILED MACAQUES IN SINGAPORE

2.1 Abstract

A central goal of host-associated microbiome research is to determine the level of influence of the host and the environment in community composition. Research into these relationships has advanced greatly in recent years, but often is performed in controlled settings, reducing or eliminating exposure to the environmental variables and microbes, which may be instrumental in bacteria community colonization and assemblage dynamics. Host phylogeny, physiology, and behavior have all explained some variation in inter- and intra-species microbiota composition and community metrics, but it is unclear how all these factors combine to influence microbial dynamics across spatial distance. To investigate this, we used 16S rRNA amplicon sequencing to analyze the oral and gut microbiomes in populations of long-tailed macaques distributed throughout Singapore to further our understanding of influences of environment and host biology on host-associated microbial community structure. Oral and gut microbiomes showed distinct differences in their overall composition and community properties. Oral microbiomes exhibited a greater number of distinctive taxa, higher diversity and evenness, and lower dissimilarity than gut microbiomes across sampling locations. Samples with a higher Firmicutes:Bacteroidetes (F/B) ratio were more common in oral than in gut

samples, and families within Firmicutes fluctuated less in gut communities. Clustering analyses showed clear differentiation between oral and gut environments, with greater variability exhibited across families in Bacteroidetes, when comparing saliva and fecal samples. Oral and gut microbiomes in free-ranging macaques reflect the influences of the environment and host biology. The higher richness seen in oral samples across the island may reflect greater exposure to the environment, whereas the oral microbiome shows greater evenness and similarity than fecal samples across geographical range, suggesting individual host differences play a greater role in the gut. Families within the Firmicutes phylum tend to fluctuate less in gut samples with a high F/B ratio, contributing to the greater overall similarity between samples that have higher Firmicutes abundances. These findings contribute to our understanding of how host biology and environmental variation interact to influence the relationship between free-ranging primates and their associated prokaryotic communities, particularly with respect to how host body site microbiomes are maintained.

2.2 Background

Animal microbiomes are composed of millions of microbes engaged in complex dynamics driven by a multitude of factors. The influences these interactions exact on the host have been the target of a rich body of medical research focused on humans, demonstrating associations between microbiomes and observable host traits, such as obesity (Turnbaugh et al., 2006), nervous system function (De Vadder et al., 2018), and gastrointestinal distress (Sartor, 2008). Differences in the composition and diversity of bacterial communities become more nuanced when sampling between different body sites. In humans, for example, the gut microbiome is influenced by diet on long- and short-term bases and environmental sur-

roundings (David et al., 2014a; Menke et al., 2017; Rothschild et al., 2018; Wu et al., 2011). In comparison, the oral microbiome often differs in composition, but is subject to similar influences as the gut, with the addition of the increased direct exposure to the environment and the complexity of microhabitats within the oral cavity, such as areas under the tongue and between the teeth (Dewhirst et al., 2010). Diversity, composition, and functional profiles of taxa present in the oral microbiomes of humans vary with oral hygiene practices and influence health through periodontal diseases, such as dental caries (Clemente et al., 2015; Mashima et al., 2017). Evidence increasingly shows the oral microbiome is convergent across primate species (Asangba et al., 2022; Sawaswong et al., 2021), showing the plasticity exhibited by hosts in adapting to niche environments.

In nonhuman primates, seasonality (Orkin et al., 2019b), reproductive status (Mallott and Amato, 2018), and diet (Greene et al., 2018; Moeller et al., 2012) can have reproducible and quantifiable effects on primate microbial composition, as can the presence of gut parasites that may act as drivers in the gut microbial community such as *Blastocystis* (Laforest-Lapointe and Arrieta, 2018; Nieves-Ramírez et al., 2018). Anthropogenic influences can impact microbial structure as well, as seen in studies examining captivity, in which the dominant taxa of the microbiome in nonhuman primates converge with those of humans, possibly as a result of shifts in diets provided in captivity as compared to wild habitats (Clayton et al., 2016). Phylogeny, heredity, and physiology have effects that contribute to the variation imparted by diet and geography in animals, both within and across species (Grieneisen et al., 2021; McCord et al., 2014; Yildirim et al., 2010).

Examining changes in microbial taxa at the phylum level allows characterizations of communities at a broad taxonomic scale in host populations, and studies framing phylum-dominated communities as enterotypes have been used to tie microbiome structure to health and diet in hosts (Costea et al., 2017; Wu et al.,

2011). Focusing more closely on intra-phylum associations, such as the ratio of Firmicutes to Bacteroidetes (F/B ratio), can provide deeper understanding of these systems and their association with shifts in specific host traits, such as body mass index and age (Koliada et al., 2017; Ley et al., 2006; Mariat et al., 2009). The F/B ratio of host-associated microbial communities has also been useful in defining possible functional profiles and microbial community-specific properties, although how these community structures are colonized and maintained is less clear (Costea et al., 2017; Wu et al., 2011), particularly in free-ranging nonhuman primates. Enterotypes can be detected in chimpanzee populations in Gombe National Park, for example, but the factors involved in their formation remain elusive (Moeller et al., 2012).

Investigating the microbial dynamics among taxa is important in understanding their associations with nonhuman primate hosts, in particular those with increased contact with humans and therefore vulnerable to zoonotic pathogen transmission (Gardner and Luciw, 2008; Sha et al., 2009b). Here, we describe the community structure and dominant taxa of the oral and gut prokaryotic microbiomes of free-ranging populations of long-tailed macaques (*Macaca fascicularis*), an edge species well-known for its success in anthropogenic environments (Gumert, 2011). We specifically focus on macaques within the island of Singapore, an environment ripe with human-macaque interactions (Fuentes et al., 2008). Using the V4 region of the 16S rRNA gene, we assess the overall bacterial composition of body sites, identify which taxa define body sites, and measure differences in community structure (e.g., richness, evenness, dispersion, and dissimilarity), and then evaluate how those results reflect the effects of the environment and the biology of the hosts on these populations. We inspect these communities at the phylum, family, and amplicon sequence variant (ASV) level to identify possible drivers underlying the changes in structure we see at higher taxonomic levels. Targeting these in-

teractions in a highly-anthropogenic population of macaques allows an approach that can highlight host plasticity and adaptability to altered environments, effects essential for understanding the role of the microbiome in host fitness.

2.3 Methods

2.3.1 Study Site, Sampling, and DNA extraction

Saliva and fecal samples (N=92, 46 saliva and 46 fecal) were collected from free-ranging long-tailed macaques across ten populations throughout Singapore (Figure 2.1) and were sequenced for the 16S rRNA genes. Fecal (five sites, N=46) and saliva (nine sites, N=46) samples were collected from Singapore in June 2013 – October 2013. Collection times fall within Southwest Monsoon season, marked by frequent rain (Singapore). Fresh fecal samples were recovered from the ground when defecation was observed, whereas saliva was collected using sugar-syrup flavored buccal swabs from Salimetrics, offered to the macaques and recovered immediately after discard when only a single individual had handled it. Swabs were kept on ice packs during collection and stored in 1% sucrose cell lysis buffer and at -80°C. Total DNA was extracted from feces using the Qiagen QIAMP Stool Minikit (Qiagen, Hilden, Germany) as per manufacturer instructions (Qiagen, 2012). DNA from saliva samples was extracted via a phenol chloroform protocol: samples were incubated overnight at 55°C in 900.00 μ L CTAB with 20.0 μ Proteinase K enzyme. Samples were shaken with phenol/chloroform for one minute, and then spun to separate the aqueous phase that was then washed with 100% EtOH, followed by washing with 70% EtOH solution, and stored at -80°C. Further details regarding sample collection and DNA extraction are described previously (Klegarth et al., 2017; Wilcox et al., 2014).

2.3.2 Target Amplification and High-Throughput Sequencing

A 287 bp sequence of the V4 region of the 16S rRNA gene was amplified from genomic DNA using the following primers (Illumina adaptors in italics): Forward (S-D-Arch-0519-a-S-15, 5' - *TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CAG CMG CCG CGG TAA* -3'), Reverse (S-D-Bact-0785-b-A-18, 5' - *GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GTA CNV GGG TAT CTA ATC C* -3') (Klindworth et al., 2013; Van Bleijswijk et al., 2015) in the following PCR conditions: for a 25.0 μ L reaction volume, 10.00 μ L ddH₂O, 12.50 μ L Kappa HiFi Hot Start Taq Ready-Mix (Kappa Biosystems, Wilmington, MA, USA), 0.75 μ L 10 uM forward and reverse primer, 1.00 μ L template DNA. Thermocycler conditions were: 3:00 initial denaturation at 95.0 °C, 30 cycles (0:20 denaturation at 95 °C, 0:15 anneal at 65 °C, 0:15 extension at 72 °C), and a final extension time of 3:00 at 72 °C. Included in the first and last wells of the 96-well plate were negative controls, containing only autoclaved water in the place of the template DNA. Next to these negative controls were reactions containing genomic DNA from a prepared mock community (ZymoBIOMICS™ Microbial Community DNA Standard, Catalog no: D6306) as template, included to identify potential bias from PCR and sequencing assays and improve reproducibility in our work.

Illumina PCR and sequencing PCR products were purified with Ampure XP Beads (Beckman Coulter, Indianapolis, IN, USA), and size distributions were taken from each library using a Bioanalyzer DNA 7500 chip (Agilent, Santa Clara, CA, USA). Size distributions were unique to each library, and no major anomalies were seen. Index PCRs, were performed on all libraries following clean-up and quality assessment using a 95 °C denaturing phase, 55 °C annealing phase, and 72 °C extension phase, for eight cycles, each lasting 30 seconds. Libraries were pooled and normalized prior to index amplification. Each library was used to create 100 μ M



Figure 2.1. Sampling sites indicate populations from which samples were obtained. Samples were collected throughout accessible macaque habitat. Geographic coordinates and number of samples for each site are as follows: AP - Admiralty Park: 1° 27' 2.9844" N, 103° 46' 38.9964" E (N = 12); BB - Bukit Batok: 1° 22' 36.2900" N, 103° 45' 30.6242" E (N = 3); BSI - Big Sister's Island: 1° 12' 52.4916" N, 103° 50' 6.3055" E (N = 8); BT - Bukit Timah: 1° 22' 47.0205" N, 103° 46' 25.42636" E (N = 11); MR - MacRitchie Reserve 1° 20' 28.0572" N, 103° 49' 48.5616" E (N = 17); OUT - Old Upper Thomson Road: 1° 22' 49.7233" N, 103° 49' 6.3001" E (N = 15); SI - Sentosa Island: 1° 15' 29.3885" N, 103° 48' 33.5866" E (N = 4); UP - Upper Pierce: 1° 22' 32.2896" N, 103° 48' 17.3988" E (N = 6); US - Upper Seletar: 1° 23' 55.5432" N, 103° 48' 23.2812 E; WW - Woodlands Waterfront: 1° 27' 6.6652" N, 103° 46' 52.2617" E (N = 11).

aliquots based on the following equation:

$$\text{Concentration in nM} = \left(\frac{\text{concentration in ng}/\mu\text{L}}{600 \text{ g/mol} \times \text{average library size in bp}} \right) \times 10^6$$

A normalized library was prepared by diluting 5.0 μL from the prepared library to 20.0 nM aliquots. Size distributions were bioanalyzed using a DNA 7500 chip, to confirm appropriate size distributions within the library. The library was sequenced on an Illumina HiSeq 2500 (Illumina, San Diego, CA, USA), using a rapid run with paired-end 250 bp reads at the University of New Hampshire Hubbard Center for Genome Studies.

2.3.3 Bioinformatic Workflow and Statistical Analysis

The R package Dada2 was used for processing sequences, including filtering, trimming, and merging, and removal of chimeras (Callahan et al., 2016). Taxonomy assignment was performed using the Dada2 function of the RDP classifier and the RDP training set (release 11.5), and all eukaryotes, archaea, and chloroplasts were removed (Callahan et al., 2016). We aligned the contents of mock community samples to a reference database of 16S ssuRNA sequences downloaded from the manufacturers website (version 2, accessed 09/29/2017) using HISAT2 (v. 2.1.0) (Kim et al., 2015); each aligned at greater than 98%, and all ASVs were retained hereafter, unless otherwise noted. Rarefaction curves were created with the rarecurve function from the vegan package (v 2.5-7)(Oksanen et al., 2008).

To measure alpha diversity, richness and Pielou's evenness were calculated in R using the vegan package (Oksanen et al., 2022). Bray-Curtis dissimilarity was measured between each sample's relative abundance and the median relative abundance for the corresponding sampling location. Dissimilarities of all samples

were then grouped by body site. To account for uneven sampling, we used Mann-Whitney U-tests to determine differences in richness, evenness, and dissimilarity between body sites. Taxa were sorted and levels combined using the R (v 4.0.5) package Phyloseq (v 1.34.0) (McMurdie and Holmes, 2013). We identified differentially abundant taxa between oral and gut samples using linear discriminant analyses (LDA) using the program LefSe (v 1.0) (Segata et al., 2011).

To further inspect taxa defining the oral and gut environments, we employed a K-means clustering method, and identified the taxa correlated with each cluster. The ASV table was transformed using the centered log-ratio method, which can decrease bias in analyzing compositional data (Gloor et al., 2017). The centered log-ratio transformed samples were clustered using a K-means algorithm. A K-value of two was selected in an attempt to show that the oral and gut bacterial communities are distinctive. If these two clusters match the sample type it would support this distinctiveness. The adjusted Rand index was used to measure agreement between the clusters and sample type. (Hubert and Arabie, 1985). Aitchison (i.e., Euclidean) distances between samples were calculated using the `vegdist` function in the `vegan` package (v 2.5-7) (Oksanen et al., 2008) in R. The `envfit` function in the `vegan` package was used to determine which bacterial taxa were strongly correlated ($R^2 > 0.7$) with which clusters.

NMDS plots were created to visualize grouping in samples in R using the `vegan` package, using Bray-Curtis distances (Oksanen et al., 2008). PERMANOVAs (`adonis2` function from the `vegan` package in R) tested for differences in unfiltered, relative abundance transformed composition between body sites at phylum and family levels, controlling for sampling location and sample read counts as well as differences between sampling sites, and controlling for body sites and read counts. Additionally, PERMANOVAs were used to test for differences in composition among sampling types for each body site at phylum and family levels. We

measured the dispersion of variance between body sites, and then between sampling locations within body site samples using the `permdisp` and `permutest` (permutations = 9999) functions in the `vegan` package. Changes in abundances of families within the phyla Firmicutes and Bacteroidetes, were tested using a series of Mann-Whitney U tests and plotted. Families composing less than one percent of their respective phylum were excluded to avoid emphasis on changes in low-abundance families. MetaCyC functional pathways were predicted using the software PICRUSt2 (Barbera et al., 2019; Caspi et al., 2014; Czech et al., 2020; Douglas et al., 2020; Louca and Doebeli, 2018; Mirarab et al., 2011; Ye and Doak, 2009). The resulting pathway table was then analyzed in R using ALDEx2, which uses a Dirichlet-multinomial model to infer abundance from counts, to determine which pathways were significantly associated with body sites (v 1.22.0) (Fernandes et al., 2013, 2014; Gloor et al., 2016a). As recommended by the authors of this package, we used an effect size of > 1.0 to determine significance of pathways associated with body sites.

2.4 Results

A total of 71,460,186 paired-end reads were obtained for processing; after filtering and removal of chimeras, eukaryotes, and chloroplasts, 52,192,561 reads remained, resulting in a total of 13,699 bacterial ASVs (rarefaction curves in Figure A.1 show all samples plateau at the read depth of our chosen sequencing platform).

We measured more overall diversity, yet less dissimilarity, across individuals in the oral microbial community than the gut. When comparing family richness (Figure 2.2A), dissimilarity (Figure 2.2B), and evenness (Figure 2.2C) between gut and oral microbiomes, the oral microbiome contained significantly higher family

richness ($W = 243.5$; $p < 0.001$) and significantly lower family dissimilarity ($W = 657$; $p = 0.002$). As we could only directly compare overlapping samples for both oral and gut microbiomes, these were the only ones directly tested. Metrics for all sites are displayed in Table A.1.

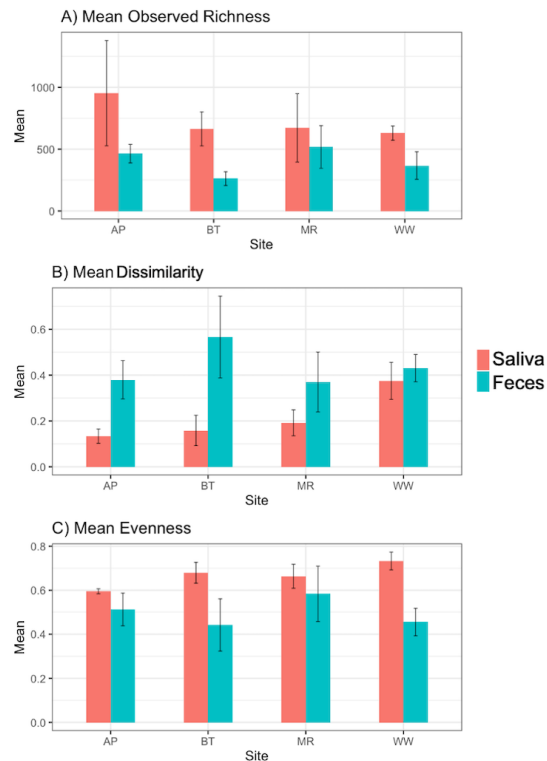


Figure 2.2. Family richness (A), dissimilarity (B), and evenness (C) of sites containing both saliva and fecal sampling - Admiralty Park (AP), Bukit Timah (BT), MacRitchie Nature Reserve (MR), and Woodlands Waterfront (WW). Saliva samples display greater richness, similarity, and evenness throughout sites.

In oral and gut microbiomes, the dominating phyla are Proteobacteria, Firmicutes, Bacteroidetes, and Actinobacteria, with 28 total phyla represented across

these two body sites (Figure 2.3). A PERMANOVA test of Bray-Curtis distances between body sites revealed significant differences in distances in both phylum and family taxa (phylum: $R^2 = 0.065$, $p < 0.005$, family: $R^2 = 0.198$, $p < 0.001$ (Table 2.1), indicating that the compositions of our samples are significantly different between these internal host environments.

TABLE 2.1
BODY SITE AND SAMPLING LOCATION FACTORS INCLUDED IN
PERMANOVA PHYLA AND FAMILIES OF ASVS. SIGNIFICANT
VARIABLES AND VALUES IN BOLD FOR CLARITY

| | | <i>Df</i> | <i>SS</i> | <i>R</i> ² | <i>F</i> | <i>p</i> |
|----------------------|--------------------------|-----------|--------------|-----------------------|--------------|--------------|
| <i>Phylum</i> | | | | | | |
| | Body Site | 1 | 0.543 | 0.065 | 6.133 | 0.004 |
| | Sampling Location | 9 | 0.58 | 0.069 | 0.728 | 0.783 |
| | Readcount | 1 | 0.191 | 0.023 | 2.158 | 0.101 |
| | Residual | 80 | 7.079 | 0.844 | | |
| | Total | 91 | 8.392 | 1 | | |
| <i>Oral</i> | Sampling Location | 8 | 0.338 | 0.222 | 1.379 | 0.126 |
| | Readcount | 1 | 0.08 | 0.053 | 2.618 | 0.059 |
| | Residual | 36 | 1.104 | 0.725 | | |
| | Total | 45 | 1.523 | 1 | | |
| <i>Gut</i> | Sampling Location | 4 | 1.299 | 0.271 | 3.864 | 0.004 |
| | Readcount | 1 | 0.131 | 0.027 | 1.559 | 0.211 |
| | Residual | 40 | 3.362 | 0.702 | | |
| | Total | 45 | 4.792 | 1 | | |

TABLE 2.1 (CONTINUED)

Table 2.1 continued from previous page

| | | <i>Df</i> | <i>SS</i> | <i>R</i> ² | <i>F</i> | <i>p</i> |
|----------------------|--------------------------|-----------|--------------|-----------------------|---------------|------------------|
| <i>Family</i> | Body Site | 1 | 5.639 | 0.198 | 21.851 | <0.001 |
| | Sampling Location | 9 | 1.723 | 0.06 | 0.742 | 0.852 |
| | Readcount | 1 | 0.502 | 0.018 | 1.946 | 0.094 |
| | Residual | 80 | 20.646 | 0.724 | | |
| | Total | 91 | 28.51 | 1 | | |
| <i>Oral</i> | Sampling Location | 8 | 0.474 | 0.24 | 1.528 | 0.073 |
| | Readcount | 1 | 0.103 | 0.052 | 2.663 | 0.053 |
| | Residual | 36 | 1.395 | 0.707 | | |
| | Total | 45 | 1.971 | 1 | | |
| <i>Gut</i> | Sampling Location | 4 | 1.241 | 0.262 | 3.676 | 0.005 |
| | Readcount | 1 | 0.118 | 0.025 | 1.398 | 0.242 |
| | Residual | 40 | 3.377 | 0.713 | | |
| | Total | 45 | 4.737 | 1 | | |

Combined oral and gut samples did not show significantly distinct variation across sampling locations; however, when viewing oral samples and gut samples separately these differences became more pronounced. Sampling site location was a significant source of variation within gut samples (PERMANOVA, phylum: $R^2 = 0.271$, $p < 0.005$, family: $R^2 = 0.262$, $p = 0.005$), but oral samples did not show significant variation across geographic sites. Dispersion of the oral versus the gut microbiome was significantly different ($p = 0.0001$), with greater variance among gut samples. In contrast, dispersion was not different between sampling sites in either oral ($p = 0.2$) or gut ($p=0.5$) samples.

LDA recovered 57 family-level taxa with an LDA score ± 3.0 following internal

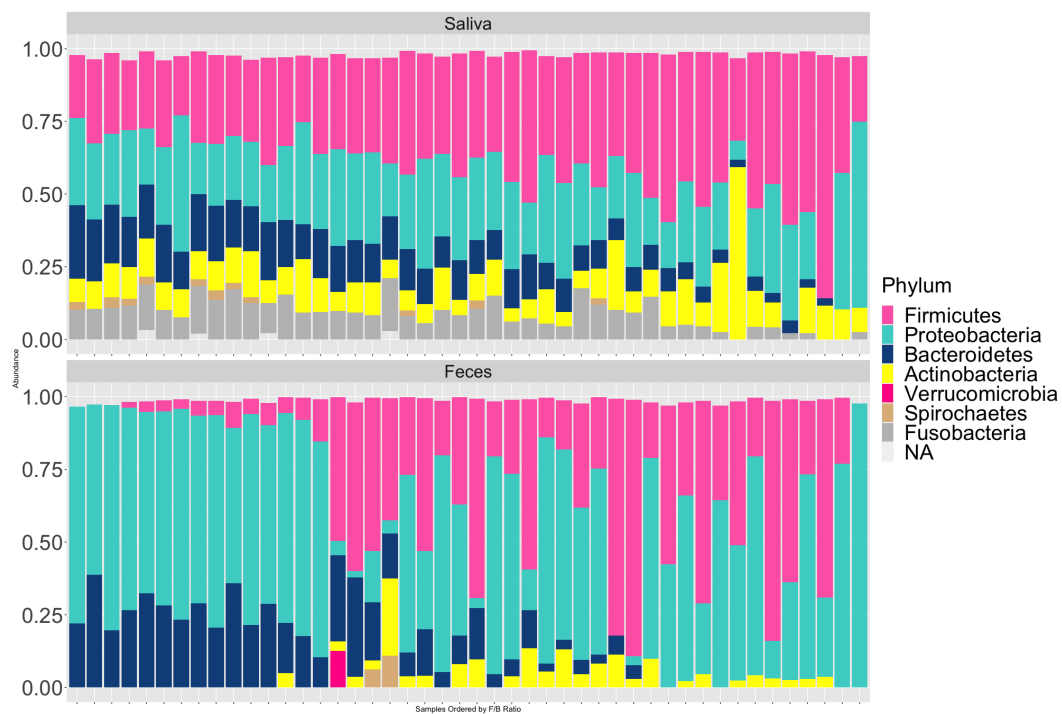


Figure 2.3. Relative abundance of bacterial phyla in oral and gut microbiomes. Samples are ordered left to right by decreasing F/B ratio.

Wilcoxon rank-sum tests ($\alpha = 0.05$) (Figure 2.4). Nine phyla were represented in 29 families defining the oral microbiome, while 28 families in six phyla were defining features of the gut microbiome. Top defining families in oral samples include Streptococcaceae, Pasteurellaceae, and Viellonellaceae, while Enterobacteraceae, Moraxellaceae, and Planococcaceae were more abundant in gut samples.

Visualizing the compositional distances between oral and gut samples showed distinct groups in NMDS ordination of ASVs (Figure 2.5). Saliva samples form a more compact grouping when compared to the fecal samples, suggesting less variation among these communities across individuals. K-means clusters overlapped completely with body sites (adjusted rand index value = 1.0). Fifty-one families strongly correlated ($R^2 > 0.7$) with these clusters and reinforced the body site-defining taxa found through in LDA results. Results of these analyses differed with respect to one family, Burkholderiaceae, which LDA associated with gut, but K-means clustering correlated with oral.

Oral microbiomes are overall more Firmicutes-rich (i.e., Firmicutes taxa are more abundant than Bacteroidetes taxa)(Figure 2.6A). We identified families in either Bacteroidetes-rich or Firmicutes-rich communities, and tested their variation in abundance within these categories (Figure 2.6, Table 2.2). We found four families within the phylum Firmicutes that were more abundant within a Firmicutes-rich environment (Ruminococcaceae, Planococcaceae, Lachnospiraceae, and Clostridia_1), yet no Firmicutes families were measured to be more abundant in a Bacteroidetes-rich environment. Conversely, the Bacteroidetes family Prevotellaceae was more abundant in a Firmicutes-rich environment, whereas the Bacteroidetes families Sphingobacteriaceae and Flavobacteriaceae were more abundant in a Bacteroidetes-rich environment. We used PICRUST2 to predict MetaCyC pathways from ASVs in our samples and then used ALDEx2 to find the effect sizes of these pathways

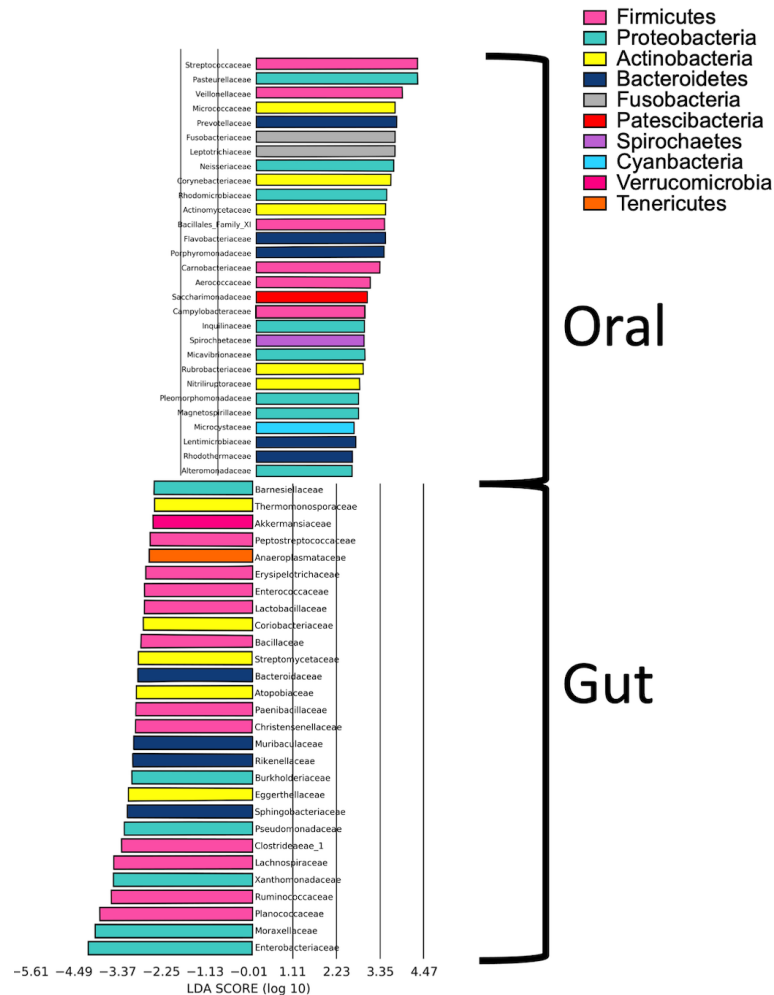


Figure 2.4. LEfSe analysis of family level taxa in oral and gut microbiome. Taxa shown scored at least 2.0 in Linear Discriminant Analysis (57 total). Bars are colored by phylum corresponding to Figure 2.3. For taxa classified as undefined or uncultured at the family level, class name is provided.

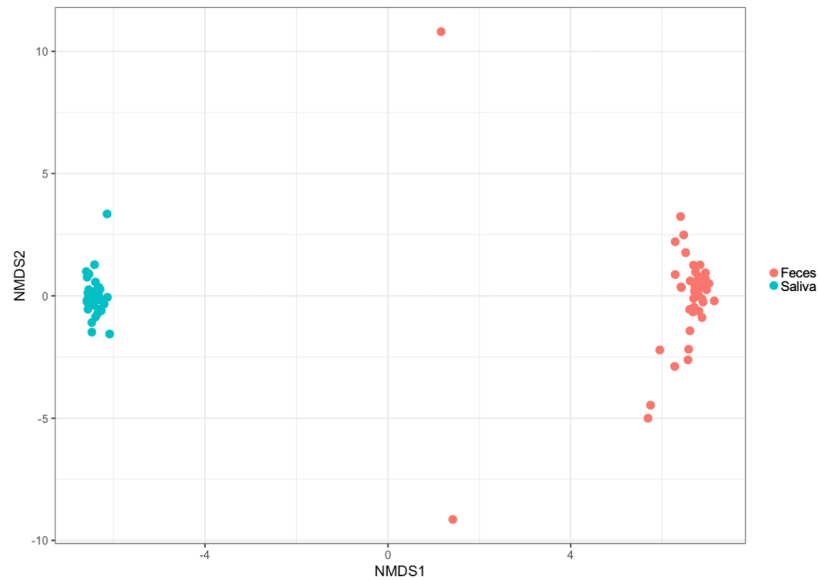
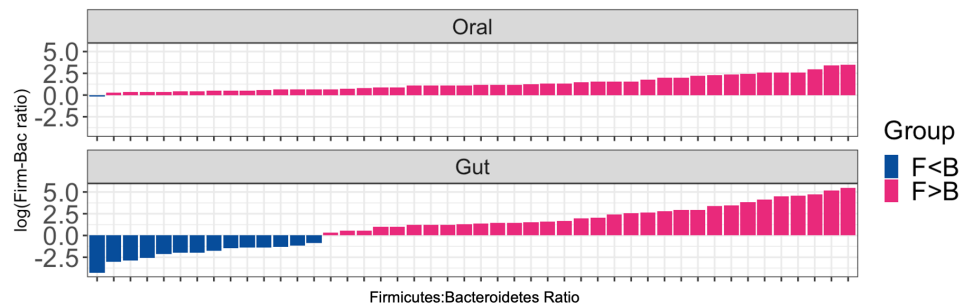


Figure 2.5. Non-metric Multidimensional Scaling of ASVs colored by body site (stress = 0.07). Saliva samples group together more tightly, while fecal samples show greater variation.

between body sites and Firmicute-s or Bacteroidetes-rich samples. For body sites, 24 pathways had an effect size > 1.0 (17 oral, 7 gut). These pathways and their descriptions are presented in Table 2.3. For Firmicutes- or Bacteroidetes-rich samples (Table 2.4), ten pathways had an effect size greater than 1.0 (five in Firmicutes-rich samples, five in Bacteroidetes-rich samples).

A) Body Site Changes of Firmicutes:Bacteroidetes ratio



B) Significant Family Changes within Firmicutes:Bacteroidetes Ratio in the Gut

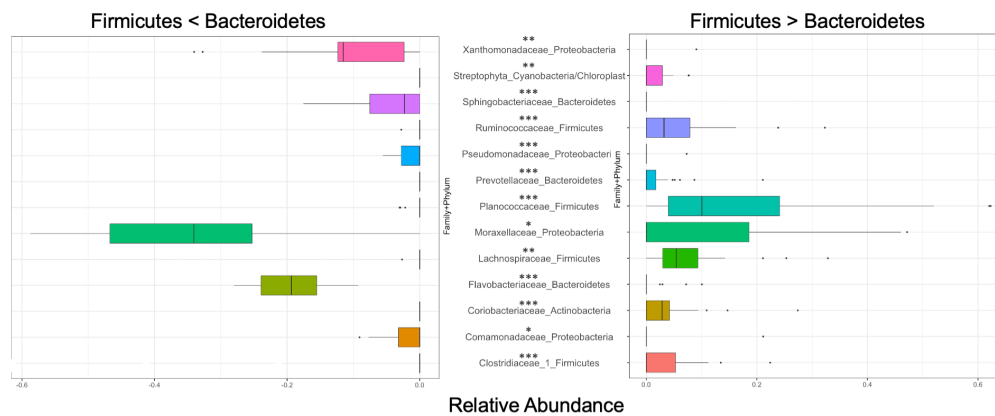


Figure 2.6. Distribution of F/B ratio in oral and gut microbiomes (A). Oral and gut samples ordered by log(F/B ratio). Families which showed significant differences between Bacteroidetes-rich and Firmicutes-rich bacterial communities in the gut environment are shown in (B). Taxa are described with family and phylum-level. Specific p-values for each taxon listed in Table 2.2

TABLE 2.2

FAMILIES WITH SIGNIFICANT CHANGES BETWEEN
BACTEROIDETES-RICH AND FIRMICUTES-RICH GUT
MICROBIOMES. SIGNIFICANCE TESTED WITH MANN-WHITNEY U
TESTS, α OF 0.05

| Phylum | Family | p |
|----------------|---------------------|-----------|
| Cyanobacteria | Streptophyta | 1.289e-02 |
| Bacteroidetes | Prevotellaceae | 4.121e-02 |
| Proteobacteria | Comamonadaceae | 5.099e-03 |
| Proteobacteria | Pseudomonadaceae | 5.099e-03 |
| Firmicutes | Clostridiaceae_1 | 8.533e-03 |
| Proteobacteria | Moraxellaceae | 1.612e-04 |
| Firmicutes | Ruminococcaceae | 2.648e-04 |
| Actinobacteria | Coriobacteriaceae | 2.998e-04 |
| Bacteroidetes | Sphingobacteriaceae | 2.701e-05 |
| Firmicutes | Lachnospiraceae | 1.877e-06 |
| Firmicutes | Planococcaceae | 2.542e-06 |
| Proteobacteria | Xanthomonadaceae | 1.856e-07 |
| Bacteroidetes | Flavobacteriaceae | 2.427e-09 |

TABLE 2.3

PREDICTED METACYC PATHWAYS WITH AN EFFECT SIZE OF > 1.0
 AS ESTIMATED BY ALDEX2 – COMPARISON FOR ORAL OR GUT
 SAMPLES.

| Pathway | Description | Associated Body Site | Effect | Overlap |
|----------|---|-------------------------|--------|---------|
| P341-PWY | Glycolysis - Pyrococcus | Gut | 1.0094 | 0.0808 |
| CODH-PWY | Reductive Acetyl Coenzyme A Pathway I (Homoacetogenic Bac- teria) | Gut | 1.0271 | 0.1185 |
| PWY-7373 | Superpathway Of Demethylmenaquinol-6 Biosyn- thesis II, | Gut | 1.1026 | 0.0822 |
| PWY-6143 | CMP-Pseudamine Biosynthesis | Gut | 1.1136 | 0.0781 |
| PWY-7031 | Protein N-Glycosylation (Bacte- rial) | Gut | 1.2783 | 0.0503 |
| PWY-7616 | Methanol Oxidation To Carbon Dioxide | Gut | 1.4764 | 0.0608 |

TABLE 2.3 (CONTINUED)

| Pathway | Description | Associated Body Site | Effect | Overlap |
|---------------------|--|----------------------|--------|---------|
| PWY-6397 | Mycolyl-Arabinogalactan-Peptidoglycan Complex Biosynthesis | Gut | 1.7127 | 0.0085 |
| DENITRIFICATION-PWY | Nitrate Reduction I (Denitrification) | Oral | 1.0112 | 0.1497 |
| PWY-5507 | Adenosylcobalamin Biosynthesis I (Anaerobic) | Oral | 1.1098 | 0.1107 |
| PWY-922 | Mevalonate Pathway I (Eukaryotes And Bacteria) | Oral | 1.1400 | 0.1247 |
| PWY-5910 | Superpathway Of Geranylgeranyl-diphosphate Biosynthesis I (Via Mevalonate) | Oral | 1.1428 | 0.1182 |
| PWY-7377 | Cob(II)Yrinate Biosynthesis I (Early Cobalt Insertion) | Oral | 1.1538 | 0.1230 |
| PWY-7376 | Cob(II)Yrinate Biosynthesis II (Late Cobalt Incorporation) | Oral | 1.2631 | 0.0849 |

TABLE 2.3 (CONTINUED)

| Pathway | Description | Associated Body Site | Effect | Overlap |
|----------|--|----------------------|--------|---------|
| PWY490-3 | Nitrate Reduction VI (Assimilatory) | Oral | 1.4550 | 0.0764 |
| PWY-5088 | L-Glutamate Degradation VIII (To Propanoate) | Oral | 1.4952 | 0.0591 |
| PWY-5677 | Succinate Fermentation To Butanoate | Oral | 1.6184 | 0.0411 |
| P381-PWY | Adenosylcobalamin Biosynthesis II (Aerobic) | Oral | 1.8604 | 0.0285 |
| PWY-5005 | Biotin Biosynthesis II | Oral | 2.1418 | 0.0221 |
| P162-PWY | L-Glutamate Degradation V (Via Hydroxyglutarate) | Oral | 2.1654 | 0.0061 |
| P163-PWY | L-Lysine Fermentation To Acetate And Butanoate | Oral | 2.1654 | 0.0071 |
| PWY-7013 | (S)-Propane-1,2-Diol Degradation | Oral | 2.2730 | 0.0211 |
| PWY-5177 | Glutaryl-CoA Degradation | Oral | 2.4111 | 0.0173 |
| PWY1G-0 | Mycothioliol Biosynthesis | Oral | 2.4316 | 0.0054 |

TABLE 2.3 (CONTINUED)

| Pathway | Description | Associated Body Site | Effect | Overlap |
|----------------|---|---------------------------------|---------------|----------------|
| PWY-6383 | Mono-Trans, Poly-Cis Decaprenyl Phosphate Biosynthesis | Oral | 2.4682 | 0.0048 |

TABLE 2.4

PREDICTED METACYC PATHWAYS WITH AN EFFECT SIZE OF > 1.0
 AS ESTIMATED BY ALDEX2 – COMPARISON FOR FIRMICUTES- OR
 BACTEROIDETES-RICH GUT SAMPLES.

| Pathway | Description | Firmicutes- /Bacteroidetes- rich | Effect | Overlap |
|----------------------|---|--|------------|------------|
| FUC-RHAMCAT- PWY | superpathway of fucose and rhamnose degradation | Bacteroidetes | 1.00388961 | 0.12932004 |
| PWY-7003 | glycerol degradation to butanol | Bacteroidetes | 1.0452088 | 0.12374591 |
| PWY-7377 | cob(II)yrinate a,c-diamide biosyn- thesis I (early cobalt insertion) | Bacteroidetes | 1.27260623 | 0.07692323 |
| FUCCAT-PWY | L-fucose degradation | Bacteroidetes | 1.33122869 | 0.06361626 |
| PWY-6588 | pyruvate fermentation to acetone | Bacteroidetes | 1.59193577 | 0.06250019 |
| METH- ACETATE-PWY | methanogenesis from acetate | Firmicutes | 1.00996758 | 0.10033456 |
| PWY-5088 | L-glutamate degradation VIII (to propanoate) | Firmicutes | 1.05700591 | 0.14285722 |

TABLE 2.4 (CONTINUED)

| | | | | |
|----------|---|------------|------------|------------|
| PWY-6478 | GDP-D-glycero- α -D-manno-heptose biosynthesis | Firmicutes | 1.17303832 | 0.10367904 |
| PWY-7210 | pyrimidine deoxyribonucleotides biosynthesis from CTP | Firmicutes | 1.22288385 | 0.07580841 |
| PWY-7198 | pyrimidine deoxyribonucleotides de novo biosynthesis IV | Firmicutes | 1.2504326 | 0.06020087 |

2.5 Discussion

Our analysis details the oral and gut microbiomes of long-tailed macaques in Singapore through community structure, variance, and taxonomic differences. We find higher oral richness and evenness, but less similarity and dispersion across our sampling sites. In contrast, gut microbiomes are more varied across our sites, have slightly fewer defining taxa, and have fewer identified predictions that might relay functions of these taxa. The phyla Proteobacteria, Firmicutes, Bacteroidetes, and Actinobacteria are highly represented in our samples, and the F/B ratio is only positive in oral samples, but divided in gut samples. Finally, we identify differences between predicted functions in F/B environment, collectively providing more insight into the utility of using phyla-level interactions to study the microbiome.

The higher oral richness observed in our samples may reflect the oral cavity's physiological properties and the constant influx of environmental microbes through foraging, social, and grooming behavior. The microhabitats of the oral cavity contribute to the composition of the bacterial community, providing colonization sites and areas where plaques and biofilms form, and may add to the increase in diversity within this environment. These oral cavity properties are unlikely to differ significantly between groups and potentially contribute to the high evenness and similarity in microbial composition reflected across the geographical space in our sampling sites (Dewhirst et al., 2010; Fellows Yates et al., 2021; Han et al., 2016). In addition, the high diversity of the oral cavity in this study could reflect the widely varying diet of free-ranging long-tailed macaques sampled in Singapore, which frequently includes anthropogenic-sourced food (Fuentes et al., 2008). This result adds to recent evidence of high diversity of oral microbiomes in free-ranging macaques (Sawasong et al., 2021), and recent reports that the oral microbiome has distinct properties across primates (Asangba et al., 2022). Our results may be compared with those presented in rhesus macaques, for example, oral micro-

biomes of Janaik et al. and Chen et al. were less rich than anal microbiomes, but the distinctive grouping of oral microbiomes is similar (Chen et al., 2018; Janiak et al., 2021). While we note differences in study design, sampling technique, and diversity measures in these studies, the collective evidence of a distinctive oral microbiome is well supported by our study.

Evenness and similarity of oral communities across sampling sites is another possible reflection of the environmental influence on the bacteria in the oral cavity, or rather, the lack of host influence. The constant influx of taxa when consuming food or water, combined with a relatively homogeneous landscape would facilitate an oral bacteria community populated from the external environment, as opposed to the gastrointestinal tract, in which established strains are adapted to a host and able to prevent the colonization and proliferation of incoming bacteria (Ochman et al., 2010). In addition, the microbes that can proliferate and colonize further down the digestive tract have been initially subjected to the harsh conditions of the stomach and are more likely to be specialized and stay with hosts, even as they migrate across long distances, an effect that can be detected more strongly than phylogeny (Poole et al., 2019). These influences could be further elucidated by comparing or expanding this study to a similar design that both investigates strains of host-associated microbes and incorporates variation in geographic and genetic structure in a single species of free-ranging primates.

More taxa defined the oral environment when compared to the gut, and more diversity overall was measured in this body site. Linear discriminant analysis identified defining oral taxa in nine phyla. Actinobacteria and Firmicutes, two highly abundant phyla in our saliva samples, have been recorded to dominate supragingival plaque in humans (Keijser et al., 2008); these phyla, along with Proteobacteria, also commonly dominate the oral cavity in humans (Dewhirst et al., 2010).

Gut microbial composition was defined by families within six phyla (Firmicutes,

Actinobacteria, Bacteroidetes, Proteobacteria, Cyanobacteria, and Tenericutes), with the highest number of defining families belonged to the phylum Firmicutes. Gut microbiome communities were less rich when compared to the oral environment, reflecting a possible bottleneck effect along the gastrointestinal tract and the inability of many abundant oral bacteria to form colonies in the gut environment (e.g., due to the presence of gastric enzymes and decreased pH). As our sampling strategy did not capture saliva and fecal samples from the same individual, these data are insufficient to thoroughly examine this possibility. Similarity of gut communities between individuals and sampling locations was lower than the saliva samples, which could be reflective of greater host influence in the gut when compared to the oral cavity. Although there is evidence that shared digestive physiology between primate species has a greater influence on microbial composition than dietary niche differences (Han et al., 2016), gut microbial communities of animals with very specialized diets are observed to cluster together over phylogeny or physiology (Trosvik et al., 2018). Another source of variation, particularly in the gut microbiome, is the diet, which can fluctuate within a group due to rank and resource availability (Fuentes et al., 2011). An examination of microbiome differences on the scale of a single species across a single island would allow more subtle influences of environmental variation, such as access to and competition for diet items, to emerge.

The F/B ratio is useful to examine in more detail the relationships between phyla in and between the oral and gut bacterial communities. Using the F/B ratio as a testing framework clarifies differences of phyla in our results by highlighting changes of family abundances underlying shifts in bacterial communities. Relative abundances of families within the Firmicutes and Proteobacteria phyla changed more than families in other phyla when looking at shifts between a Firmicutes- or Bacteroidetes-rich gut community. Out of thirteen families that changed abun-

dance between these compositions, the family Prevotellaceae, belonging to the phylum Bacteroidetes, increased significantly when the overall F/B ratio shifted away from Bacteroidetes abundance. This may indicate that changes in this ratio are more due to the fluctuations of families in Firmicutes than Bacteroidetes. The difference in ratios between the oral and gut environments may be related to the functions for which a particular bacterial community is optimized. For example, the Bacteroidetes phylum contains taxa that can protect the host against colonization of pathogens, which may be especially important in free-ranging mammals (Buffie and Pamer, 2013). Conversely, Firmicutes contains taxa that contribute to energy harvest in human guts and likely have an advantage in the oral environment, where some carbohydrate digestion takes place and can be associated to the taxa present (Flint et al., 2007; Poole et al., 2019). In this study, four of the families with higher abundances associated with a Firmicutes-rich gut environment (Prevotellaceae, Ruminococcaceae, Lachnospiraceae, and Clostridaceae_1) are producers of short chain fatty acids, which are important in host energy metabolism and homeostasis (Xu et al., 2020). In an analysis of predicted functional pathways to compare samples with contrasting F/B ratios, we found an equal number of pathways predicted in both Firmicutes- and Bacteroidetes-rich samples. In both of these types of samples, we detected degradation and biosynthesis pathways with high effect sizes and metabolism products indicative of energy production (Oliphant and Allen-Vercoe, 2019).

A predicted function analysis revealed more pathways with a significant effect size in the oral microbiome, which is consistent with our expectations, given the higher taxonomic richness seen in these samples. The greater number of pathways and patterns of diversity we find associated with the oral cavity, in which free-living microbial communities are more likely to be observed than in the gut, reflects the higher richness and diversity seen in these communities worldwide (Thompson

et al., 2017). Expanding this study to capture the metabolic profile of prokaryotes distinguishing host body sites, would grant deeper insight into how the composition of bacterial communities are maintained and their role in the host-microbe relationship (Lloyd-Price et al., 2017). Additionally, while ASVs, phylum-level, and family-level dynamics are powerful tools useful for categorizing communities, particularly across multiple populations, capturing genomic variation would greatly enhance sensitivity in detecting the genetic pathways and mechanisms driving the dynamics of host-microbiome relationships captured in this work. Eukaryotic microbes that co-occur in the gut were not included in this study, but are probable additional factors that can greatly influence the variation seen in the fecal samples (Leung et al., 2018; Wilcox and Hollocher, 2018). In particular, protozoans, such as *Blastocystis*, which feed on prokaryotes, and helminths, which have immunomodulatory effects on hosts, likely have a profound impact on the prokaryotic gut community composition through important bottom-up and top-down processes, respectively, not as readily associated with oral microbiomes (Deng et al., 2021; Wilcox et al., 2014).

Our sampling sites on the island of Singapore are similar with respect to anthropogenic influence, vegetation, and proximity to water sources; yet despite these shared environmental characteristics, we found significant inter-site variation in gut microbiomes (Sha et al., 2009a). Although our analysis did not examine the diet of the macaques, it is possible that small-scale variation between sampling sites reflected in our analysis can be attributed to different access to diet items and competition with other animals, including other macaques not part of our study. These differences were not observed in oral samples, however, which is reflected by the similarity of these samples across geographical variation in our study. The similarity exhibited by oral samples may provide insight into the plasticity of this environment, particularly in comparison with the gut. As shown by our testing

of composition by environmental variation, the similarity of the oral samples was strong enough to obfuscate the variation in the gut samples when these samples were analyzed together. When separated, however, this variation was significant in the gut, indicating that gut samples more closely resemble those at the same sampling location rather than those separated by larger geographical distances.

2.6 Conclusions

Our findings highlight multiple structural differences between the oral and gut microbial communities of a population of free-ranging primates in close proximity to anthropogenic activity. We find clearly distinct communities of prokaryotic microbes, and evidence of shifts in phylum-level composition between the oral and gut environments, exemplified by the shifting F/B ratio between the saliva and fecal samples. The similarity of oral bacterial communities across individuals and geographic space, in addition to the higher alpha diversity found here, may echo the essentially unconstrained entry and availability of environmental bacteria. As the oral cavity acts at the first line of defense to pathogens, a more diverse pool of bacteria present can be advantageous to the host via colonization resistance. Dissimilarity between sampling sites seen in the gut could, therefore, be reflective of a community that is highly specialized to an individual host.

CHAPTER 3

DIET PRESENTS OPPORTUNITIES TO STUDY THE ROLE OF HUMAN INTERACTIONS IN SHAPING GUT MICROBIOME IN URBAN ENVIRONMENTS

3.1 Abstract

Dietary and environmental variation well known for their role in structuring the microbiomes of humans and non-human primates in wild and captive environments. These effects are not well understood in the edges of urban environments, and the species who occupy them, despite their demonstrated flexibility in adaptation. In this study, macaque gut bacterial microbiomes were assayed simultaneously with diet across a range of anthropogenic environments. We used a K-means clustering algorithm based on bacterial Amplicon Sequence Variants (ASVs) data and found that clusters do not correspond well to broad geographical structuring. However, bacteria associated with the clusters recovered show characteristics relating to different microenvironments and digestive functions. Chloroplast and plant items are correlated in our samples, and are more abundant in samples from sites where provisioning is forbidden. Conversely, in sites where provisioning is common, we see higher crop abundance and richness, but do not observe a linear relationship to chloroplasts, likely due to provisioned foods being most fruits or tubers, rather than leafy parts of a plant. Bacteria richness is inversely related to dietary richness, which may be explained by the food availability in highly urban sites. We find bacterial taxa correlating to clusters reflecting dietary information. Our analysis reveals the importance of considering the interactions

between diet and environmental influences on the gut bacterial communities of primates in close proximity to humans, which has cascading effects important in understanding the role of plasticity in the microbiome.

3.2 Background

Gut microbiomes of primates are essential in understanding host adaptability to environment and are strongly influenced through a multitude of external factors. Abiotic factors such as soil variation and seasonal fluctuations shape the taxonomic composition and metabolic properties of bacterial communities in primates (Baniel et al., 2020; Grieneisen et al., 2019). Short-term changes in diet can also induce rapid changes in taxonomic composition and community structure, while long-term diet patterns are linked to evolutionary patterns across mammals (David et al., 2014b; Muegge et al., 2011). Environmental and dietary factors interact as seen in the impact of human-induced habitat degradation on primate gut microbiomes via changes in access to dietary items (Amato et al., 2013; Barelli et al., 2015). Research focusing on captive primates has helped to elucidate these interactions by analyzing responses to effects controlled environments and diets exert on gut microbiomes, particularly in the cases of primates with specialized dietary niches (Clayton et al., 2016; Frankel et al., 2019; Greene et al., 2018). How host microbiomes adapt to environments in species adapted to edge environments, however, which can neither be considered captive or truly wild, has not been well characterized. One such example is the macaque (genus *Macaca*), a cercopithid known for its adaptability and plasticity in response to shifting ecological landscapes, exemplified by a broad distribution throughout Southeast Asia (Abegg, 2002; Thierry, 2007). In free-ranging macaques, home ranges often cross the boundaries between more natural, agricultural, and urban habitats, and macaques

demonstrate their ecological flexibility by utilizing anthropogenic resources, such as crops and deliberately provisioned human food sources (Gumert, 2011; Ilham et al., 2018; Wheatley et al., 2002). A recent study by Lee et al. finds macaque (*M. fuscata*) gut microbiomes respond across a range of human disturbance types and habitats (i.e., wild, captive, provisioned, etc.), but point to diet effects to explain variation within these disturbance types (Lee et al., 2019). We combine this data with 16S and 18S rRNA data extracted from free-ranging long-tailed macaque (*M. fascicularis*) fecal samples collected across a range of environments in Bali and Singapore. Our aim is to target and identify sources of bacterial variation in these edge species by testing the similarity of microbiomes throughout this dataset. We then use diet to explain variation in Singapore and Bali samples. We predict that microbial variation will reflect the dietary and location-based differences in the macaque environment, through community composition and differences in specific taxa which reflect to diet function. Further, we expect to find patterns in diet which reflect the location of macaques and aim to connect these to the context of microbiome variation.

3.3 Methods

3.3.1 Sampling and DNA Extraction

We non-invasively collected fresh fecal samples (N=86) from macaques (*Macaca fascicularis*) at eight sites in Singapore (N =41) and nine sites in Bali (N=45). Samples were placed on ice immediately, shipped to Notre Dame, IN, US on dry ice, and remained frozen at -85°C until DNA extraction [detailed procedure as described in previous publications (Klegarth et al., 2017; Lane-deGraaf et al., 2014a). DNA was extracted from fecal samples using the Qiagen DNA Stool minikit. The V9 region of the 18S rRNA gene and the V3-V4 region of the 16S rRNA gene were

amplified and PCR was performed to isolate gene regions as described in Chapter 2 and (Wilcox and Hollocher, 2018). Further details regarding proportions of DNA and primer used in PCR master mix can be found in Appendix A.2.1. The resulting DNA library was sequenced on an Illumina HiSeq 2500 (Illumina, San Diego, CA, USA), using a rapid run with paired-end 250bp reads at the University of New Hampshire Hubbard Center for Genome Studies.

3.3.2 16S and 18S rRNA Read Processing

To incorporate samples from Lee et al., we downloaded 16S rRNA reads hosted on the NCBI Short Read Archive (BioProject Number: PRJDB8636) (Lee et al., 2019; Leinonen et al., 2011). Reads from Lee et al. were downloaded in fastq format, and filtered using the `filterAndTrim` function in the R package Dada2 v. 1.22.0 (Callahan et al., 2016; R Core Team, 2020). Reads sequenced from sample collection in Bali and Singapore were filtered separately using the same function. After filtering, reads were combined and used as input for the `dada2` function to create Amplicon Sequence Variants (ASVs), create a count table, and assign taxonomy to ASVs (Benjamin Callahan, 2017; Callahan et al., 2016; Maidak et al., 1997). The packages Phyloseq v 1.38.0 and Microbiome v 1.16.0 were used to filter ASV data to the kingdom Bacteria and to format for downstream analyses (Lahti and Shetty, 2012; McMurdie and Holmes, 2013). Reads assigned to taxonomy order chloroplast were summed and used to as a potential estimate for plant ingestion and fiber intake in downstream analyses (Ochman et al., 2010). Reads from 18S rRNA amplicons were received in fastq format, and followed a pipeline including filtering, denoising, and operational taxonomic unit (OTU) construction using the Usearch v 8.1 and Vsearch v 2.18 software tools (Edgar, 2010; Rognes et al., 2016). SILVA v128 was used as a reference database for taxonomic assignment (Quast et al., 2012). Diet was filtered from 18S rRNA OTUs using the taxonomic

framework combined with known macaque diet information. Specifically, taxa in the clade Embryophyta and the phyla Mollusca and Arthropoda were isolated and merged at the genus level. This decision was made after reviewing diet information of macaques in (Gumert, 2011; Sha and Hanya, 2013; Yeager, 1996). Dietary genera were then filtered using an intelligent cutoff of 6, following (Cirtwill and Hambäck, 2020). Embryophyte items were subsequently categorized as crop or non-crop items. A complete list of crop items can be found in Appendix A.9.

3.3.3 Statistical Analysis

The 16S rRNA ASV table was filtered (`codaSeq.filter` function from `CoDaSeq` (Gloor and Reid, 2016; Gloor et al., 2016b) and transformed using the Hellinger transformation functionality in the `decostand` function from the `vegan` package (v 2.6.2) (Oksanen et al., 2022). The transformed ASV table was then clustered using a K-means algorithm (`K-means` function in R stats package), setting $K = 3$ to view recovery of samples by location (i.e., Singapore, Bali, and Japan). To validate clusters, we used `Fossil` to calculate the Adjusted Rand Index, which ranges from 0 to 1, where 1 indicates perfect agreement between clusters and input sample data (Steinley, 2004; Vavrek, 2011). The function `envfit` from the `vegan` (v 2.6) package was used to identify bacterial genera correlated with clusters, which were then plotted as vectors to overlay clusters. Bacteria were categorized as ‘moderate’ when correlations were greater than 0.5 and as ‘strong’ when greater than 0.7. We then compared samples by cluster to identify differences between them more thoroughly. To see how bacterial diversity associated with clusters, we calculated ASV richness using the `vegan` package and compared samples using Welch’s T-test in R. We then compared chloroplast read counts (Welch’s T-test) to estimate plant consumption between clusters. At this point, we removed samples from Lee et al., as downstream analyses focused on plant reads derived from 18S rRNA

reads. Monkeys in our sampling design are provisioned with roots, legumes, and fruits of crops (e.g., potatoes, bananas, peanuts), which we did not expect to be represented by chloroplasts in barcoding (Fuentes et al., 2011; Leca et al., 2008). Therefore, we sought to find whether relationships exist in our data linking chloroplast to crop and/or non-crop items, and we used linear models (stats package, R) to test the strength of these relationships. Finally, we compared the abundance and richness of crop and non-crop plants between clusters.

3.4 Results

Readcounts of samples from 16S rRNA in our dataset totaled 11,000,795. Combined with Lee et al.'s data, our 16S rRNA dataset analyzed 20,710,247 reads. We recovered 13,953 ASVs from our pipeline. Our 18S rRNA reads totaled 219,422,829 representing 53,217 taxa. Out of these, 1,020 were counted as diet items (Bali = 745, Singapore 783), 62 plants were counted as crop plants (Bali = 58, Singapore = 61). K-means clustering found that the recovery of islands in clusters was poor (Figure 3.1, Rand index 0.37), indicating that location is less important than other factors in structuring our dataset (Table 3.1). Clusters 2 and 3 group together, showing these two clusters are more similar to each other than Cluster 1.

Genera correlating to clusters are listed in Table 3.2. *Prevotella* correlated highly with Cluster 3 (Green). Cluster 1 has lower ASV richness than either Cluster 2 ($t = 4.662$ $df = 54.415$ $p = 2.07 \times 10^{-05}$) or Cluster 3 ($t = -4.3242$ $df = 48.057$ $p = 7.681 \times 10^{-05}$). ASV richness between Clusters 2 and 3 does not differ significantly.

Chloroplast reads differed significantly between clusters. Cluster 1 contains more chloroplast reads than either Cluster 2 ($t = -2.786$, $df = 70.03$, $p = 0.007$) or 3 ($t = -3.6828$, $df = 50.729$, $p \leq 0.001$). Crop plants do not correlate to chloro-

TABLE 3.2

HIGHLY CORRELATING BACTERIAL GENERA WITH CLUSTERS

| Cluster 1 (Red) | Cluster 2 (Brown) | Cluster 3 (Green) |
|-----------------------------|-----------------------|------------------------------|
| <i>Acinetobacter</i> | <i>Treponema</i> | <i>Prevotella</i> |
| <i>Escherichia.Shigella</i> | <i>Oscillobacter</i> | <i>Alloprevotella</i> |
| | <i>Clostridium IV</i> | <i>Sutterella</i> |
| | <i>Intestimonas</i> | <i>Faecalibacterium</i> |
| | <i>Sporobacter</i> | <i>Roseburia</i> |
| | | <i>Phascolarctobacterium</i> |

Note: All taxa listed have a correlation value >0.5 . Bolded genera have a correlation >0.7 . The genus *Escherichia.Shigella* combines two genera due to sequence overlap in taxonomy assignment.

plast reads (beta = -2.48×10^{-5} , $R^2 = -0.012$, $p = 0.979$), while non-crop plants have a positive linear relationship with chloroplast abundance (beta = 3.51×10^{-3} , $R^2 = 0.156$, $p < 0.001$) (Figure 3.2).

Diets of samples in Cluster 1 contain fewer crop plants than Cluster 2 ($t = -2.631$ df = 20.124 $p = 0.016$) and more non-crop plants than either Cluster 2 or 3 (1 and 2: $t = -5.105$ df = 66.85 $p = 2.95 \times 10^{-6}$, 1 and 3: $t = 2.495$ df = 23.517 $p = 0.020$; Figure 3.3). In comparing 16S rRNA to 18S rRNA counts, diet genera richness decreases as ASV richness increases (beta = -0.11169 , $R^2 = 0.190$, $p = 1.67 \times 10^{-5}$). This relationship is also seen in crop plants (beta = -0.629 , $R^2 = 0.109$, $p = 0.001$; Figure 3.4). Cluster 1 has a higher diet richness than either Cluster 2 ($t = -2.989$, df = 67.498, $p = 0.004$) or Cluster 3 ($t = -2.830$, df = 39.392, $p = 0.007$; Figure 3.5). Crop genera richness displays a similar pattern: Cluster 1 has higher crop richness than Cluster 2 ($t = -2.093$, df = 67.56, $p = 0.040$) or 3 ($t = 2.144$, df = 32.784, $p = 0.040$).

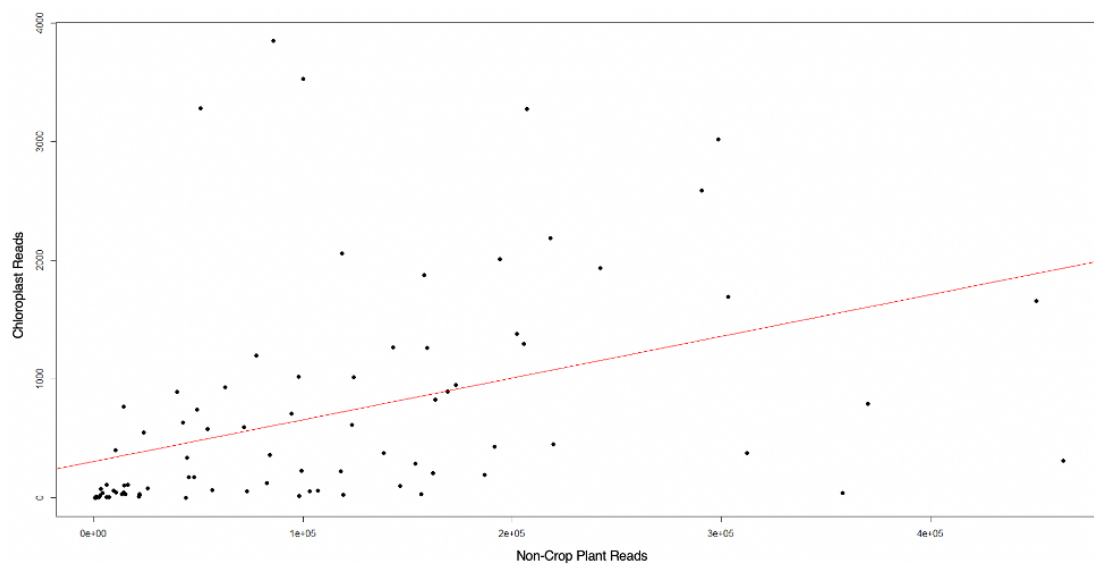


Figure 3.2. Linear Relationship Between Chloroplasts and Non-Crop Plant Reads. Relationship is positive and significant ($\beta = 3.51e^{-03}$, $R^2 = 0.156$, $p < 0.001$). Chloroplast counts obtained from 16S rRNA reads and non-crop counts are calculated from 18S rRNA reads

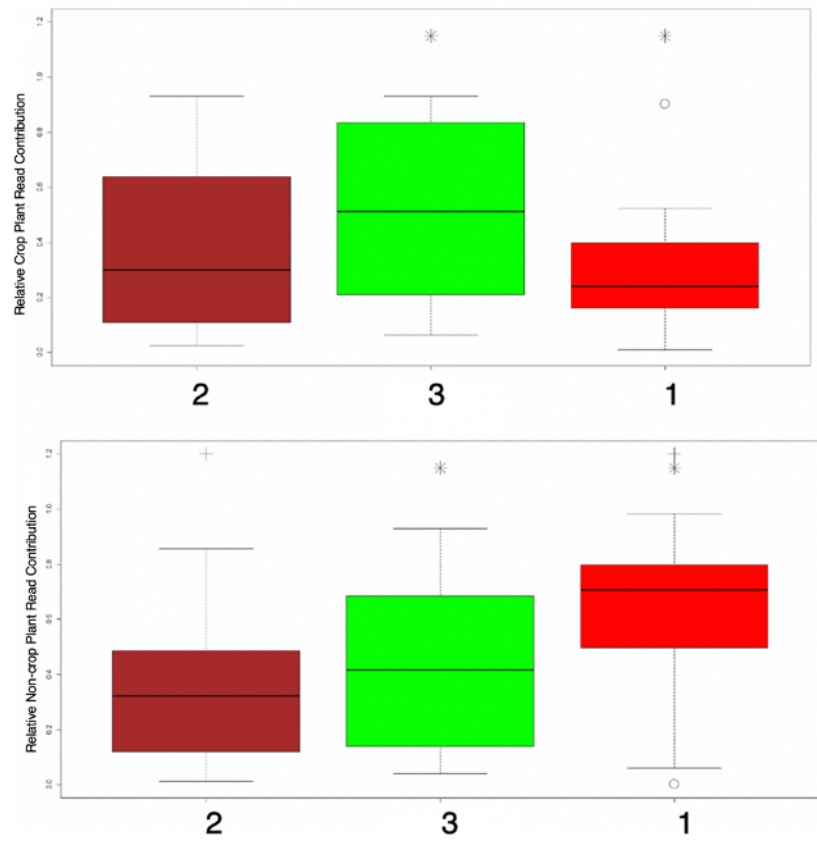


Figure 3.3. Crop and Non-Crop Plant Reads Between Clusters. Samples in Cluster 1 (red) have fewer relative crop plants than samples in Cluster 3 (green; $t = -2.6305$, $df = 20.124$, $p = 0.016$). Miscellaneous plants make up more of the diet in samples from Cluster 1 than samples in Clusters 2 and 3 (brown; 1 and 2: $t = -5.1049$, $df = 66.85$, $p = 2.95 \times 10^{-06}$; 1 and 3: $t = 2.495$, $df = 23.517$, $p = 0.020$).

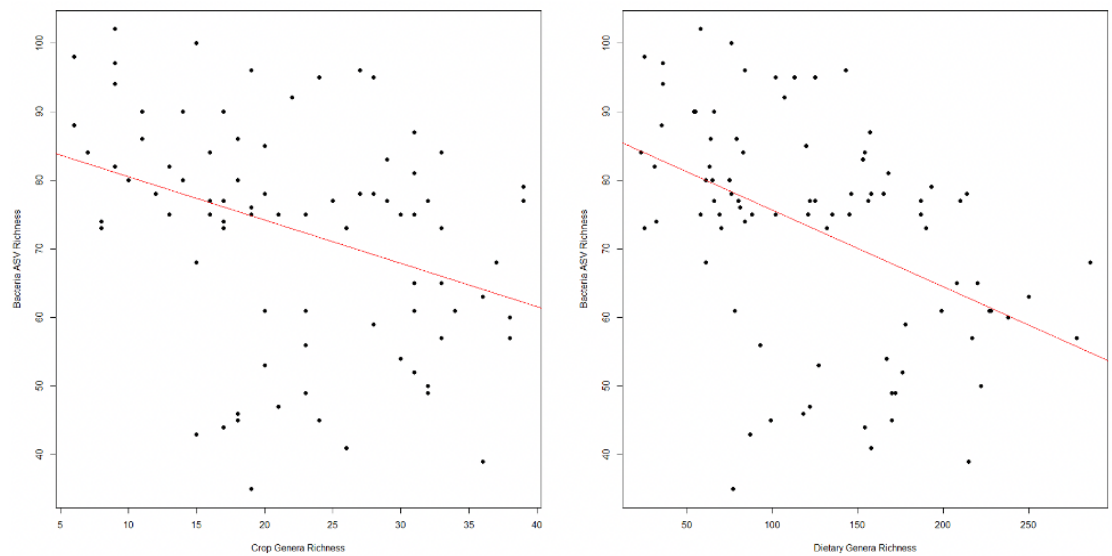


Figure 3.4. Linear relationships between bacterial ASV richness and crop richness (left) and dietary genera richness (right) in Singapore and Bali samples. Crop richness and dietary genera richness both show a linear relationship with bacterial richness (crop: $R^2 = 0.109$, $p = 0.001$, dietary genera $R^2 = 0.190$, $p = 1.67 \times 10^{-05}$).

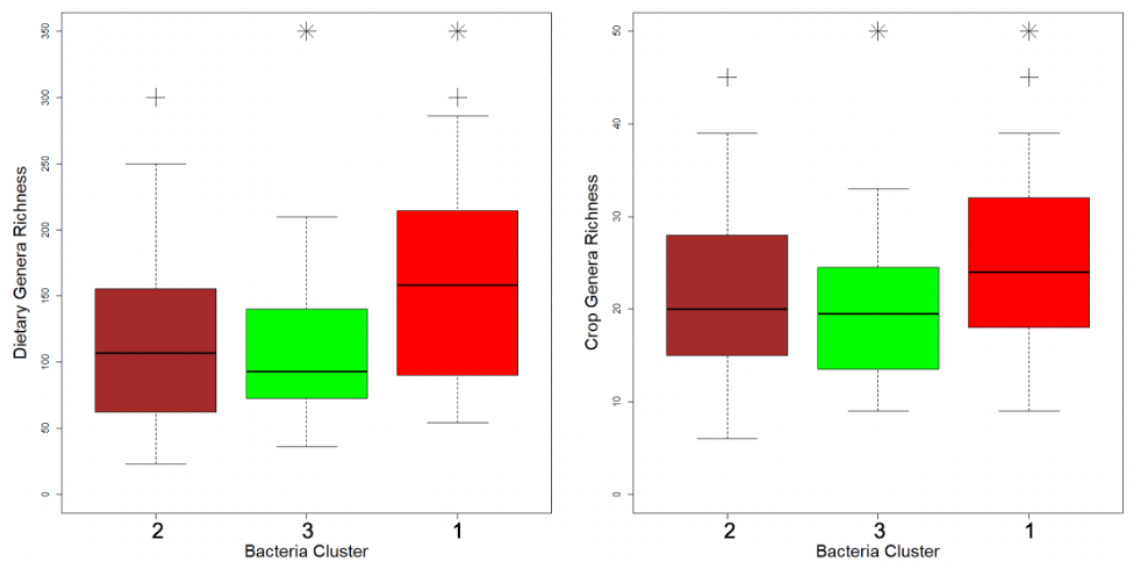


Figure 3.5. Dietary genera richness (left) and crop genera richness (right) by bacteria cluster. Only samples in Singapore and Bali are tested and plotted. Cluster 1 has higher dietary and crop richness than Clusters 2 or 3 (dietary richness: 1-2 $p=0.004$, 1-3 $p = 0.007$; crop richness: 1-2 $p = 0.04$, 1-3 $p = 0.04$).

3.5 Discussion

Our analysis of uses a clustering algorithm to detect variation in gut microbiome data collected from Southeast Asia. We predicted samples would cluster according to location, however our analyses showed location be a poor predictor of variation. We then examined bacterial and diet differences between clusters and show differences in diet measured through chloroplast presence, diet richness and crop consumption associated with samples in clusters. Together these data relay differences in diet shaping the microbiome across and throughout populations.

Our initial expectations predicted location would exert a strong effect in the formation of sample clusters given the importance of location in previous microbiome studies in wild primates (Degnan et al., 2012; Eschweiler et al., 2021; Fogel, 2015; Gomez et al., 2015). Comparisons of wild and captive primate microbiomes strengthen this distinction, but conditions of captivity such as a controlled diet, leave the role of location confounded (Clayton et al., 2016; Frankel et al., 2019). Species are also not defined clearly in clusters, suggesting that diet and environmental factors in these samples impact the bacteria present in the gut to a greater degree than phylogeny, which has previously been observed in Colobine monkeys in captivity (Hale et al., 2018).

Macaques in Southeast Asia have an extremely broad diet which reflects food availability in the environment, and provisioning shifts consumption away from wild fruits and flowers (Fuentes et al., 2011; Sha and Hanya, 2013). Macaques in Indonesia in highly provisioned sites have high consumption of human-acquired food and adjust their behavior in response to patterns of provisioning (Brotcorne, 2014; Ilham et al., 2017, 2018; Wheatley, 1999). Crop plants in our samples show a heavy representation of plants used for provisioning, such as sweet potatoes (*Ipomoea*) and corn (*Zea*). Provisioning practices commonly supply fruits or roots of plants, which are unlikely to leave traces of chloroplasts detectable by 16S rRNA

sequencing, and could explain low chloroplast presence in monkeys in Clusters 2 and 3 (Loudon et al., 2006).

Conversely, non-crop plants in the macaque diet are more likely to be consumed as leafy material during foraging, which would explain the both the high count of chloroplasts in Cluster 1, which has high representation of less-provisioned sites, and the linear relationship we observe between these plants and chloroplast reads. Effects of leafy diets on the microbiome have similarly been observed in wild gorillas (Moeller et al., 2015). The impacts in diet we observe on formation of ASV clusters lead us to conclude that differences in provisioning and foraging behavior can be differentiated in macaque gut microbiomes.

We found Cluster 1 had the highest amount of chloroplast reads and the highest dietary genera richness in comparison to Clusters 2 and 3. This finding combined with our observation of a relationship between non-crop plants and chloroplasts suggests that monkeys in Cluster 1 consume more volume and types of foliage than monkeys in Clusters 2 and 3, which are more likely to be heavily provisioned (Lane et al., 2011; Lee et al., 2019). Sites in Singapore are considered less-provisioned due to land management policies in these sites which are implemented to discourage people from feeding wildlife, though macaques are able to obtain food directly from people, and through raiding of trash and household properties (Fuentes et al., 2008; R Core Team, 2020; Sha et al., 2009b).

Literature directly measuring relatedness of dietary richness and bacterial richness is conflicting, possibly due to a lack of cohesiveness in measuring dietary richness (Bowyer et al., 2018) and focus on microbial changes in response to niche or non-specialized diets (Ley et al., 2008). Claesson et al. found dietary richness associated with bacterial richness in humans, for example, while Ren et al. found no relationship in wild baboons (Claesson et al., 2012; Ren et al., 2016). Cluster 1, which has the highest dietary and lowest ASV richness, is primarily com-

posed of samples from the highly urban environment of Singapore, may be related to decreases in ASV richness (Ayeni et al., 2018; Das et al., 2018).

Bacteria correlating with clusters provides further insight into the diets of samples in clusters. Only two genera correlate to Cluster 1, *Acinetobacter* and *Escherichia.Shigella*; these are common inhabitants of gastrointestinal tracts of humans and non-human primates (Clayton et al., 2018). Comparatively, Clusters 2 and 3 have more taxa correlated with them, and these taxa provide further insight into the diet of these samples. Genera correlating with Cluster 2 have a high association with the production of butyrate and diet fiber breakdown (*Treponema*, *Oscillibacter*, *Clostridium IV*, *Intestimonas*, and *Sporobacter*) (Bui et al., 2015; Chen et al., 2018; Claesson et al., 2012; GRECH-MORA et al., 1996; Ma et al., 2014; Pryde et al., 2002; Seekatz et al., 2013). Cluster 3 correlates strongly (>0.7) with the genus *Prevotella*, which has carbohydrate digesting and fermentation properties and is a common member of macaque and other primate microbiomes (Chen et al., 2018; Lee et al., 2019; Ma et al., 2014; McKenna et al., 2008; Ren et al., 2016; Sawaswong et al., 2021). Connections between bacterial taxa such as *Prevotella* and diet, however, are not well supported (Moeller et al., 2012).

We did not directly test individual environmental variables in our samples due to a lack of comparative metrics for sampling sites between islands. For instance, samples in our Bali populations have been measured for proximity to urban areas and amount of provisioning, but these metrics are not available in Singapore (Lane et al., 2011). In addition, the land mass of Singapore is very small in comparison to Bali (685 km² and 5,780 km², respectively) and the human density is 8019 people/km², while the city with the highest population density in Bali is roughly half of that (Bal). Thus, measurements that attempt to capture variation between sites such as proximity to urban areas in Singapore are unlikely to withstand confounding effects. Even so, there are some trends we observed in the data that we deem

worthy of discussion.

Cluster 1 contains samples from Singapore and Balinese samples which are primarily from the site LM, which contains the most urban area of any sites on Bali (Lane et al., 2011). This finding resembles other examples of changes seen in non-human primate microbiomes in close proximity to an anthropogenic environment due to diet changes and habitat fragmentation or degradation (Amato et al., 2013; McCord et al., 2014). Cluster 2 contains primarily Bali sites, several of which are highly-provisioned. Sites U and PU in particular contain large populations of macaques who are routinely provisioned and a feature of tourism in Bali (Fuentes et al., 2011). Samples from Singapore which fall into Cluster 2 are from sites bordering the Central Catchment Nature Reserve (MR, US, LP), which is a popular tourist and recreational destination with visitors who, despite fines imposed by the National Parks Board, are observed feeding macaques (200, 2008). Half of samples from a site in Japan visited frequently by tourists (J-SH) are found in Cluster 2 as well. We suggest effects of tourism, which has demonstrable effects on the wildlife attractions, on microbiomes be centered in aims to source of variation in primate diets and microbial communities (Cui et al., 2021). Cluster 3 contains samples from Singapore which are informative from a conservation perspective, as they come from the only two sites in our study which have direct access to the little primary forest remaining on Singapore, the Bukit Timah Nature Reserve and the Yee Soon Swamp Forest (US)(Corlett, 1992; Klegarth, 2015).

3.6 Conclusions

These relationships could be further elucidated by combining this barcoding approach with direct observations of provisioning in macaques and other urban animals on a large scale, and a more quantitative understanding of which and how

properties of urban environments relate to the gut microbiomes of animals in their vicinity. Our findings collectively highlight variation in the gut bacterial microbiome driven by differences in diet of macaque hosts. The environmental and anthropological effects of this variation are difficult to untwine, and holistic approaches are integral to elucidating the dynamics underlying microbial communities.

CHAPTER 4

EVIDENCE FOR SELF-MEDICATION IN LONG-TAILED MACAQUES IN SOUTHEAST ASIA

4.1 Abstract

Primate consumption of plants with medicinal properties has been hypothesized as a strategy to alleviate or combat parasitic infections. Limited observations of feeding behavior in free-ranging primates have provided support for this hypothesis, yet many questions regarding the extent to which this behavior might occur remain. In this study, we employ barcoding techniques to evaluate interactions between parasites and medicinal plants in populations of free-ranging macaques (*Macaca fascicularis*) in Southeast Asia. Using these data, we cataloged 663 plant items in the macaque diet and cataloged their medicinal properties (antibacterial, antiprotozoal, antifungal, anthelmintic, or general human medicinal use). We find more than a quarter (286) of plants in the macaque diet have medicinal properties, and 48 of these have all medicinal properties targeted in our queries. We then used this database to test hypotheses related to the presence and abundance of medicinal plants and their potential role in parasitic infections. Widespread associations between medicinal plants and protozoans were uncovered, driven by protozoan parasites and particularly the genus *Plasmodium*. Our findings reveal the presence and abundance of parasites to be closely related to the consumption of plants exhibiting the greatest number of medicinal attributes. The presence and abundance of parasites also show a direct relationship to the diversity of medicinal plants consumed. These findings add support to literature surrounding the

role of self-medication as a potentially adaptive response to parasite presence in free-ranging primates.

4.2 Background

Deciphering the dynamics underlying primate and parasite interactions is integral to understanding infectious disease transmission dynamics in a global health context. Primates have a high capacity to host a range of infectious agents with low thresholds for host specificity (Pedersen et al., 2005). Due to their abilities to adapt and wide geographic dispersal, long-tailed macaques have long been counted as one of the most successful primates in Southeast Asia, although their status was recently updated to Vulnerable on the IUCN Red List (Hansen et al., 2021). Commonly found in edge habitats at the interface of forests and human construction, macaques have a high degree of contact with humans, making them of interest for transmission of zoonoses notable when considering infectious disease dynamics (Gumert, 2011; Shano et al., 2021). Long-tailed macaques are natural hosts to a multitude of helminths and protozoans, including several *Plasmodium* species (Lempang et al., 2022; Schurer et al., 2019; Zhang et al., 2016). *Plasmodium* presence in wild primates is a cause for spillover concern due to its extreme importance in humans and risk of transmission (Cox-Singh et al., 2008; Imwong et al., 2019; Jongwutiwes et al., 2004; Ng et al., 2008; Singh et al., 2004).

Although not all parasite infections present symptoms in a host, stress caused by malnutrition and external environment factors can be exacerbated by infection. Primate strategies of self-medication with plants to ameliorate infection-induced stress have been hypothesized and supported previously (Huffman et al., 1996). Reports of bitter pith chewing, whole leaf swallowing, and consumption of plants with medicinal compounds are some of the observed behaviors for self-medication,

primarily in African primates (Huffman, 1997; McLennan and Huffman, 2012). Nematode expulsion is associated with whole leaf swallowing in chimpanzees in Tanzania, and evidence of rough leaf-swallowing, marked by undigested whole leaves in feces, has been indirectly observed in gibbons in Thailand (Barelli and Huffman, 2017; Huffman et al., 1996).

Self-medication evidence provides potential avenues for the control of emerging infectious diseases, and highlight the importance of a biodiversity in plant landscapes. This research is primarily based on direct observation of primates ingesting plant material, such as the mastication and digestion of rough-textured leaves or bitter pith chewing (Huffman, 1997). These methods rely on observing the primate ingesting a plant, quantifying the volume of the plant eaten, and describing in detail the manner in which a plant was consumed (e.g., folding and swallowing leaves whole, or chewing the bark or pith of trees). Drawbacks to relying solely on observational data include difficulty in monitoring an individual animal's complete diet and increased logistical complications in scaling research to the level of populations. Observational studies are also costly when considering sampling effort, and field training requires an extensive knowledge of primate and plant taxonomy and the need to be highly skilled in systematic primate observational techniques. The development of barcoding techniques provides a way to bypass the logistical difficulties and complications associated with behavioral studies and allows for a rapid, simultaneous assessment of both parasite infection status and diet through the use of curated databases which harness the information becoming increasingly available as the use of sequencing technology becomes more widespread (Coissac et al., 2012; Valentini et al., 2016). This methodology provides broader insight into the variety of plants consumed and can include protozoans difficult to survey through microscopy alone (Aivelo and Medlar, 2018; Wilcox and Hollocher, 2018).

Southeast Asia feature landscapes full of rich plant life and biodiversity intermingled with highly urban areas. Long-tailed macaques (*Macaca fascicularis*) in Singapore and Bali are a free-ranging edge species and have unbarred access to tropical forested landscapes and anthropogenic habitat. They also demonstrate a capacity to host a range of parasites (Lane-deGraaf et al., 2014b; Zhang et al., 2016), providing an ideal system in which to study the role of medicinal plants in relation to parasite infection. Here we utilize a barcoding approach to capture the diversity of parasites and medicinal plants in this region (Wilcox and Hollocher, 2018), and predict that macaque parasitic infections, especially *Plasmodium*, can be reflected through consumption of medicinal plants.

4.3 Methods

4.3.1 Sample Collection

In Singapore, macaques are highly managed, partially due to ongoing conflicts, and generally contained in the Central Catchment of the island (Riley et al., 2016; Sha et al., 2009c). In contrast, macaques in Bali, Indonesia, are highly provisioned, a draw of wildlife tourism, and ingrained in the religious landscape of the human populations (Lane-deGraaf et al., 2014b). To evaluate the presence of eukaryotic parasites and medicinal plants, we collected fecal samples (N = 127 from populations of long-tailed macaques in Bali, Indonesia in the summer of 2007 and Singapore during the summers of 2012 and 2013 (Table A.10). Fresh fecal samples were placed on ice immediately, shipped to the US on dry ice, and remained frozen at -85°C until DNA extraction [detailed collection procedure described in Lane-deGraaf et al. (2014b) for Bali samples, and Klegarth et al. (2017) for Singapore samples].

4.3.2 DNA extraction/Sequencing

DNA was extracted from fecal samples using the Qiagen DNA Stool minikit. PCR of extracted DNA targeted and amplified the V9 region of the 18S rRNA gene and the V4 region of the 16S rRNA gene. Barcoding for 16S rRNA targeted a 287 bp sequence from genomic DNA using the following primers (Illumina adaptors in italics): Forward (S-D-Arch-0519-a-S-15, 5' - *TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CAG CMG CCG CGG TAA* -3'), Reverse (S-D-Bact-0785-b-A-18, 5' - *GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GTA CNV GGG TAT CTA ATC C* -3') (Klindworth et al., 2013; Van Bleijswijk et al., 2015). Methods for amplification of the 18S rRNA gene have been described previously (Wilcox and Hollocher, 2018). Further details regarding proportions of DNA and primer used in PCR master mix for library prep can be found in Appendix A.2.1. The resulting DNA library was sequenced on an Illumina HiSeq 2500 (Illumina, San Diego, CA, USA), using a rapid run with paired-end 250bp reads at the University of New Hampshire Hubbard Center for Genome Studies.

4.3.3 16S and 18S Read Processing

The program Usearch (v 8.0, 264GB) was used as previously described in Wilcox and Hollocher (2018) to process sequences directly from raw files (fastq format). Briefly, the sequences were quality filtered, dereplicated, and clustered at 97% to form operational taxonomic units (OTUs). These OTUs were assigned taxonomy using the SILVA reference database and a matching threshold of 80% and the SILVA database (release 132)(Quast et al., 2012; Yilmaz et al., 2014). From matched reads, a count table was constructed and consolidated to genus. From this table, we isolated all items which we classify as 'resident taxa', i.e. eukaryotic parasites with high probability of capacity to infect vertebrates (full list referenced

from Wilcox and Hollocher (2018). We then further isolated taxa to create helminth and protozoa tables and filtered counts using a cutoff value of 6, following Cirtwill and Hambäck (2020). Plant items were filtered from the 18S OTU table using the term “Embryophyte” and were the source data for the construction of a medicinal plant database.

4.3.4 Medicinal Plant Database Construction

Plants detected in the macaque diet were individually searched and compared to literature (764 citations, referenced in database available in supplementary Excel file) for anti-pathogenic activity and general human medicinal use. Queried plants were labeled as antiprotozoal, anthelmintic, antibacterial, antifungal, and/or human (plants with documented medicinal uses in humans, but which were not tested against pathogens directly) according to supporting literature. Based on these documented properties, plants were categorized with the following groups: general medicinal, Anti-all, or non-medicinal for statistical analysis. General medicinal is the broadest category and includes plants with at least one labeled medicinal property. Anti-all is the narrowest category and contains plants that possess all medicinal properties queried (i.e., these have documented antiprotozoal, anthelmintic, antifungal, antibacterial properties, and are used medicinally). Non-medicinal plants are those with no documented medicinal properties reported in the literature survey.

4.3.5 Statistical analysis

To investigate broad correlative relationships between medicinal plants and resident taxa, we constructed distance matrices of the following count tables: general medicinal plants, Anti-all plants, non-medicinal plants, helminths, protozoans,

helminths and protozoans combined, *Plasmodium* read counts, and all plant read counts using the vegan package in R [v. 4.1.3](R Core Team, 2022; RStudio Team, 2022) (Oksanen et al., 2022). Bray-Curtis distances were used to calculate all matrices except read counts, for which Euclidean distances were more appropriate. Resulting p-values were adjusted using the Bonferroni correction (p.adjust function in R stats package). Partial mantel tests were used to first compare the presence of all parasites to the medicinal, Anti-all, and non-medicinal plant items detected in the macaques (Pearson correlation, 9999 permutations). We performed separate tests comparing helminths, protozoans, and the genus *Plasmodium*.

To determine whether medicinal plant abundance differed in samples with *Plasmodium*, and account for uneven samples, Welch's t-tests were performed (stats package, R). Linear models (stats package, R) were then constructed to inspect relationships between *Plasmodium* abundance and medicinal plants abundance and richness. To ensure read count was not a confounding variable in linear models, we tested *Plasmodium* read counts against all plant read counts. As we were interested in how parasites and medicinal plants interacted in all samples and on smaller scales, we analyzed all samples together and separately for Bali and Singapore for all tests.

4.4 Results

4.4.1 Medicinal plant categorization

After documenting the known medicinal qualities and uses of 663 plants identified from 18S rRNA reads, we calculated overlap of the medicinal properties. A majority of medicinal plants detected in the macaque diet across both islands have multiple medicinal properties (Figure 4.1). General medicinal plants were the most common type of medicinal plant detected in the macaque diet. The second most

common type of medicinal plant were Anti-all plants. The ten most abundant plants in general medicinal, Anti-all, and non-medicinal categories can be found in Table 4.1.

TABLE 4.1
TEN MOST ABUNDANT PLANTS IN EACH CATEGORY USED FOR
ANALYSIS.

| General Medicinal | Anti-all | Non-Medicinal |
|-------------------------------|-------------------------------|---|
| <i>Octomeles sumatrana</i> | <i>Morus alba</i> | <i>Cryptocarya meisneriana</i> |
| <i>Actaea simplex</i> | <i>Gossypium arboreum</i> | <i>Rheum kialense</i> |
| <i>Vitis vinifera</i> | <i>Eucalyptus grandis</i> | <i>Perilla frutescens</i> var. <i>crispa</i> |
| <i>Elaeis guineensis</i> | <i>Anacardium occidentale</i> | <i>Orontium aquaticum</i> |
| <i>Calophyllum inophyllum</i> | <i>Ananas comosus</i> | <i>Freycinetia formosana</i> |
| <i>Solanum lycopersicum</i> | <i>Capsicum annuum</i> | <i>Eustoma exaltatum</i> subsp. <i>russellianum</i> |
| <i>Psidium cattleianum</i> | <i>Sesamum indicum</i> | <i>Setaria italica</i> |
| <i>Carya cathayensis</i> | <i>Zea mays</i> | <i>Trithrinax campestris</i> |
| <i>Ongokea gore</i> | <i>Cannabis sativa</i> | <i>Silene antirrhina</i> |
| <i>Colchicum autumnale</i> | <i>Olea europaea</i> | <i>Cyrilla racemiflora</i> |

Similar numbers of medicinal plant species were detected in Bali (255) and Singapore (253) (Table 4.2). A majority of the medicinal plant species were also common to both islands: 221 general medicinal plant species and 44 Anti-all plant species were common to samples from both Bali and Singapore. A complete list of medicinal plants documented and cataloged in our study, along with the referenced literature, can be found in the supplementary Excel file available with

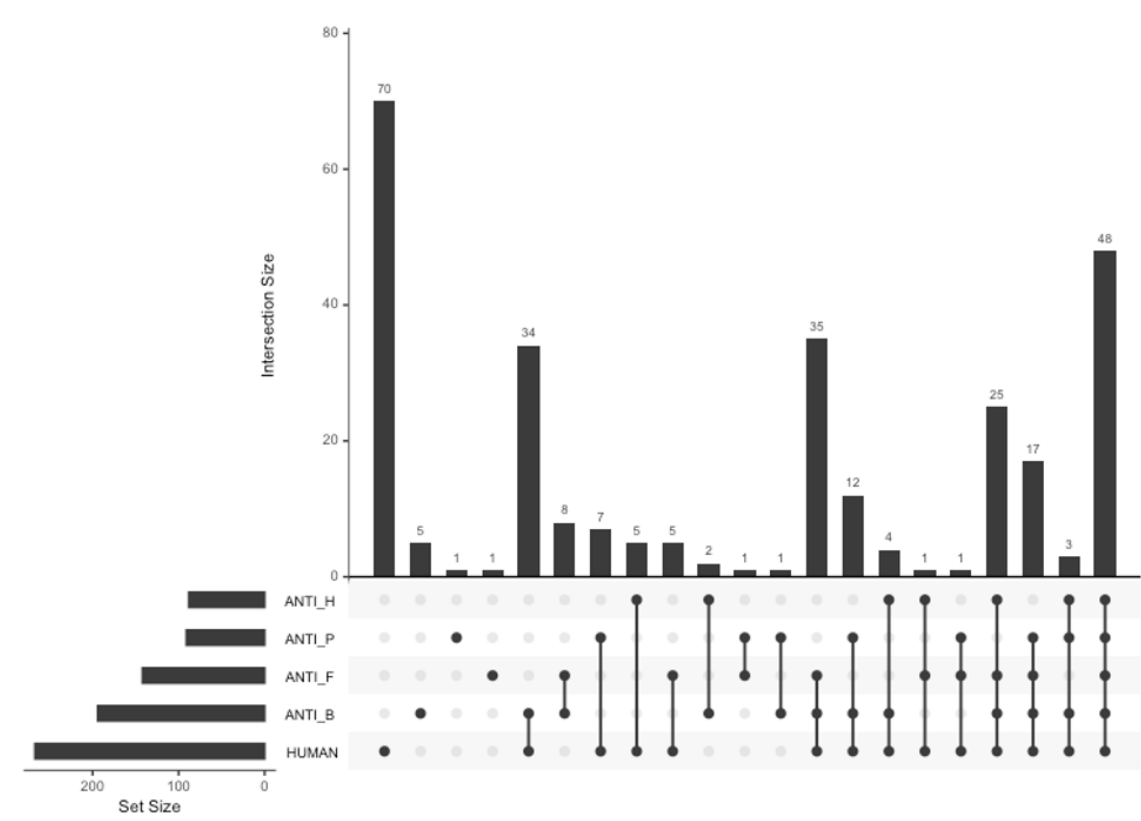


Figure 4.1. Intersections of plants in medicinal categories. The medicinal plant database categorizes plants by their documented medicinal properties (Anti-helminthic, Anti-protozoan, Anti-fungal, Anti-bacterial, and human medicinal use). The far-right column contains plant taxa which have combined medicinal properties, also referred to as "Anti-All" throughout analyses.

this dissertation.

TABLE 4.2

UNIQUE PLANT ITEMS IN MEDICINAL CATEGORIES

| | All Samples | Singapore | Bali |
|---------------------------|--------------------|------------------|-----------------|
| | (N = 127) | (N = 63) | (N = 64) |
| General Medicinal Plants | 286 | 253 | 255 |
| Anti-all Plants | 48 | 46 | 46 |
| Non-medicinal Plant Items | 377 | 301 | 327 |
| Total Plant Diet Items | 663 | 554 | 582 |

4.4.2 Protozoan and Helminth Associations with Medicinal Plants

Selected genera abundance and prevalence are presented in Table 4.3, and a complete list of recovered genera is included in the supplementary excel file available with this dissertation. We recovered 25 total helminth genera in all samples (Bali = 18, Singapore = 19), and 224 protozoan genera (Bali = 198, Singapore = 157).

TABLE 4.3

SELECTED HELMINTH AND PROTOZOAN GENERA RECOVERED IN
18S RRNA DATA SORTED BY ABUNDANCE

| | Abundance (N reads post-filtering) | Prevalence |
|-----------------------|---------------------------------------|------------|
| Helminths | | |
| <i>Haemonchus</i> | 1,559,359 | 56% |
| <i>Strongyloides</i> | 851,455 | 49% |
| <i>Trichinella</i> | 12,796 | 84% |
| Protozoans | | |
| <i>Blastocystis</i> | 29,426,416 | 71% |
| <i>Plasmodium</i> | 14,627,860 | 80% |
| <i>Entamoeba</i> | 6,540,856 | 27% |
| <i>Neobalantidium</i> | 1,280,041 | 10% |

NOTE: A complete list of genera can be accessed in supplementary excel file.

Partial Mantel tests revealed significant relationships between parasite abundance distances and medicinal plant abundance distances (Table 4.4). When examining helminths and protozoans together, there were significant relationships in both islands combined ($r = 0.111$, $p = 0.011$); when separating the two islands, there was only a significant relationship in Bali ($r = 0.152$, $p = 0.018$). When examining helminths and medicinal plants separately, we did not find a relationship. However, when examining protozoans and medicinal plants, the relationships in all samples ($r = 0.123$, $p = 0.004$) and Bali individually ($r = 0.1503$, $p = 0.001$) remained. In Bali, there was also a significant relationship between protozoans and Anti-all plants ($r = 0.150$, $p = 0.032$). Relationships between *Plasmodium* abun-

dance and general medicinal plant abundance were significant for both islands ($r = 0.286$, $p = 0.004$), Singapore ($r = 0.251$, $p = 0.004$), and Bali ($r = 0.156$, $p = 0.007$). Significant relationships were observed when looking more specifically at *Plasmodium* and Anti-all plant among both islands ($r = 0.302$, $p = 0.004$), in Singapore ($r = 0.241$, $p = 0.004$), and in Bali ($r = 0.261$, $p = 0.004$). Non-medicinal plants were not associated with any parasites in our samples.

TABLE 4.4

MEDICINAL AND NON-MEDICINAL PLANT CORRELATIONS TO

PARASITES

| All Samples | | Singapore | | | Bali | | |
|---------------------------------|----------------|----------------|-------------------|----------------|----------------|-------------------|-------------------|
| | <i>r value</i> | <i>p value</i> | <i>Adjusted p</i> | <i>r value</i> | <i>p value</i> | <i>Adjusted p</i> | <i>Adjusted p</i> |
| Protozoans and Helminths | | | | | | | |
| Gen. med. | 0.111 | 0.000 | 0.011 | 0.084 | 0.067 | 1.000 | 0.018 |
| Anti-all | 0.095 | 0.002 | 0.061 | 0.111 | 0.018 | 0.641 | 0.194 |
| Non-med. | 0.048 | 0.023 | 0.839 | 0.148 | 0.002 | 0.061 | 0.187 |
| Helminths | | | | | | | |
| Gen. med. | 0.026 | 0.083 | 1.000 | 0.021 | 0.205 | 1.000 | 1.000 |
| Anti-all | 0.036 | 0.034 | 1.000 | 0.053 | 0.036 | 1.000 | 1.000 |
| Not med. | 0.052 | 0.003 | 0.094 | 0.065 | 0.016 | 0.569 | 1.000 |
| Protozoans | | | | | | | |
| Gen. med. | 0.123 | 0.000 | 0.004 | 0.080 | 0.076 | 1.000 | 0.032 |
| Anti-all | 0.102 | 0.002 | 0.054 | 0.109 | 0.019 | 0.691 | 0.187 |
| Non-med. | 0.033 | 0.074 | 1.000 | 0.127 | 0.005 | 0.162 | 0.166 |

TABLE 4.4 (CONTINUED)

| | All Samples | | | Singapore | | | Bali | | |
|--------------------------|----------------|----------------|-------------------|----------------|----------------|-------------------|----------------|----------------|-------------------|
| | <i>r value</i> | <i>p value</i> | <i>Adjusted p</i> | <i>r value</i> | <i>p value</i> | <i>Adjusted p</i> | <i>r value</i> | <i>p value</i> | <i>Adjusted p</i> |
| <i>Plasmodium</i> | | | | | | | | | |
| Gen. med. | 0.286 | 0.000 | 0.004 | 0.251 | 0.000 | 0.004 | 0.156 | 0.000 | 0.007 |
| Anti-all | 0.302 | 0.000 | 0.004 | 0.241 | 0.000 | 0.004 | 0.261 | 0.000 | 0.004 |
| Non-med. | 0.086 | 0.002 | 0.065 | 0.033 | 0.199 | 1.000 | 0.119 | 0.002 | 0.083 |

NOTE: Protozoans and helminths combined, helminth, protozoan, and *Plasmodium* distance matrices to general medicinal, Anti-all, and non-medicinal plant distance matrices. Significant results (Bonferroni-adjusted $p \leq 0.05$) are bolded.

Plasmodium presence was associated with general medicinal plant abundance in Singapore (Welch's $t = -4.200$, $p = 4.05 \times 10^{-4}$), Bali (Welch's $t = -3.449$, $p = 1.03 \times 10^{-3}$), and all samples combined (Welch's $t = -4.65$, $p = 1.101 \times 10^{-5}$; Figure 4.2A-C). Linear models were not significant between general medicinal plant abundance and *Plasmodium* abundance in any of our samples (Figure 4.2D-F).

We found higher abundance of Anti-all plants when *Plasmodium* was present throughout our samples (all samples: Welch's $t = -5.709$, $p = 8.64 \times 10^{-8}$; Singapore: Welch's $t = -5.273$, $p = 2.02 \times 10^{-6}$; Welch's $t = -2.955$, $p = 4.63 \times 10^{-3}$; Figure 4.3A-C). Linear modeling of Anti-all plant abundance and *Plasmodium* abundance revealed a highly-significant model for both islands combined ($R^2 = 0.304$, $p = 1.78 \times 10^{-11}$; Figure 4.3D), which further underscores the positive relationship detected by the Mantel test. This significant relationship held when examining only Singapore samples ($R^2 = 0.3046$, $p = 2.75 \times 10^{-6}$; Figure 4.3E), but was not observed in Bali samples (Figure 4.3F).

General medicinal plant richness was higher in *Plasmodium* positive samples across our samples (Welch's $t = -6.750$, $p = 3.63 \times 10^{-8}$), and when samples were separated by island (Bali: Welch's $t = -2.846$, $p = 0.01$, Singapore: Welch's $t = -3.914$, $p = 9.71 \times 10^{-3}$; Figure 4.4A-C). *Plasmodium* abundance and general medicinal plant richness have a positive relationship in combined samples and Bali, but not in Singapore (all samples: $R^2 = 0.100$, $p = 3.46 \times 10^{-4}$; Bali: $R^2 = 0.146$, $p = 0.002$; Figure 4.4D-F).

This difference was strengthened when we tested Anti-all plant richness across samples ($t = -9.527$, $p = 5.54 \times 10^{-14}$), and in islands separately (Singapore: $t = -5.429$, $p = 0.001$, Bali: $t = -4.750$, $p = 4.219 \times 10^{-5}$) (Figure 4.5A-C).

We again saw relationship strengthen when we narrowed our analyses to Anti-all plant richness (all samples: $R^2 = 0.115$, $p = 9.369 \times 10^{-5}$; Singapore: $R^2 = 0.010$, $p = 0.01179$; Bali: $R^2 = 0.010$, $p = 0.013$; Figure 4.5D-F). *Plasmodium* read

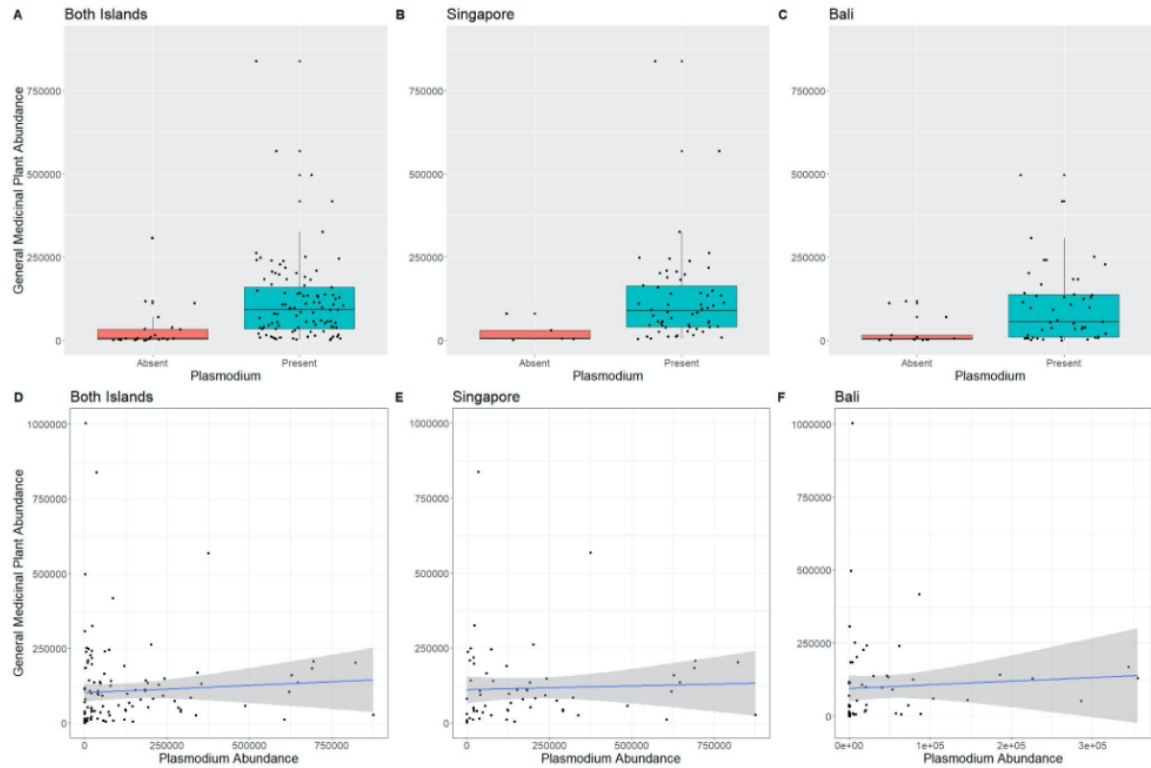


Figure 4.2. General medicinal plant abundance in *Plasmodium*-positive and *Plasmodium*-negative samples. (A-C) Medicinal plants were more abundant when *Plasmodium* was present (Welch's $t = -4.654$, $p = 1.101 \times 10^{-5}$), Singapore (Welch's $t = -4.190$, $p = 4.055 \times 10^{-4}$), and Bali (Welch's $t = -3.449$, $p = 1.028 \times 10^{-3}$). (D-F) General medicinal plant abundance is not associated with *Plasmodium* abundance.

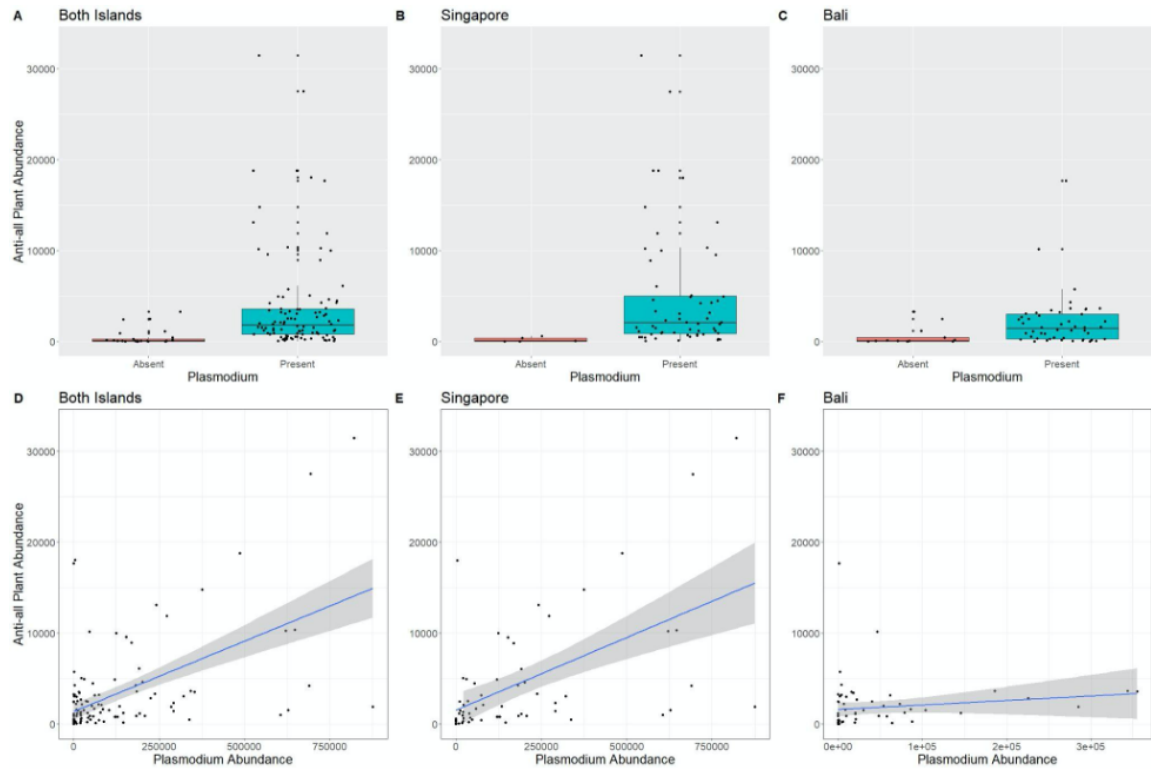


Figure 4.3. Anti-all plant abundance in *Plasmodium*-positive and *Plasmodium*-negative samples. (A-C) Anti-all plants are more abundant in *Plasmodium*-positive across samples (combined samples: Welch's $t = -5.709$, $p = 8.64 \times 10^{-6}$; Singapore: Welch's $t = -5.274$, $p = 2.02 \times 10^{-6}$; Bali: Welch's $t = -2.955$, $p = 4.63 \times 10^{-3}$). (D-F) Anti-all plants are positively associated with *Plasmodium* abundance in all samples and Singapore (combined samples: $R^2 = 0.304$, $p = 1.78 \times 10^{-11}$; Singapore: $R^2 = 0.305$, $p = 2.75 \times 10^{-6}$, but not Bali).

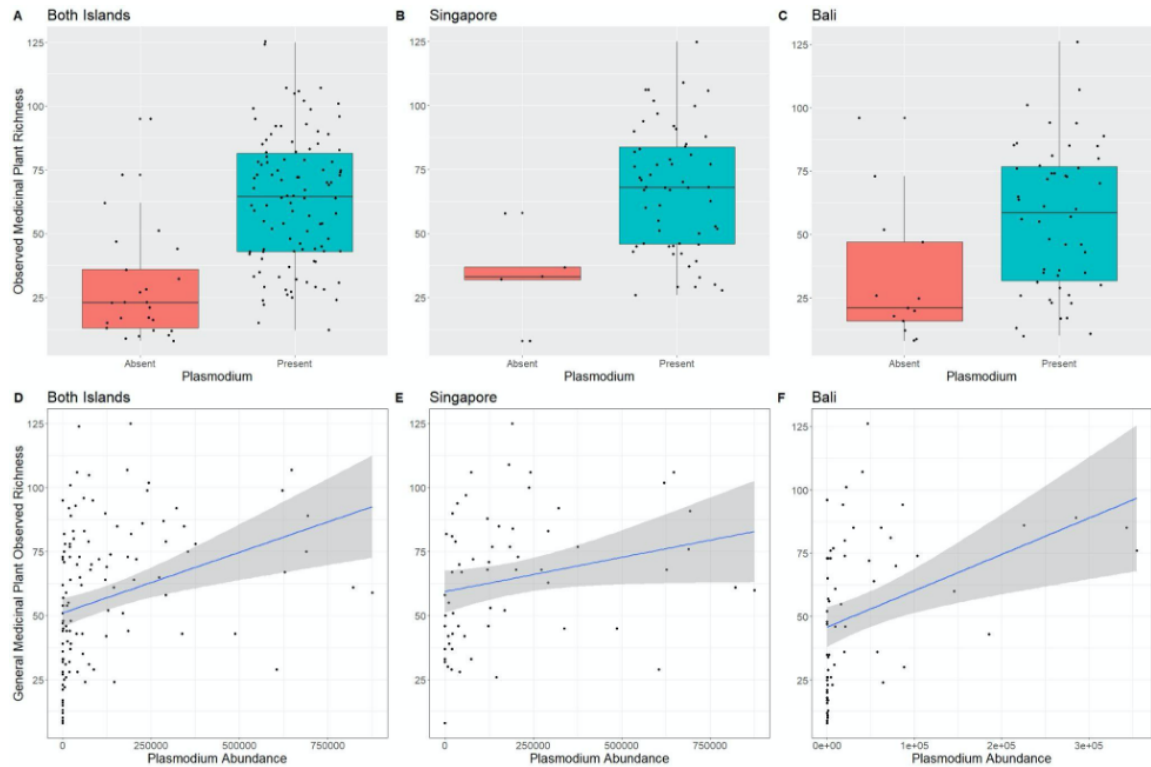


Figure 4.4. General Medicinal plant richness in *Plasmodium*-positive and *Plasmodium*-negative samples. (A-C) Medicinal plant richness is higher in *Plasmodium*-positive samples (all samples: Welch's $t = -6.750$, $p = 3.63 \times 10^{-6}$; Singapore: Welch's $t = -3.914$, $p = 9.71 \times 10^{-3}$; Bali: Welch's $t = -2.846$, $p = 0.010$). (D-F) *Plasmodium* abundance is positively related to medicinal plant richness in combined samples and Bali, but not Singapore (all samples: $R^2 = 0.09774$, $p = 3.455 \times 10^{-4}$; Bali: $R^2 = 0.146$, $p = 0.002$).

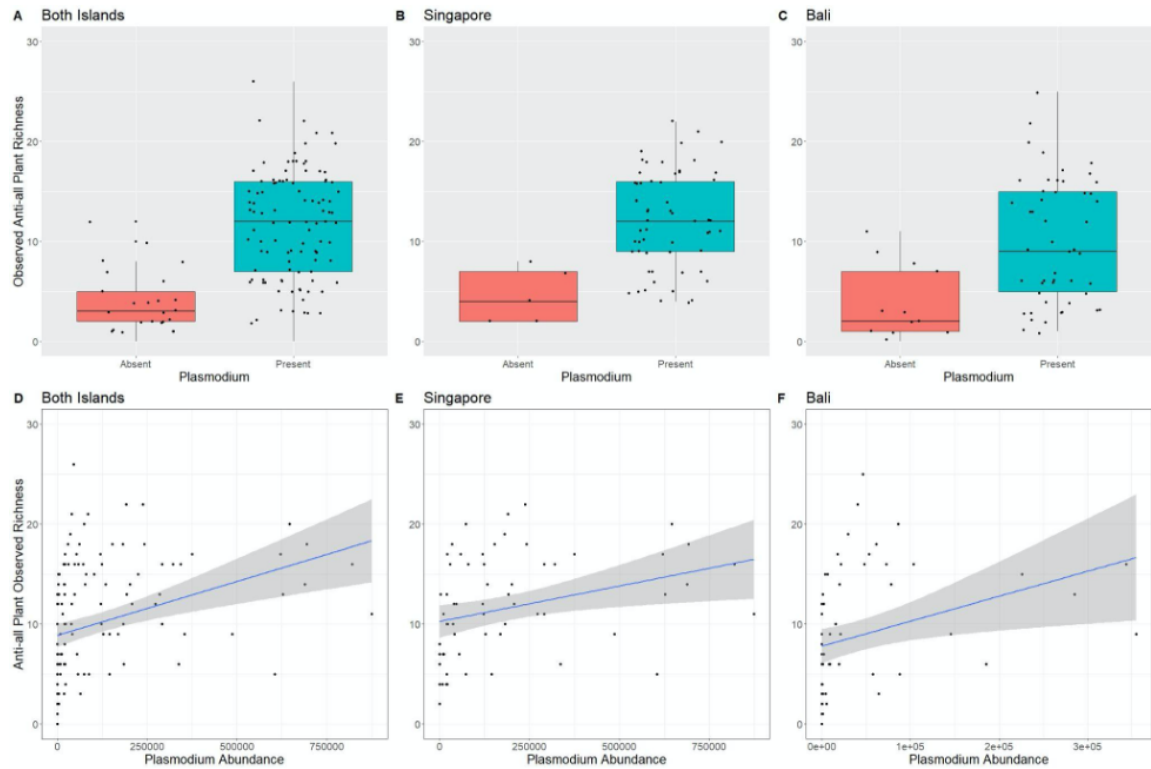


Figure 4.5. Anti-all plant observed richness in *Plasmodium*-positive and *Plasmodium*-negative samples. (A-C) Anti-all plant richness is higher in *Plasmodium*-positive samples across data (all samples: Welch's $t = -9.527$, $p = 5.54 \times 10^{-14}$; Singapore: Welch's $t = -5.429$, $p = 0.001$; Bali: Welch's $t = -4.750$, $p = 4.22 \times 10^{-5}$). (D-F) *Plasmodium* abundance is positively related to Anti-all plant richness across samples (combined samples: $R^2 = 0.115$, $p = 9.37 \times 10^{-5}$; Singapore: $R^2 = 0.100$, $p = 0.01$; Bali $R^2 = 0.098$, $p = 0.01$).

counts were not significantly related to plant read counts.

4.5 Discussion

We designed and constructed a medicinal plant database to investigate our questions centered on the consumption of plants by primates, and share here with the goal of facilitating some of the many challenges encountered when researching wild primate populations. To our knowledge, a resource like this does not exist. In our focus on medicinal plants, we aggregate a rich body of work documented in over 750 publications describing and testing properties of plants. This resource is designed to be able search and filter plants quickly based on medicinal properties and apply this knowledge to address a variety of biological questions. Our database is based entirely from the diet of long-tailed macaques in Singapore and Bali, Indonesia, and may be less inclusive of plants beyond the geographical boundaries of our study area, although we note we placed no restrictions on whether a plant was native to the area. Close to half of the plants we cataloged are medicinal plants and have widely overlapping presence on both islands in our study, making them generally accessible to the macaques we sampled.

Employing this plant database to look at self-medication in long-tailed macaques, we find positive correlations between medicinal plant consumption and parasitic infections. Medicinal plant consumption in macaques on both islands is consistently higher when *Plasmodium* is present, and this relationship is even more pronounced when we focus on Anti-all plants, suggesting that plants with more medicinal properties underlie our findings. Abundance of medicinal plants is a predictor of *Plasmodium*, a connection rooted in our Singapore samples, whereas richness of medicinal plants is driven by Bali samples. Collectively, our results relay new insights into self-medicating behaviors in primates made possible by employing bar-

coding techniques previously underutilized in this field. We find consistent positive correlations between medicinal plant consumption and the presence of *Plasmodium*: these results were seen on both islands and driven primarily by associations between *Plasmodium* and anti-all plant consumption.

We found no association between medicinal plants and helminths, which was somewhat unexpected given the array of previous work focusing self-medication of helminthic infections (Huffman et al., 1996). Helminth detection is generally detected via field assays or microscopy, and is aided by examining undigested rough leaves to slough intestinal walls (Anderson, RC, 1992; Huffman et al., 1996). Because digestion status of plant material would not be decipherable in our study, we cannot speak to the prevalence of this mechanical mechanism for reducing parasite loads.

We hypothesize that Anti-all plants, having documented anti-helminthic, anti-bacterial, anti-protozoan, and anti-fungal properties have greater medicinal value, and are ingested more frequently in macaques exhibiting symptoms from parasite infestations. Although we see the relationship between *Plasmodium* infection and an increase in consumption of medicinal plants, our data is insufficient to test directly if this is related to the onset of illness in macaques, as we did not collect any information regarding symptoms or health status of our sample sources, but this is a promising direction for future research. Another important factor in medicinal plant consumption concerns the ingestion of specific plant structures, e.g., the chewing of bark or, again, the rough surfaces of foliage (Tasdemir et al., 2020). However, secondary compounds of plants underlie the medicinal properties in many cases, and are not necessarily restricted to botanical structures (Hussein and El-Anssary, 2019). *Morus alba*, for example, is the most abundant Anti-all plant in our samples and has medicinal properties in extracts taken from the root bark, leaves, stems, and fruits of the plant (Sohn et al., 2004; Wang et al.,

2012). Water content of ingested plants could have a dilutive effect on secondary compounds (i.e., whether it was fresh or dried), and whether any fermentation occurred, which could affect the concentration and composition of antimicrobial phenolic compounds such as tannins and flavonoids (Septembre-Malaterre et al., 2018).

Landscape differences offer a contrast between Singapore and Bali and are likely to underlie some of the differences seen in our data. Macaque provisioning is widely seen throughout our sample sites in Bali, encouraged by wildlife tourism and religious practices, and the geographic spread of sampling sites in Bali is large compared to that in Singapore (Lane et al., 2011). Conversely in Singapore, the government places restrictions on provisioning which is enforced by fines, although opportunities for obtaining food in the highly urbanized areas abound (Riley et al., 2016). The plant taxa we detected in these islands have a great deal of overlap, but there are likely to be differences in availability to macaques due to landscape management practices and land dedicated to crop farming. Many of the plants in our database are crop plants, such as cotton (*Gossypium*), palm oil (*Elaeis*), corn (*Zea*), and soybean (*Glycine*).

Plasmodium taxa in our data could not be assigned to the species level, and long-tailed macaques are capable of hosting multiple species in Southeast Asia (Lee et al., 2011; Lempang et al., 2022; Zhang et al., 2016). Plant extracts that have been shown to inhibit *Plasmodium* species are numerous and often involve specific plant structures (Bagavan et al., 2011a,b; Bauri et al., 2015). Our work presents the use of barcoding in identifying the associations between medicinal plant consumption and parasite presence, demonstrating the importance of further research into self-medication of macaques infected with *Plasmodium*.

4.6 Conclusions

Our work presents a sizeable collection of plants with documented medicinal properties captured through barcoding which can be used in future investigations into the role of medicinal plants in ecology and self-medication in primates. Our analyses provide support for self-medication behaviors by showing consistent relationships between medicinal plant consumption and parasite presence, greatly increasing insight to these dynamics. *Plasmodium* reservoirs and fluxes are particularly important from infectious disease transmission perspectives, and here we argue that incorporating this knowledge into efforts to preserve biodiversity in wild habitats is essential in allowing what behaviors might exist to remain undisturbed.

CHAPTER 5

BLASTOCYSTIS IS ASSOCIATED WITH SHIFTS IN PROKARYOTIC DIVERSITY AND COMPOSITION IN FREE-RANGING LONG-TAILED MACAQUES

5.1 Abstract

Host-associated microbial communities influence a range of important biological functions of the host, including development, metabolism, and immune defenses. Several factors govern the assemblage of different microbial communities, such as host traits, environment, diet, and interactions with eukaryotic gut parasites – a relationship that is understudied. It has been hypothesized that the influences gut parasites exert on bacterial community richness and diversity occur through ecological patterns such as top-down processes mediated through host immune responses and bottom-up processes governed by competition and resource availability in the gut. *Blastocystis* is one of the most widespread intestinal parasites known, but limited information is available regarding its role in the gut microbiome. To investigate the influence of protozoan presence on microbiomes, we targeted the prokaryotic and protozoan members of the gut microbial communities of long-tailed macaques on the islands of Singapore and Bali using amplicon sequencing, and focusing on the shifts in community structure found in connection with *Blastocystis* presence and abundance. We find *Blastocystis* to drive variation among protozoans in the gut, and see an increase in bacterial alpha diversity associated with *Blastocystis* presence and abundance. We measure sources of

variation on the bacterial community and find that *Blastocystis* presence explains a significant amount of variation in community composition alongside sampling location. We also identify nine Amplicon Sequence Variants (ASVs) associated with the presence of *Blastocystis* in macaques that are also associated with infections in humans. Together, these results present insight into how protozoan infections can influence variation of the gut microbiome of primates, which is essential in understanding how overall ecological dynamics of eukaryotes and prokaryotes in the gut affect host health.

5.2 Background

Microbiomes increasingly provide new directions to understand host health. Functions of immune physiology, gastrointestinal illnesses, and even nervous systems have all been linked to changes in microbiomes (Cho and Blaser, 2012; De Vadder et al., 2018; Sartor, 2008). Gut bacterial communities are quantitatively impacted by fluctuations in the host environment through abiotic factors, seasonal climatic variables, anthropogenic factors, and parasitic exposure (Clayton et al., 2016, 2018; David et al., 2014b; Greene et al., 2021).

Protozoans are widely varied and abundant in wild animals, and frequently considered parasites with detrimental impacts on host health. The pathogenicity of *Blastocystis*, a non-flagellated Stramenopile ubiquitous throughout Metazoa, is debated as it has been associated with infection in asymptomatic and symptomatic patients, and may even contribute to positive health status in hosts (Deng et al., 2021; Nieves-Ramírez et al., 2018; Scanlan et al., 2014; Stensvold and van der Giezen, 2018; Wawrzyniak et al., 2013). Genetic diversity of *Blastocystis* has been investigated to contextualize pathogenicity in humans, but these efforts have not been sufficient to form a consistent conclusion (Stensvold and Clark,

2016). Interactions with bacteria may affect protozoan ability to colonize and proliferate in a host, and pathogenicity of a parasite may be modulated by bacteria (Bär et al., 2015; Burgess et al., 2017; Shukla et al., 2016; Yooseph et al., 2015). Certain protozoans, like *Blastocystis*, have debatable pathogenicity and may be more appropriately considered commensals in certain contexts, acting as mediators of bacterial communities by keeping bacterial populations from becoming too abundant and causing a population imbalance (Mann et al., 2020; Parfrey et al., 2014; Wilcox and Hollocher, 2018). *Blastocystis* infection is frequently associated with an increase in gut bacterial alpha diversity and change in community composition, including the changes of bacterial taxa (Audebert et al., 2016; Even et al., 2021; Nieves-Ramírez et al., 2018). Protozoans have been proposed as ecological drivers in these environments, exerting top-down effects on diversity through consuming bacteria, for example, but these relationships and mechanisms are not well understood (Chabé et al., 2017; Laforest-Lapointe and Arrieta, 2018).

To further investigate these relationships, we collected fecal samples from free-ranging long-tailed macaques (*Macaca fascicularis*) and surveyed the microbiome using combined 16S and 18S rRNA amplicons. Our objective was to detect the extent of variation of *Blastocystis* in relation to other gut parasites and the bacterial community. We aim to elucidate the coinfection dynamics of *Blastocystis* in relation to other protozoans in the gut environment of macaques, and relate this to changes in the prokaryotic microbiome, focusing on shifts in bacterial diversity and community composition, across different taxonomic scales.

5.3 Methods

5.3.1 Sample Collection and DNA Processing

Fecal samples were collected non-invasively from free-ranging long-tailed macaques in Bali, Indonesia (N=45, from nine sites) and Singapore (N=41 from eight sites) using methods described previously (Klegarth et al., 2017; Lane-deGraaf et al., 2014b; Wilcox and Hollocher, 2018). Macaques in these populations commonly reside in edge habitats bordering forested areas and urban environments, and are subsequently exposed to a range of anthropogenic influences which differ drastically between site location (Lane et al., 2011; Sha et al., 2009c).

Total DNA was extracted from each sample using the Qiagen QIAMP Stool Minikit (Qiagen, Hilden, Germany) as per manufacturer instructions. PCR targeted the V4 region of the 16S rRNA gene and the V9 region of the 18S rRNA gene. For 16S rRNA, a 287 bp sequence was amplified from genomic DNA using the following primers (Illumina adaptors in italics): Forward (S-D-Arch-0519-a-S-15, 5' - *TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CAG CMG CCG CGG TAA* -3'), Reverse (S-D-Bact-0785-b-A-18, 5' - *GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GTA CNV GGG TAT CTA ATC C* -3'; (Klindworth et al., 2013; Van Bleijswijk et al., 2015). Methods for amplification of the 18S rRNA gene have been described previously (Wilcox and Hollocher, 2018). Further details for library prep, including proportions of DNA and primer used in PCR master mix available in Appendix A.2.1. The resulting DNA library was sequenced on an Illumina HiSeq 2500 (Illumina, San Diego, CA, USA), using a rapid run with paired-end 250 base pair reads at the University of New Hampshire Hubbard Center for Genome Studies.

5.3.2 Read Processing

Raw sequencing reads were processed via two differing pipelines. To process 16S rRNA gene reads, raw fastq files were imported directly into R and processed using the Dada2 package (Dada2 v 1.18.0, R v 4.0.5) (Callahan et al., 2016; R Core Team, 2020). Amplicon sequence variants (ASVs) produced by this pipeline were then assigned taxonomy using the RDP database [RDP training set 16, formatted from 11.5 database release] (Callahan, 2017). ASVs were filtered to retain the kingdom Bacteria and remove chloroplast/Cyanobacteria ASVs before constructing a count table and importing into the phyloseq package for downstream statistical processing (McMurdie and Holmes, 2013). To process 18S reads, the Usearch pipeline was used as previously described in Wilcox and Hollocher (Wilcox and Hollocher, 2018). Briefly, sequences were quality filtered, dereplicated, and clustered at 97% to form operational taxonomic units (OTUs). All singleton and doubleton OTUs were removed during clustering. Pre-clustered reads were then mapped back onto these OTUs with a 97% identity threshold. Matching reads were then searched against the full Silva 132 non-redundant database using an 80% threshold and constructed into a count table (Quast et al., 2012; Yilmaz et al., 2014). Taxa were consolidated to the genus level and then filtered using a cutoff level of 6, following Cirtwill and Hambäck (2021). Genera were then filtered to isolate protozoan eukaryotes for downstream analysis. The genera *Plasmodium* and *Hepatocystis* were removed from samples as they are blood parasites not known to reside in the gut.

5.3.3 Statistical Analysis

We first inspected the presence and prevalence of *Blastocystis* in our samples. To evaluate *Blastocystis* in the gut environment, and its presence in relation

to other protozoans present in our samples, we began by performing a Principal Coordinates Analysis (PCA) on Hellinger-transformed protozoan taxa using the `prcomp()` function in the `stats` package in R. We then compared the presence and abundance of *Blastocystis* to the prokaryotic microbiome to analyze whether prokaryotic diversity was associated with the presence of *Blastocystis*. To evaluate how alpha diversity reflects the ecological roles *Blastocystis* holds in the gut, we selected three metrics which differentially weight abundant and rare taxa to compare to *Blastocystis* presence and abundance. To balance weight between rare and dominant prokaryotic taxa, we used Shannon diversity, followed by richness (observed taxa) to weight rare taxa equally, and Pielou evenness to emphasize more abundant taxa while still accounting for richness. Wilcoxon Rank Sum tests were used to test prokaryotic alpha diversity when *Blastocystis* is present (`wilcox.test()` function, `stats` package). To see differences between islands, we separated samples by location and repeated tests. Linear models were used (`lm()` function, `stats` package) to assess the linear associations between alpha diversity metrics (again, richness, Shannon diversity, and Pielou evenness).

We used a series of permutational multivariate analysis of variances (PERMANOVA) using the `adonis2()` function in the `vegan` package to test the variation in composition in the bacterial microbiome explained by variables in our dataset (Table 2; Oksanen et al. (2022)). Models tested were euclidean distance matrices of Hellinger-transformed ASVs in samples to represent community composition and *Blastocystis* presence/absence, location (e.g. Bali or Singapore), site (within-island sample collection sites, used in tests comparing only Singapore or only Bali samples), and readcount (sum of reads within samples) (Legendre and Gallagher, 2001). Permutations were set at 9999 for all tests. We used multidimensional scaling of Bray-Curtis distances to ordinate samples based on ASV community composition using the `phyloseq` package (function `ordinate()` and `plot_ordination()`).

To find bacterial taxa which were differentially abundant when *Blastocystis* was present, we used the ALDEx2 package (Fernandes et al., 2013, 2014; Gloor and Reid, 2016). Briefly, center-log ratios (CLR) of each bacterial taxa were calculated and tested using a Dirichlet distribution and Wilcoxon Rank Sum test to identify differentially abundant ASVs. Resulting P-values were adjusted using a Benjamini-Hochberg correction. An α of 0.05 was used throughout this analysis as a threshold for significance.

5.4 Results

From 16S rRNA reads, 5,961 ASVs were recovered, after filtering taxa to exclude non-bacterial ASVs and chloroplast/Cyanobacteria. From 18S rRNA, we recovered 224 protozoan genera (excluding *Plasmodium* and *Hepatocystis*). From these, Singapore contained 146 genera and Bali samples contained 189 genera. *Blastocystis* is present in 62.79% of total samples (Bali 68.89%, Singapore 56.10%). We found *Blastocystis* to be a leading correlating genus for PC1 when all samples were combined, as well as for Singapore and Bali individually when PCAs were performed on separated samples (Figure 5.1, Table 5.1. In all samples, PC1 accounted for 35.9% of variation. In Bali and Singapore, PC1 accounted for 39.1%, and 35.4% of variation, respectively. Other genera correlating with PC1 were *Entamoeba*, *Neobalantidium*, and *Adelina*.

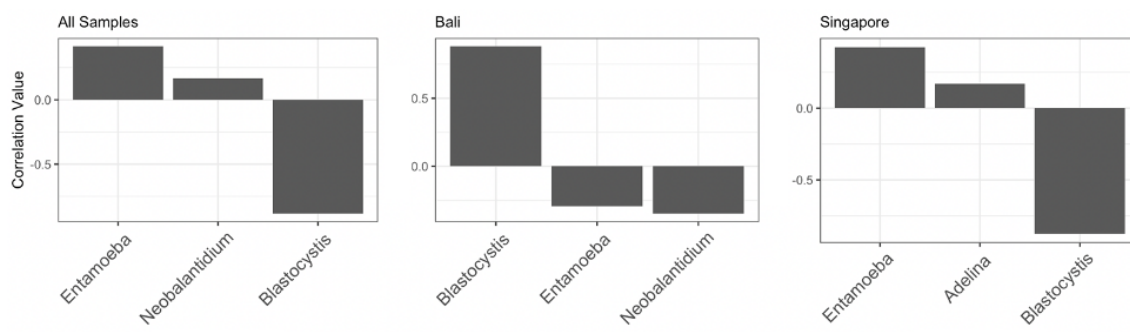


Figure 5.1. Top correlating taxa (loading scores > ± 0.1). *Blastocystis* is a leading taxon for PC1 in all samples combined, and Singapore and Bali individually. Taxa were Hellinger transformed before PCA.

TABLE 5.1

COMPONENT VARIATION AND LOADING SCORES OF
CORRELATING PROTOZOA

| | PC1 | PC2 | PC3 |
|------------------------|--------|--------|--------|
| All Samples | | | |
| Standard Deviance | 0.441 | 0.333 | 0.268 |
| Proportion of Variance | 0.358 | 0.204 | 0.132 |
| Cumulative Variance | 0.358 | 0.561 | 0.694 |
| Loading Scores | | | |
| <i>Entamoeba</i> | 0.416 | 0.848 | -0.231 |
| <i>Neobalantidium</i> | 0.168 | -0.375 | -0.877 |
| <i>Blastocystis</i> | -0.884 | 0.297 | -0.237 |
| <i>Vannella</i> | | | 0.160 |
| <i>Acanthamoeba</i> | | | 0.127 |
| <i>Adelina</i> | | | 0.118 |

TABLE 5.1 (CONTINUED)

| | PC1 | PC2 | PC3 |
|------------------------|------------|------------|------------|
| <i>Bresslaua</i> | | | 0.104 |
| Singapore | | | |
| Standard Deviance | 0.436 | 0.335 | 0.237 |
| Proportion of Variance | 0.354 | 0.209 | 0.105 |
| Cumulative Variance | 0.354 | 0.562 | 0.667 |
| Loading Scores | | | |
| <i>Entamoeba</i> | 0.365 | -0.897 | |
| <i>Adelina</i> | 0.170 | 0.150 | -0.102 |
| <i>Trichinella</i> | 0.169 | 0.123 | |
| <i>Blastocystis</i> | -0.885 | -0.270 | |
| <i>Heligmosomoides</i> | | 0.106 | |
| <i>Phytomonas</i> | | 0.169 | |
| <i>Acanthamoeba</i> | | 0.138 | -0.220 |
| <i>Neobalantidium</i> | | | 0.942 |
| <i>Balantidium</i> | | | 0.124 |
| Bali | | | |
| Standard Deviation | 0.459 | 0.358 | 0.279 |
| Proportion of Variance | 0.391 | 0.238 | 0.144 |
| Cumulative Variance | 0.391 | 0.628 | 0.772 |
| Loading Scores | | | |
| <i>Blastocystis</i> | 0.882 | | 0.376 |
| <i>Entamoeba</i> | -0.293 | -0.801 | 0.447 |
| <i>Neobalantidium</i> | -0.348 | 0.581 | 0.680 |
| <i>Neobodo</i> | | | -0.114 |

TABLE 5.1 (CONTINUED)

| | PC1 | PC2 | PC3 |
|------------------------|-----|-----|--------|
| <i>Geochelone</i> | | | -0.118 |
| <i>Parabodo</i> | | | -0.125 |
| <i>Heligmosomoides</i> | | | -0.148 |
| <i>Vannella</i> | | | -0.228 |

NOTE: First three components and taxa correlating $> \pm 0.1$ shown.

We observed higher richness and Shannon diversity in samples when *Blastocystis* was present (richness: $W = 433.5$, $p < 0.001$, Shannon: $W = 574$, $p = 0.010$), but not evenness ($W = 666$, $p = 0.078$) (Figure 5.2A-C). In Singapore, richness was higher when *Blastocystis* was present ($W = 115.5$, $p = 0.017$), whereas Shannon diversity and evenness showed no differences (Shannon: $W = 151$, $p = 0.146$, evenness: $W = 168$, $p = 0.316$). In Bali, richness and Shannon diversity were higher when *Blastocystis* was present (richness: $W = 126$, $p = 0.027$; Shannon: $W = 134$, $p = 0.042$), and evenness was not different ($W = 164$, $p = 0.201$).

Linear modeling of *Blastocystis* presence and Shannon diversity of prokaryotes showed a significant, but weak, positive relationship in all samples combined ($p = 0.02$, $R^2 = 0.046$), and a stronger relationship when Singapore was analyzed separately ($p < 0.001$, $R^2 = 0.236$; Figure 5.3A). We saw similar relationships with evenness (all samples: $p = 0.03$, $R^2 = 0.043$; Singapore: $p = 0.002$, $R^2 = 0.201$; Figure 5.3B). Bali samples did not show a significant relationship between *Blastocystis* presence and Shannon diversity, nor did we see an effect when testing evenness in prokaryotes (Shannon: $p = 0.6$; evenness: $p = 0.6$). Prokaryotic richness increased with *Blastocystis* abundance in Singapore, but this was not

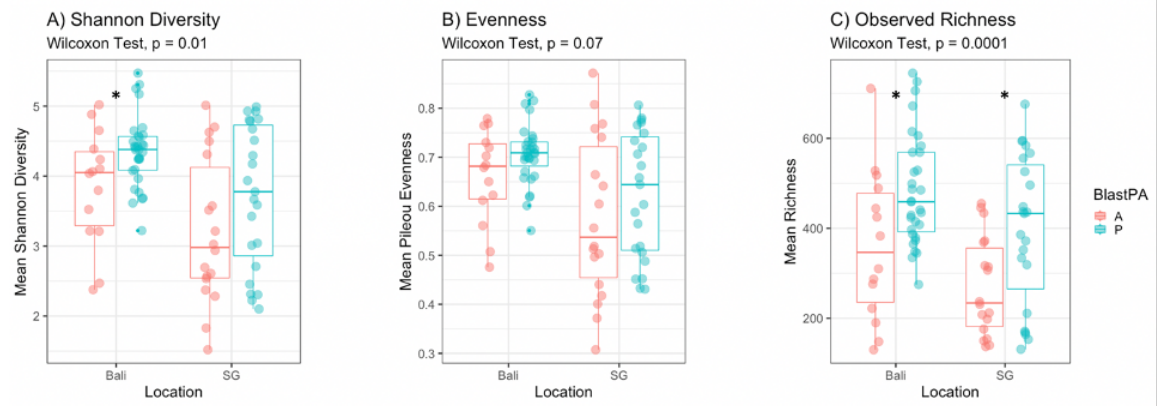


Figure 5.2. Boxplots of alpha diversity metrics compared to *Blastocystis* presence (P) and absence (A). Prokaryotic alpha diversity is significantly higher in the presence of *Blastocystis* using Shannon Diversity (A) and richness (C) as metrics. Evenness (B) does not differ in the presence of *Blastocystis*. P-values shown in title represents test from all samples combined. In Singapore, Shannon Diversity and richness are significantly higher when *Blastocystis* is present. In Bali, Shannon Diversity and richness are not different when *Blastocystis* is present, while evenness is higher. Asterisks denote significance from Wilcoxon-Rank Sum tests controlled for location.

seen in Bali or in combined samples (Singapore: $p = 0.06$, $R^2 = 0.024$; Bali: $p = 0.79$, combined samples: $p = 0.08$; Figure 5.3C).

When testing the community composition of all samples, location explained 7.1% of variation, followed by *Blastocystis* presence (3.1 %; Table 5.2). We then used PERMANOVAs to test within-location sampling site effect along with *Blastocystis* presence and read count. Results in Singapore and Bali are similar: site has the greatest effect (Singapore: 33.5%, Bali: 32.2%), followed by *Blastocystis* presence (Singapore: 4.0 %, Bali: 5.4%). In Bali alone, read count had an effect of 2.7%. Multidimensional scaling displayed a wider Euclidean distribution of samples containing *Blastocystis*, and neither *Blastocystis* presence nor location separated samples without overlap (Figure 5.4). Axis 1 explained 14.0 % of variation and axis 2 explained 6.4 %.

TABLE 5.2

PERMANOVA: VARIATION IN COMMUNITY COMPOSITION

| | Df | SumOfSqs | R ² | F | Pr(>F) |
|-------------------------------|----|----------|----------------|-------|-------------|
| All Samples | | | | | |
| Blastocystis Presence/Absence | 1 | 2.083 | 0.031 | 2.902 | 2.00x10-04 |
| Readcount | 1 | 0.738 | 0.011 | 1.028 | 0.366 |
| Location | 1 | 4.711 | 0.071 | 6.564 | 1.00 x10-04 |
| Residual | 82 | 58.852 | 0.887 | | |
| Total | 85 | 66.384 | 1.000 | | |
| Singapore | | | | | |
| Blastocystis Presence/Absence | 1 | 1.202 | 0.040 | 2.039 | 0.011 |
| Readcount | 1 | 0.592 | 0.020 | 1.005 | 0.4203 |
| Site | 7 | 10.117 | 0.335 | 2.452 | 1.00 x10-04 |

TABLE 5.2 (CONTINUED)

| | Df | SumOfSqs | R² | F | Pr(>F) |
|-------------------------------|-----------|-----------------|----------------------|----------|------------------|
| Residual | 31 | 18.274 | 0.605 | | |
| Total | 40 | 30.185 | 1.000 | | |
| Bali | | | | | |
| Blastocystis Presence/Absence | 1 | 1.686 | 0.054 | 3.079 | 2.00 x10-04 |
| Readcount | 1 | 0.844 | 0.027 | 1.541 | 0.028 |
| Site | 8 | 10.048 | 0.322 | 2.294 | 1.00 x10-04 |
| Residual | 34 | 18.613 | 0.597 | | |
| Total | 44 | 31.191 | 1.000 | | |

Eight ASVs were more abundant when *Blastocystis* was present and none was abundant in the absence of *Blastocystis* (Table 5.3). No differentially abundant taxa were recovered from Bali or Singapore samples when analyzed separately. Six taxa that were more abundant in *Blastocystis*-positive samples were in the family *Ruminococcaceae*, and in the genera *Succinivibrio*, *Sporobacter*, and *Oscillibacter*. One taxa in this family could not be assigned to the genus level. Other two taxa that were more abundant in the presence of *Blastocystis* were in the family Porphyromonadaceae (genus not assigned) and the family Succinivibrionaceae (genus *Succinivibrio*).

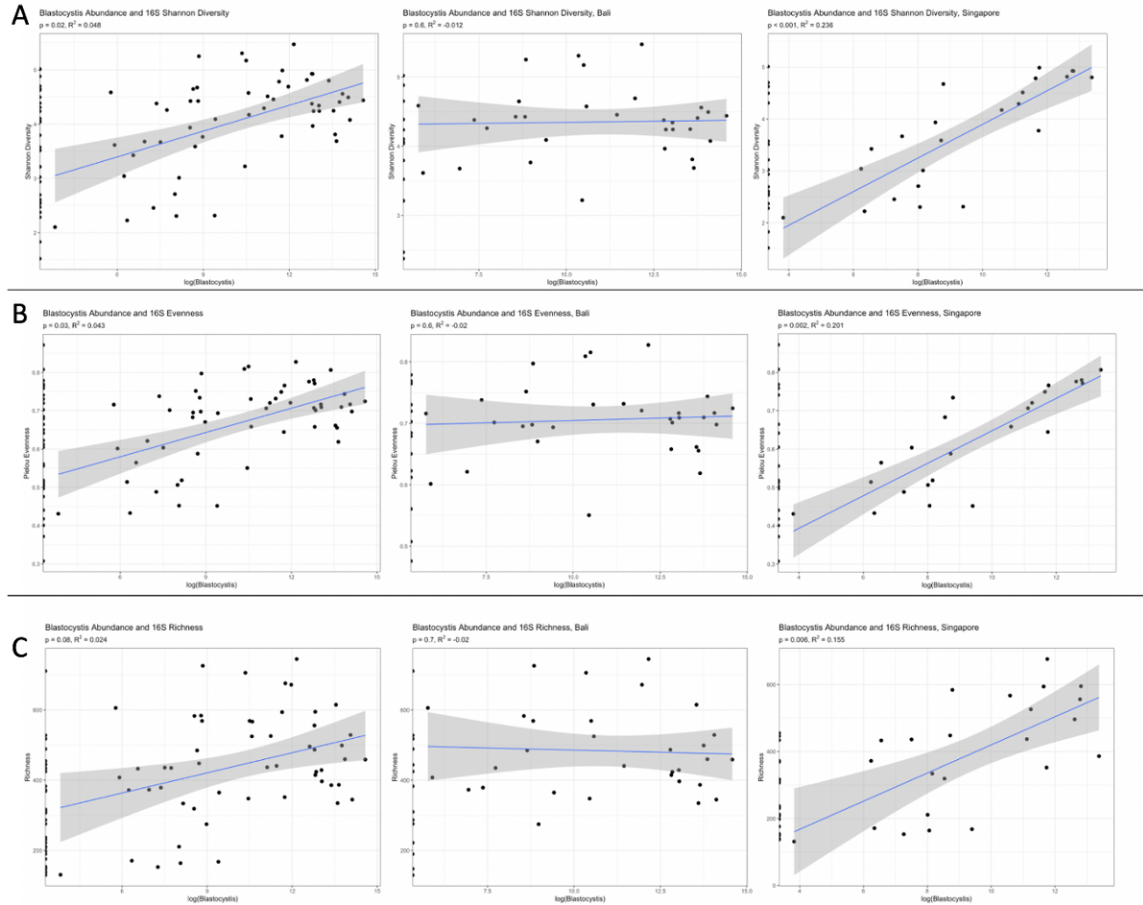


Figure 5.3. Prokaryotic alpha diversity metric relationships with *Blastocystis* abundance. A) Shannon diversity of prokaryotes increases as *Blastocystis* abundance increases in all samples and with Singapore alone. B) Pielou evenness of prokaryotes is significantly associated with *Blastocystis* abundance in all samples and in Singapore alone. C) Richness, or observed taxa, is significantly associated with *Blastocystis* abundance in Singapore but not in all samples or Bali. *Blastocystis* displayed on a logarithmic scale for visibility of samples.

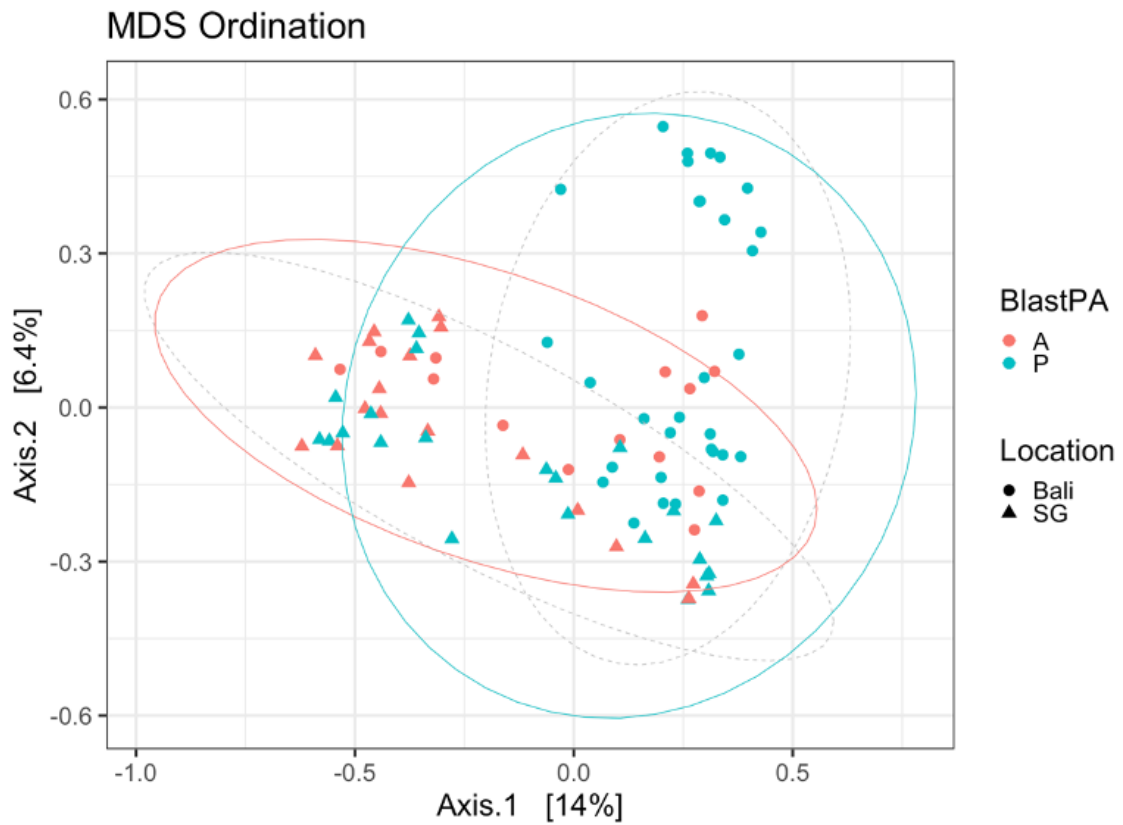


Figure 5.4. Multidimensional scaling using Euclidean distances. Ellipses group samples by *Blastocystis* presence (represented by colors) and location (grey dashed lines)

TABLE 5.3
DIFFERENTIALLY ABUNDANT BACTERIAL TAXA IN
BLASTOCYSTIS-POSITIVE SAMPLES.

| ASV | Taxonomy | Median CLR Absence | Median CLR Presence | Effect | P-val | Adjusted P-val |
|--------|-----------------------|--------------------------|---------------------------|--------|--------|-------------------|
| ASV_10 | K: Bacteria | | | | | |
| | P: Bacteroidetes | | | | | |
| | C: Bacteroidia | 5.6379 | 11.5672 | 0.6156 | 0.0001 | 0.0428 |
| | O: Bacteroidales | | | | | |
| | F: Porphyromonadaceae | | | | | |
| | G: NA | | | | | |

TABLE 5.3 (CONTINUED)

Table 5.3 continued from previous page

| ASV | Taxonomy | Median CLR Absence | Median CLR Presence | Effect - | P-val | Adjusted P-val |
|--------|------------------------|--------------------------|---------------------------|-------------|--------|-------------------|
| ASV_13 | K: Bacteria | | | | | |
| | P: Proteobacteria | | | | | |
| | C: Gammaproteobacteria | 5.9142 | 10.4479 | 0.6133 | 0.0001 | 0.0474 |
| | O: Aeromonadales | | | | | |
| | F: Succinivibrionaceae | | | | | |
| | G: Succinivibrio | | | | | |
| ASV_83 | K: Bacteria | | | | | |
| | P: Firmicutes | | | | | |
| | C: Clostridia | 2.1051 | 8.9813 | 0.6814 | 0.0001 | 0.0443 |
| | O: Clostridiales | | | | | |
| | F: Ruminococcaceae | | | | | |
| | G: Faecalibacterium | | | | | |

TABLE 5.3 (CONTINUED)

Table 5.3 continued from previous page

| ASV | Taxonomy | Median CLR Absence | Median CLR Presence | Effect - | P-val | Adjusted P-val |
|--------|--------------------|--------------------------|---------------------------|-------------|--------|-------------------|
| ASV_93 | K: Bacteria | | | | | |
| | P: Firmicutes | | | | | |
| | C: Clostridia | 3.3448 | 9.9210 | 0.6217 | 0.0000 | 0.0281 |
| | O: Clostridiales | | | | | |
| | F: Ruminococcaceae | | | | | |
| | G: Sporobacter | | | | | |
| ASV_95 | K: Bacteria | | | | | |
| | P: Firmicutes | | | | | |
| | C: Clostridia | 5.0875 | 9.5794 | 0.7063 | 0.0000 | 0.0239 |
| | O: Clostridiales | | | | | |
| | F: Ruminococcaceae | | | | | |
| | G: Sporobacter | | | | | |

TABLE 5.3 (CONTINUED)

Table 5.3 continued from previous page

| ASV | Taxonomy | Median CLR Absence | Median CLR Presence | Effect - | P-val | Adjusted P-val |
|---------|--------------------|--------------------------|---------------------------|-------------|--------|-------------------|
| ASV_140 | K: Bacteria | | | | | |
| | P: Firmicutes | | | | | |
| | C: Clostridia | 5.3740 | 9.5598 | 0.6310 | 0.0000 | 0.0234 |
| | F: Clostridiales | | | | | |
| | G: Ruminococcaceae | | | | | |
| | S: Sporobacter | | | | | |
| ASV_248 | K: Bacteria | | | | | |
| | P: Firmicutes | | | | | |
| | C: Clostridia | 1.2641 | 8.4892 | 0.7581 | 0.0000 | 0.0258 |
| | O: Clostridiales | | | | | |
| | F: Ruminococcaceae | | | | | |
| | G: NA | | | | | |

TABLE 5.3 (CONTINUED)

Table 5.3 continued from previous page

| ASV | Taxonomy | Median CLR | Median CLR | Effect | P-val | Adjusted P-val |
|---------|--------------------|---------------|---------------|--------|--------|-------------------|
| | | Absence | Presence | - | | |
| | K: Bacteria | | | | | |
| | P: Firmicutes | | | | | |
| | C: Clostridia | 1.2054 | 7.9453 | 0.7542 | 0.0000 | 0.0240 |
| ASV_309 | O: Clostridiales | | | | | |
| | F: Ruminococcaceae | | | | | |
| | G: Oscillibacter | | | | | |

NOTE: Taxonomic level abbreviations precede names (e.g., K represents kingdom). CLR values show median value of ASV in condition. P-value and Adjusted P-value were calculated using Wilcoxon-Rank Sum tests (Benjamini-Hochberg correction). All taxa listed are more abundant in *Blastocystis*-positive samples.

5.5 Discussion

Blastocystis has a high prevalence in southeast Asia and across primates collectively, including in long-tailed macaques studied previously (Lane et al., 2011; Vaisusuk et al., 2018; Wilcox and Hollocher, 2018), which was further confirmed here. In addition, our analysis reveals *Blastocystis* driving variation in the protozoan and prokaryotic communities residing in the gut of long-tailed macaques. The presence of *Blastocystis* is tied to higher alpha diversity across our samples, and is associated with differentially abundant ASVs. The presence of *Blastocystis* is also tied to the abundance of other protozoans present in the gut and we demonstrate a negative correlation between *Blastocystis* to *Entamoeba* abundance. This points to a resource-based competition relationship between these two parasites, given the similar niches these parasites occupy in the gut environment (Blessmann et al., 2003; Iebba et al., 2016). Competition between *Blastocystis* and *Entamoeba* in the context of the gut bacteria has been evidenced in humans, but is largely understudied in other groups (Alzate et al., 2020; Even et al., 2021). One species of *Entamoeba*, *E. histolytica*, exhibits pathogenic properties in humans and is present in Southeast Asia (Mahmud et al., 2013). Although our data does not have sufficient resolution to identify *Entamoeba* to the species level or evaluate health effects, future research directions should include a focus on competition-based protozoan interactions between these parasites and their effect on the microbiome.

Bacterial diversity is higher in *Blastocystis*-positive samples across our samples using diversity metrics which emphasize rare taxa, supporting similar results across the literature (Andersen and Stensvold, 2016; Audebert et al., 2016; Forsell et al., 2017). Phagocytic behavior has been observed in *Blastocystis*, demonstrating its potential to exert top-down effects on the gut prokaryotic community which result in an increase in prokaryotic diversity (Dunn et al., 1989; Singh et al., 1995; Tan and Suresh, 2006). Our data show no difference in evenness between *Blas-*

tocystis-positive and *Blastocystis*-negative samples, but we did see evenness increase with *Blastocystis* abundance, a result driven by Singapore samples. Linear relationships were stronger in all measures of alpha diversity in Singapore, and non-existent in Bali, implying the responsiveness of these interactions are affected by differences in environment.

Bacterial community composition relays stark differences between gut microbiomes of Bali and Singapore (see Chapter 3). In the present study, location effect across islands and sampling site effect within islands account for sizeable variation. Environmental factors, such as seasonal fluctuations, diet, and captivity, on the prokaryotic microbiome composition are known sources of variation in free-ranging primates, and this effect in our samples is not unexpected (Clayton et al., 2018; Hicks et al., 2018). Despite this, we were still able to observe significant variation correlating with *Blastocystis*. *Blastocystis* presence explained more variation in composition in Bali, even though we did not detect a significant linear relationship between abundance and prokaryotic diversity. Ellipses enclosing samples in Figure 5.4 show greater dissimilarity of community composition when *Blastocystis* is present, which further demonstrates the increase in variation in bacterial communities residing with this protozoan.

Differentially abundant taxa in *Blastocystis* positive samples echo results in similar analyses and have properties associated with inflammation and fiber digestion. *Faecalibacterium* is associated with *Blastocystis* colonization and has anti-inflammatory properties in the gut environment (Audebert et al., 2016; Kodio et al., 2019; Sokol et al., 2008). *Sporobacter* and *Oscillibacter* are also associated with *Blastocystis*-positive samples, an association also seen in humans harboring with *Blastocystis* (Audebert et al., 2016). Differentially abundant taxa in *Blastocystis*-positive samples in our data (*Faecalibacterium*, *Succinivibrio*, and *Oscillibacter*) overlap with recent results from Stensvold et al. (2022). Collectively, our results

suggest that the response of the prokaryotic gut microbiome to *Blastocystis* infection in macaques reflects trends seen in human trials.

5.6 Conclusions

We elucidate the effects of *Blastocystis* on the protozoan and bacterial microbiome in free-ranging macaques in Southeast Asia. We show *Blastocystis* driving variation in the protozoans present in our samples, and that the presence of *Blastocystis* is associated with a consistently measurable increase in diversity. Our results show the effect of *Blastocystis* on community composition of prokaryotes in free-ranging macaques, and identify individual bacterial taxa which are differentially abundant in the presence of *Blastocystis*. These results support the role of *Blastocystis* as an ecological driver in the gut microbiome environment of primates, which underscores the importance of including protozoans in future microbiome research.

APPENDIX A

APPENDIX

The following text, figures, and tables are referenced throughout this dissertation, and provide supplementary information to the methods and analysis provided.

A.1 Chapter 2 Supplementary Materials

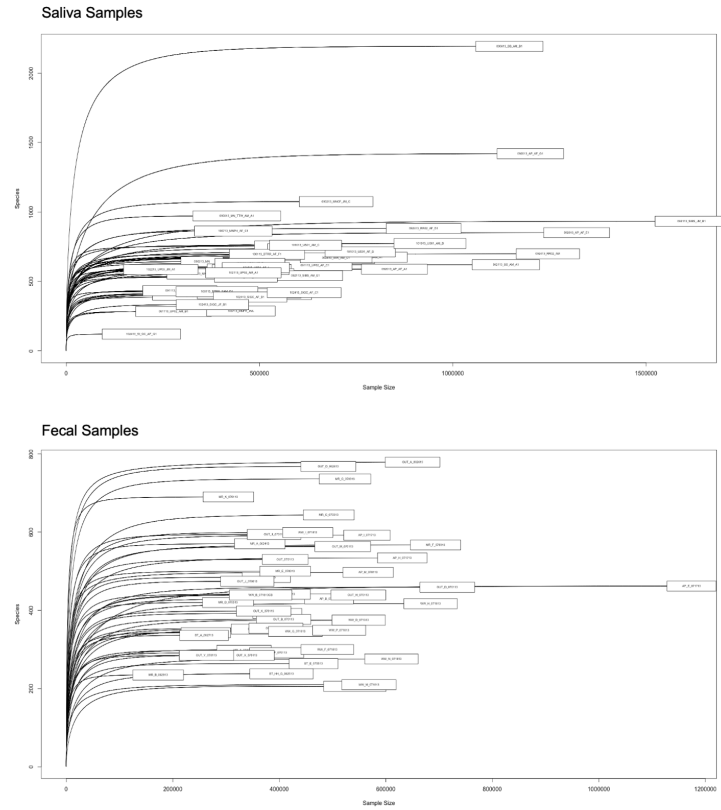


Figure A.1. Rarefaction curves for saliva samples (top) and fecal samples (bottom). Curves were created after filtering non-prokaryotic ASVs from reads, all samples show plateau before reaching read depth.

TABLE A.1

MEAN ALPHA DIVERSITY METRICS AMONG ALL SITES

| Site | Body Site Sample | N | Mean \pm SD Evenness (Pileau) | Mean \pm SD Stability | Mean \pm SD Richness |
|------|---------------------|----|---------------------------------------|-------------------------|---------------------------|
| AP | saliva | 3 | 0.6 \pm 0.01 | 0.13 \pm 0.03 | 953 \pm 425.28 |
| | feces | 9 | 0.51 \pm 0.08 | 0.38 \pm 0.08 | 463.78 \pm 74.46 |
| BB | saliva | 3 | 0.7 \pm 0.05 | 0.11 \pm 0.02 | 1204.67 \pm 857.08 |
| BT | saliva | 7 | 0.68 \pm 0.05 | 0.16 \pm 0.07 | 663.57 \pm 137.06 |
| | feces | 4 | 0.43 \pm 0.13 | 0.57 \pm 0.18 | 261 \pm 55.34 |
| MR | saliva | 8 | 0.66 \pm 0.05 | 0.19 \pm 0.06 | 672.13 \pm 276.27 |
| | feces | 9 | 0.58 \pm 0.13 | 0.37 \pm 0.13 | 517.78 \pm 171.94 |
| OUT | feces | 15 | 0.49 \pm 0.12 | 0.34 \pm 0.15 | 463.27 \pm 158.94 |
| SIBS | saliva | 8 | 0.65 \pm 0.04 | 0.1 \pm 0.04 | 577.13 \pm 185.07 |
| SIGC | saliva | 4 | 0.5 \pm 0.03 | 0.49 \pm 0.27 | 314.5 \pm 134.55 |
| UP | saliva | 6 | 0.62 \pm 0.1 | 0.24 \pm 0.15 | 501.67 \pm 124.09 |
| US | saliva | 5 | 0.71 \pm 0.08 | 0.27 \pm 0.09 | 695 \pm 78 |

TABLE A.1 (CONTINUED)

| Site | Body Site Sample | N | Mean \pm SD Evenness (Pileau) | Mean \pm SD Stability | Mean \pm SD Richness |
|------|---------------------|---|---------------------------------------|-------------------------|---------------------------|
| WW | saliva | 2 | 0.73 \pm 0.04 | 0.37 \pm 0.08 | 629.5 \pm 57.28 |
| | feces | 9 | 0.46 \pm 0.06 | 0.43 \pm 0.06 | 367.89 \pm 111.66 |

A.2 Chapter 3 Supplementary Materials

A.2.1 16S 18S Supplementary Methods

A.2.1.1 Library prep for Illumina sequencing

The 16S hypervariable V4 region is amplified with Prk-V4-F and Prk-V4-R primer pairs. A combination of two forward primers and one reverse primer are used to amplify the 18S hypervariable V9 region. The 18S forward primers target the same region; however, 1380F is 9nt upstream of 1389F, allowing for a slightly longer amplicon with the 1380F-1510R primer combination. All amplicons are created with a single primer pair in separate, parallel reaction. Multiplex PCR strategy is not applied here. After amplification, the amplicons are mixed and then dual-indexed by PCR in parallel. Indexed libraries are combined into a single mixture and sequenced on a single lane of Illumina HiSeq 2500 Rapid Run, paired-end with 250 x 250 nt read length and 8 x 8 nt index read length. PhiX is added to diversify the library's base composition to maintain base call quality. Illumina sequencing services are outsourced to the Hubbard Center for Genomics, Sequencing Core Facility at University of New Hampshire.

A.2.1.2 18S Ribosomal RNA Gene Information

The 18S ribosomal RNA gene V9 region primer sequences, 1380F (5'-NNN NNC CCT GCC HTT TGT ACA CAC-3'), 1389F (5'-NNN NNT TGT ACA CAC CGC CC-3'), and 1510R (5'-CCT TCY GCA GGT TCA CCT AC-3'). Full sequences listed in Table A.2, Illumina overhang sequence is identified with lower case letters and gene-specific sequence with uppercase letters. Information describing PCR mix and thermocycler protocols is described in Tables A.3, A.4, and A.5.

TABLE A.2

18S RIBOSOMAL RNA GENE PRIMER INFORMATION

| Primer Name | Primer Sequence (5' – 3') |
|---------------------------|---|
| Illumina overhang + 1380F | tcgtcggcagcgtcagatgtgtataagagacagNNNNNCCCTGCCHTTTGTACACAC |
| Illumina overhang + 1389F | tcgtcggcagcgtcagatgtgtataagagacagNNNNNTTGTACACACCGCCC |
| Illumina overhang + 1509R | gtctcgtgggctcggagatgtgtataagagacagCCTTCYGCAGGTTCACCTAC |

TABLE A.3

18S RIBOSOMAL RNA GENE PCR REACTION FOR BOTH 1380F AND
1389F WITH 1510R

| Component | Amount per reaction | Final Concentration |
|----------------------------------|----------------------------|----------------------------|
| DNA template, undiluted | 2.5 ml | unknown |
| Forward Primer, 10 mM | 5.0 ml | 2 mM |
| Reverse Primer, 10 mM | 5.0 ml | 2 mM |
| Kapa HiFi HotStart Ready Mix, 2X | 12.5 ml | 1X |

A.2.1.3 16S Ribosomal RNA Gene Information

The 16S ribosomal RNA gene V4 region primer sequences, S-D-Arch-0519-a-S-15 (5'-CAG CMG CCG CGG TAA-3') and S-D-Bact-0785-b-A-18 (5'-TAC NVG GGT ATC TAA TCC-3'). The primers amplify the 16S ribosomal RNA gene V4 region from position 519 to 802. The primers were renamed from S-D-Arch-0519-a-S-15 to "PrkV4-F" and from S-D-Bact-0785-b-A-18 to "PrkV4-R". Full sequences listed in Table A.6, Illumina overhang sequence is identified with lower case letters and gene-specific sequence with uppercase letters. Information describing PCR mix and thermocycler protocols is described in Tables A.7 and A.8.

TABLE A.4

18S RIBOSOMAL RNA GENE PCR THERMOCYCLER CONDITIONS
FOR PRIMERS 1380F AND 1510R

| Step | Temperature | Time | |
|------------------|-------------|------------|-------------|
| Initial Denature | 95 C | 3 minutes | — |
| Denature | 98C | 20 seconds | x 10 cycles |
| Annealing | 70C* | 15 seconds | |
| Extension | 72C | 15 seconds | |
| Denature | 98C | 20 seconds | x 25 cycles |
| Annealing | 65C | 15 seconds | |
| Extension | 72C | 15 seconds | |
| Final Extension | 72C | 5 minutes | — |

*annealing temperature decreases 0.5 C per cycle

TABLE A.5

18S RIBOSOMAL RNA GENE PCR THERMOCYCLER CONDITIONS
FOR PRIMERS 1389F AND 1510R

| Step | Temperature | Time | |
|------------------|-------------|------------|---------------|
| Initial Denature | 95 C | 3 minutes | — |
| Denature | 98 C | 20 seconds | 3*x 35 cycles |
| Annealing | 65 C | 15 seconds | |
| Extension | 72 C | 15 seconds | |
| Final Extension | 72 C | 5 minutes | — |

TABLE A.6

16S RIBOSOMAL RNA GENE PRIMER INFORMATION

| Primer Name | Primer Sequence (5' – 3') |
|-----------------------------|--|
| Illumina overhang + PRKV4-F | tcgtcggcagcgtcagatgtgtataagagacagCAGCMGCCGCGGTAA |
| Illumina overhang + PRKV4-R | gtctcgtgggcicggagatgtgtataagagacagTACNVGGGTATCTAATCC |

TABLE A.7

16S RIBOSOMAL RNA GENE PCR REACTION FOR PRIMERS
PRVKV4-F AND PRVKV4-R

| Component | Amount per reaction | Final Concentration |
|----------------------------------|---------------------|---------------------|
| DNA template, undiluted | 3.0 ml | unknown |
| Forward Primer, 10 mM | 0.75 ml | 0.3 mM |
| Reverse Primer, 10 mM | 0.75 ml | 0.3 mM |
| Kapa HiFi HotStart Ready Mix, 2X | 12.5 ml | 1X |
| Water | 10.0 ml | |

TABLE A.8

16S RIBOSOMAL RNA GENE PCR THERMOCYCLER CONDITIONS
FOR PRIMERS PRKV4-F AND PVKV4-R

| Step | Temperature | Time | |
|------------------|-------------|------------|---------------|
| Initial Denature | 95C | 3 minutes | — |
| Denature | 98C | 20 seconds | 3*x 30 cycles |
| Annealing | 65C | 15 seconds | |
| Extension | 72C | 15 seconds | |
| Final Extension | 72C | 3 minutes | — |

Amplicon ratio for sequencing: AMP_1380F, AMP_1389F, and AMP_PrkvV4 at a ratio of 7:7:16, respectively. Amplicon ratio in this library prep is 7.5 ng/ μ L. Molarity for each indexed library is calculated with the following formula:

$$\text{Concentration in } nM = \left(\frac{\text{concentration in ng}/\mu L}{600 \text{ g/mol} \times \text{average library size in bp}} \right) \times 10^6$$

A.3 Chapter 3 Supplementary Data

TABLE A.9

GENERA OF CROPS DETECTED IN BALI AND SINGAPORE

| Singapore | Bali |
|--------------------|--------------------|
| <i>Panax</i> | <i>Panax</i> |
| <i>Aphandra</i> | <i>Aphandra</i> |
| <i>Borassus</i> | <i>Borassus</i> |
| <i>Elaeis</i> | <i>Elaeis</i> |
| <i>Allium</i> | <i>Allium</i> |
| <i>Asparagus</i> | <i>Asparagus</i> |
| <i>Polianthes</i> | <i>Polianthes</i> |
| <i>Helianthus</i> | <i>Helianthus</i> |
| <i>Brassica</i> | <i>Brassica</i> |
| <i>Carica</i> | <i>Carica</i> |
| <i>Beta</i> | <i>Beta</i> |
| <i>Chenopodium</i> | <i>Chenopodium</i> |

TABLE A.9 (CONTINUED)

| Singapore | Bali |
|-------------------|-------------------|
| <i>Rheum</i> | <i>Rheum</i> |
| <i>Citrullus</i> | <i>Citrullus</i> |
| <i>Cucumis</i> | <i>Cucumis</i> |
| <i>Cucurbita</i> | <i>Cucurbita</i> |
| <i>Dioscorea</i> | <i>Dioscorea</i> |
| <i>Camellia</i> | <i>Camellia</i> |
| <i>Manilkara</i> | <i>Manilkara</i> |
| <i>Vaccinium</i> | <i>Vaccinium</i> |
| <i>Arachis</i> | <i>Arachis</i> |
| <i>Cajanus</i> | <i>Cajanus</i> |
| <i>Cicer</i> | <i>Cicer</i> |
| <i>Glycine</i> | <i>Glycine</i> |
| <i>Phaseolus</i> | <i>Phaseolus</i> |
| <i>Vigna</i> | <i>Vigna</i> |
| <i>Carya</i> | <i>Carya</i> |
| <i>Coffea</i> | <i>Olea</i> |
| <i>Olea</i> | <i>Perilla</i> |
| <i>Perilla</i> | <i>Sesamum</i> |
| <i>Sesamum</i> | <i>Persea</i> |
| <i>Persea</i> | <i>Myristica</i> |
| <i>Myristica</i> | <i>Flacourtia</i> |
| <i>Flacourtia</i> | <i>Plukenetia</i> |
| <i>Manihot</i> | <i>Punica</i> |
| <i>Plukenetia</i> | <i>Ananas</i> |

TABLE A.9 (CONTINUED)

| Singapore | Bali |
|-------------------|-------------------|
| <i>Punica</i> | <i>Hordeum</i> |
| <i>Ananas</i> | <i>Oryza</i> |
| <i>Hordeum</i> | <i>Saccharum</i> |
| <i>Oryza</i> | <i>Sorghum</i> |
| <i>Saccharum</i> | <i>Triticum</i> |
| <i>Sorghum</i> | <i>Zea</i> |
| <i>Triticum</i> | <i>Fragaria</i> |
| <i>Zea</i> | <i>Malus</i> |
| <i>Zizania</i> | <i>Morus</i> |
| <i>Fragaria</i> | <i>Prunus</i> |
| <i>Malus</i> | <i>Pyrus</i> |
| <i>Morus</i> | <i>Rubus</i> |
| <i>Prunus</i> | <i>Ziziphus</i> |
| <i>Pyrus</i> | <i>Anacardium</i> |
| <i>Ziziphus</i> | <i>Citrus</i> |
| <i>Anacardium</i> | <i>Litchi</i> |
| <i>Citrus</i> | <i>Mangifera</i> |
| <i>Litchi</i> | <i>Capsicum</i> |
| <i>Mangifera</i> | <i>Ipomoea</i> |
| <i>Capsicum</i> | <i>Solanum</i> |
| <i>Ipomoea</i> | <i>Vitis</i> |
| <i>Solanum</i> | <i>Curcuma</i> |
| <i>Vitis</i> | <i>Musa</i> |
| <i>Curcuma</i> | |

TABLE A.9 (CONTINUED)

Singapore Bali

Musa

A.4 Chapter 4 Supplementary Materials

TABLE A.10

SAMPLE SITES, LOCATION, SAMPLING FREQUENCY

| <i>Site</i> | Bali | | | Singapore | | |
|----------------------|------------------|-----------------|------------------|----------------------|-----------------|------------------|
| | <i>N samples</i> | <i>Latitude</i> | <i>Longitude</i> | <i>N samples</i> | <i>Latitude</i> | <i>Longitude</i> |
| AS | 2 | -8.356 | 115.161 | 4 | 1.446 | 103.778 |
| BG | 8 | -8.505 | 115.59 | 12 | 1.352 | 103.759 |
| BP | 3 | -8.838 | 115.093 | 12 | 1.352 | 103.776 |
| CK | 6 | -8.189 | 114.444 | 5 | 1.372 | 103.828 |
| LM | 6 | -8.237 | 115.383 | 6 | 1.352 | 103.818 |
| MK | 7 | -8.392 | 115.035 | 11 | 1.352 | 103.823 |
| PL | 4 | -8.146 | 114.682 | 9 | 1.399 | 103.806 |
| PU | 14 | -8.519 | 115.258 | 4 | 1.453 | 103.78 |
| TK | 8 | -8.14 | 115.32 | | | |
| U | 4 | -8.829 | 115.084 | | | |
| Total Samples | 62 | | | Total Samples | 63 | |

Additional tables available in supplementary Excel file includes:

- Protozoan and Helminth Counts in Samples
- Medicinal Plant Database

BIBLIOGRAPHY

- Stiffer fines to deter monkey feeding. *The Straits Times*, 1 2008. URL Archivedat<https://www.asiaone.com/News/AsiaOne%2BNews/Singapore/Story/A1Story20080131-47593.html>. [Online; accessed 2022-05-27].
- Balinese population and society – tourist information. URL <https://ubud.id/bali-island/balinese-population-and-society/>. [Online; accessed 2022-06-08].
- Structure, function and diversity of the healthy human microbiome. *nature*, 486 (7402):207–214, 2012.
- C. Abegg. Macaque evolution and dispersal in insular south-east asia. *Biological journal of the Linnean Society*, 2002.
- T. Aivelo and A. Medlar. Opportunities and challenges in metabarcoding approaches for helminth community identification in wild mammals. *Parasitology*, 145(5):608–621, Apr. 2018. ISSN 0031-1820, 1469-8161. doi: 10.1017/S0031182017000610. Publisher: Cambridge University Press.
- M. A. Alfellani, C. R. Stensvold, A. Vidal-Lapiedra, E. S. U. Onuoha, A. F. Fagbenro-Beyioku, and C. G. Clark. Variable geographic distribution of blastocystis subtypes and its potential implications. *Acta Tropica*, 126(1):11–18, 2013. ISSN 0001706X. doi: 10.1016/j.actatropica.2012.12.011.
- J. F. Alzate, M. Toro-Londoño, F. Cabarcas, G. Garcia-Montoya, and A. Galvan-Diaz. Contrasting microbiota profiles observed in children carrying either blastocystis spp. or the commensal amoebas entamoeba coli or endolimax nana. *Scientific Reports*, 10(1):15354, 9 2020. ISSN 2045-2322. doi: 10.1038/s41598-020-72286-y. number: 1 publisher: Nature Publishing Group.
- K. R. Amato. Co-evolution in context: The importance of studying gut microbiomes in wild animals. *Microbiome Science and Medicine*, 1(1):10–29, 2013. ISSN 2084-7653. doi: 10.2478/micsm-2013-0002.
- K. R. Amato, C. J. Yeoman, A. Kent, N. Righini, F. Carbonero, A. Estrada, H. Rex Gaskins, R. M. Stumpf, S. Yildirim, M. Torralba, M. Gillis, B. A. Wilson, K. E. Nelson, B. A. White, and S. R. Leigh. Habitat degradation impacts black howler monkey (alouatta pigra) gastrointestinal microbiomes. *The ISME*

- Journal*, 7(7):1344–1353, 7 2013. ISSN 1751-7370. doi: 10.1038/ismej.2013.16. number: 7 publisher: Nature Publishing Group.
- K. R. Amato, J. G. Sanders, S. J. Song, M. Nute, J. L. Metcalf, L. R. Thompson, J. T. Morton, A. Amir, V. J. McKenzie, G. Humphrey, G. Gogul, J. Gaffney, A. L. Baden, G. A.O. Britton, F. P. Cuzzo, A. Di Fiore, N. J. Dominy, T. L. Goldberg, A. Gomez, M. M. Kowalewski, R. J. Lewis, A. Link, M. L. Sauter, S. Tecot, B. A. White, K. E. Nelson, R. M. Stumpf, R. Knight, and S. R. Leigh. Evolutionary trends in host physiology outweigh dietary niche in structuring primate gut microbiomes. *ISME Journal*, 13(3):576–587, 2019. ISSN 17517370. doi: 10.1038/s41396-018-0175-0.
- L. O. Andersen and C. R. Stensvold. Blastocystis in health and disease: Are we moving from a clinical to a public health perspective? *Journal of Clinical Microbiology*, 54(3):524–528, 2016. doi: 10.1128/JCM.02520-15.
- M. Arumugam, J. Raes, E. Pelletier, D. Le Paslier, T. Yamada, D. R. Mende, G. R. Fernandes, J. Tap, T. Bruls, J. M. Batto, M. Bertalan, N. Borruel, F. Casellas, L. Fernandez, L. Gautier, T. Hansen, M. Hattori, T. Hayashi, M. Kleerebezem, K. Kurokawa, M. Leclerc, F. Levenez, C. Manichanh, H. B. Nielsen, T. Nielsen, N. Pons, J. Poulain, J. Qin, T. Sicheritz-Ponten, S. Tims, D. Torrents, E. Ugarte, E. G. Zoetendal, J. Wang, F. Guarner, O. Pedersen, W. M. De Vos, S. Brunak, J. Doré, J. Weissenbach, S. D. Ehrlich, and P. Bork. Enterotypes of the human gut microbiome. *Nature*, 473(7346):174–180, 2011. ISSN 14764687. doi: 10.1038/nature09944.
- A. E. Asangba, L. Mugisha, J. Rukundo, R. J. Lewis, A. Halajian, L. Cortés-Ortiz, R. E. Junge, M. T. Irwin, J. Karlson, A. Perkin, et al. Large comparative analyses of primate body site microbiomes indicate that the oral microbiome is unique among all body sites and conserved among nonhuman primates. *Microbiology Spectrum*, pages e01643–21, 2022.
- C. Audebert, G. Even, A. Cian, B. I. Group, A. Loywick, S. Merlin, E. Viscogliosi, M. Chabé, D. El Safadi, G. Certad, L. Delhaes, B. Pereira, C. Nourrisson, P. Poirier, I. Wawrzyniak, F. Delbac, C. Morelle, P. Bastien, L. Lachaud, A. P. Bellanger, F. Botterel, E. Candolfi, G. Desoubes, F. Morio, C. Pomares, and M. Rabodonirina. Colonization with the enteric protozoa blastocystis is associated with increased diversity of human gut bacterial microbiota. *Scientific Reports*, 6(April):1–11, 2016. ISSN 20452322. doi: 10.1038/srep25255.
- F. A. Ayeni, E. Biagi, S. Rampelli, J. Fiori, M. Soverini, H. J. Audu, S. Cristino, L. Caporali, S. L. Schnorr, V. Carelli, P. Brigidi, M. Candela, and S. Turroni. Infant and adult gut microbiome and metabolome in rural bassa and urban settlers from nigeria. *Cell Reports*, 23(10):3056–3067, 6 2018. ISSN 2211-1247. doi: 10.1016/j.celrep.2018.05.018.

- A. Bagavan, A. A. Rahuman, C. Kamaraj, N. K. Kaushik, D. Mohanakrishnan, and D. Sahal. Antiplasmodial activity of botanical extracts against *Plasmodium falciparum*. *Parasitology Research*, 108(5):1099–1109, May 2011a. ISSN 1432-1955. doi: 10.1007/s00436-010-2151-0. URL <https://doi.org/10.1007/s00436-010-2151-0>.
- A. Bagavan, A. A. Rahuman, N. K. Kaushik, and D. Sahal. In vitro antimalarial activity of medicinal plant extracts against *Plasmodium falciparum*. *Parasitology Research*, 108(1):15–22, Jan. 2011b. ISSN 1432-1955. doi: 10.1007/s00436-010-2034-4. URL <https://doi.org/10.1007/s00436-010-2034-4>.
- A. Baniel, K. R. Amato, J. C. Beehner, T. J. Bergman, A. Mercer, R. F. Perlman, L. Petrullo, L. Reitsema, S. Sams, A. Lu, and N. Snyder-Mackler. Seasonal shifts in the gut microbiome indicate plastic responses to diet in wild geladas. *bioRxiv*, page 2020.07.07.192336, 7 2020. doi: 10.1101/2020.07.07.192336. publisher: Cold Spring Harbor Laboratory section: New Results.
- A.-K. Bär, N. Phukan, J. Pinheiro, and A. Simoes-Barbosa. The interplay of host microbiota and parasitic protozoans at mucosal interfaces: implications for the outcomes of infections and diseases. *PLoS neglected tropical diseases*, 9(12): e0004176, 2015.
- N. R. Barash and J. G. Maloney. crossm giardia alters commensal microbial. *American Society for Microbiology*, 85(6):1–18, 2017.
- P. Barbera, A. M. Kozlov, L. Czech, B. Morel, D. Darriba, T. Flouri, and A. Stamatakis. Epa-ng: Massively parallel evolutionary placement of genetic sequences. *Systematic Biology*, 68(2):365–369, 3 2019. ISSN 1063-5157. doi: 10.1093/sysbio/syy054.
- C. Barelli and M. A. Huffman. Leaf swallowing and parasite expulsion in Khao Yai white-handed gibbons (*Hylobates lar*), the first report in an Asian ape species. *American Journal of Primatology*, 79(3):e22610, 2017. ISSN 1098-2345. doi: 10.1002/ajp.22610. URL <http://onlinelibrary.wiley.com/doi/abs/10.1002/ajp.22610>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/ajp.22610>.
- C. Barelli, D. Albanese, C. Donati, M. Pindo, C. Dallago, F. Rovero, D. Cavalieri, K. M. Tuohy, H. C. Hauffe, and C. De Filippo. Habitat fragmentation is associated to gut microbiota diversity of an endangered primate: implications for conservation. *Scientific reports*, 5(1):1–12, 2015. publisher: Nature Publishing Group.
- R. K. Bauri, M. N. Tigga, and S. S. Kullu. A review on use of medicinal plants to control parasites. *Indian Journal of Natural Products and Resources (IJNPR) [Formerly Natural Product Radiance (NPR)]*, 6(4):268–277, Dec. 2015. ISSN 0976-0512. URL <http://14.139.47.23/index.php/IJNPR/article/view/8837>. Number: 4.

- F. Beghini, E. Pasolli, T. D. Truong, L. Putignani, S. M. Cacciò, and N. Segata. Large-scale comparative metagenomics of blastocystis, a common member of the human gut microbiome. *The ISME journal*, 11(12):2848–2863, 2017.
- R. Blekhman, J. K. Goodrich, K. Huang, Q. Sun, R. Bukowski, J. T. Bell, T. D. Spector, A. Keinan, R. E. Ley, D. Gevers, et al. Host genetic variation impacts microbiome composition across human body sites. *Genome biology*, 16(1):1–12, 2015.
- J. Blessmann, I. K. M. Ali, P. A. Ton Nu, B. T. Dinh, T. Q. Ngo Viet, A. Le Van, C. G. Clark, and E. Tannich. Longitudinal study of intestinal entamoeba histolytica infections in asymptomatic adult carriers. *Journal of Clinical Microbiology*, 41(10):4745–4750, 10 2003. doi: 10.1128/JCM.41.10.4745-4750.2003. publisher: American Society for Microbiology.
- R. C. E. Bowyer, M. A. Jackson, T. Pallister, J. Skinner, T. D. Spector, A. A. Welch, and C. J. Steves. Use of dietary indices to control for diet in human gut microbiota studies. *Microbiome*, 6(1):1–11, 12 2018. ISSN 2049-2618. doi: 10.1186/s40168-018-0455-y. number: 1 publisher: BioMed Central.
- A. Brand, J. D. Barnes, K. S. Mackenzie, F. C. Odds, and N. A. Gow. Cell wall glycans and soluble factors determine the interactions between the hyphae of candida albicans and pseudomonas aeruginosa. *FEMS Microbiology Letters*, 287(1):48–55, 2008. ISSN 03781097. doi: 10.1111/j.1574-6968.2008.01301.x.
- F. Brotcorne. Behavioral ecology of commensal long-tailed macaque (macaca fascicularis) populations in bali, indonesia: impact of anthropic factors. 2014.
- C. G. Buffie and E. G. Pamer. Microbiota-mediated colonization resistance against intestinal pathogens. *Nature Reviews Immunology*, 13(11):790–801, 11 2013. ISSN 1474-1741. doi: 10.1038/nri3535. number: 11 publisher: Nature Publishing Group.
- T. P. N. Bui, J. Ritari, S. Boeren, P. de Waard, C. M. Plugge, and W. M. de Vos. Production of butyrate from lysine and the amadori product fructoselysine by a human gut commensal. *Nature Communications*, 6(1):10062, 12 2015. ISSN 2041-1723. doi: 10.1038/ncomms10062. number: 1 publisher: Nature Publishing Group.
- S. L. Burgess, C. A. Gilchrist, T. C. Lynn, and W. A. Petri. Parasitic protozoa and interactions with the host intestinal microbiota. *Infection and Immunity*, 85(8), 8 2017. ISSN 0019-9567, 1098-5522. doi: 10.1128/IAI.00101-17. URL <https://iai.asm.org/content/85/8/e00101-17>. PMID: 28584161.
- F. Cabana, J. B. Clayton, K. A. Nekar, W. Wirdateti, D. Knights, and H. Seedorf. Nutrient-based diet modifications impact on the gut microbiome of the javan slow loris (nycticebus javanicus). *Scientific Reports*, 9(1):1–11, 2019. ISSN 20452322. doi: 10.1038/s41598-019-40911-0.

- B. Callahan. Rdp taxonomic training data formatted for dada2 (rdp trainset 16/release 11.5). 6 2017. URL https://zenodo.org/record/801828#.X0_QdtYpBBw. DOI: 10.5281/zenodo.801828 type: dataset.
- B. J. Callahan, P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. Dada2: High-resolution sample inference from illumina amplicon data. *Nature Methods*, 13(7):581–583, 7 2016. ISSN 1548-7105. doi: 10.1038/nmeth.3869. number: 7 publisher: Nature Publishing Group.
- R. Caspi, T. Altman, R. Billington, K. Dreher, H. Foerster, C. A. Fulcher, T. A. Holland, I. M. Keseler, A. Kothari, A. Kubo, M. Krummenacker, M. Latendresse, L. A. Mueller, Q. Ong, S. Paley, P. Subhraveti, D. S. Weaver, D. Weerasinghe, P. Zhang, and P. D. Karp. The metacyc database of metabolic pathways and enzymes and the biocyc collection of pathway/genome databases. *Nucleic Acids Research*, 42(D1):D459–D471, 1 2014. ISSN 0305-1048. doi: 10.1093/nar/gkt1103.
- M. Chabé, A. Lokmer, and L. Ségurel. Gut protozoa: friends or foes of the human gut microbiota? *Trends in Parasitology*, 33(12):925–934, 2017.
- Z. Chen, Y. K. Yeoh, M. Hui, P. Y. Wong, M. C. Chan, M. Ip, J. Yu, R. D. Burk, F. K. Chan, and P. K. Chan. Diversity of macaque microbiota compared to the human counterparts. *Scientific Reports*, 8(1):1–15, 2018. doi: 10.1038/s41598-018-33950-6.
- I. Cho and M. J. Blaser. The human microbiome: at the interface of health and disease. *Nature Reviews Genetics*, 13(4):260–270, 2012. publisher: Nature Publishing Group.
- A. R. Cirtwill and P. Hambäck. Building food networks from molecular data: Bayesian or fixed-number thresholds for including links. *Basic and Applied Ecology*, 11 2020. ISSN 1439-1791. doi: 10.1016/j.baae.2020.11.007. URL <http://www.sciencedirect.com/science/article/pii/S1439179120301237>. [Online; accessed 2020-12-07].
- A. R. Cirtwill and P. Hambäck. Building food networks from molecular data: Bayesian or fixed-number thresholds for including links. *Basic and Applied Ecology*, 50:67–76, 2021. publisher: Elsevier.
- M. J. Claesson, I. B. Jeffery, S. Conde, S. E. Power, E. M. O’Connor, S. Cusack, H. M. B. Harris, M. Coakley, B. Lakshminarayanan, O. O’Sullivan, G. F. Fitzgerald, J. Deane, M. O’Connor, N. Harnedy, K. O’Connor, D. O’Mahony, D. van Sinderen, M. Wallace, L. Brennan, C. Stanton, J. R. Marchesi, A. P. Fitzgerald, F. Shanahan, C. Hill, R. P. Ross, and P. W. O’Toole. Gut microbiota composition correlates with diet and health in the elderly. *Nature*, 488(7410):178–184, 8 2012. ISSN 1476-4687. doi: 10.1038/nature11319. number: 7410 publisher: Nature Publishing Group.

- C. G. Clark, M. van der Giezen, M. A. Alfellani, and C. R. Stensvold. *Recent Developments in Blastocystis Research*, volume 82. Elsevier, 2013. ISBN 978-0-12-407706-5. URL <http://dx.doi.org/10.1016/B978-0-12-407706-5.00001-0>. DOI: 10.1016/B978-0-12-407706-5.00001-0.
- J. B. Clayton, P. Vangay, H. Huang, T. Ward, B. M. Hillmann, G. A. Al-Ghalith, D. A. Travis, H. T. Long, B. V. Tuan, V. V. Minh, F. Cabana, T. Nadler, B. Toddes, T. Murphy, K. E. Glander, T. J. Johnson, and D. Knights. Captivity humanizes the primate microbiome. *Proceedings of the National Academy of Sciences*, 113 (37):10376–10381, 2016. doi: 10.1073/pnas.1521835113.
- J. B. Clayton, A. Gomez, K. Amato, D. Knights, D. A. Travis, R. Blekhman, R. Knight, S. Leigh, R. Stumpf, T. Wolf, K. E. Glander, F. Cabana, and T. J. Johnson. The gut microbiome of nonhuman primates: Lessons in ecology and evolution. *American Journal of Primatology*, pages e22867–e22867, 6 2018. doi: 10.1002/ajp.22867.
- J. C. Clemente, E. C. Pehrsson, M. J. Blaser, K. Sandhu, Z. Gao, B. Wang, M. Margis, G. Hidalgo, M. Contreras, Ó. Noya-Alarcón, et al. The microbiome of uncontacted amerindians. *Science advances*, 1(3):e1500183, 2015.
- E. Coissac, T. Riaz, and N. Puillandre. Bioinformatic challenges for DNA metabarcoding of plants and animals. *Molecular Ecology*, 21(8):1834–1847, 2012. ISSN 1365-294X. doi: 10.1111/j.1365-294X.2012.05550.x. URL <http://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-294X.2012.05550.x>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-294X.2012.05550.x>.
- R. T. Corlett. The ecological transformation of singapore, 1819-1990. *Journal of Biogeography*, 19(4):411–420, 1992. ISSN 0305-0270. doi: 10.2307/2845569. publisher: Wiley.
- P. I. Costea, F. Hildebrand, A. Manimozhiyan, F. Bäckhed, M. J. Blaser, F. D. Bushman, W. M. De Vos, S. D. Ehrlich, C. M. Fraser, M. Hattori, C. Huttenhower, I. B. Jeffery, D. Knights, J. D. Lewis, R. E. Ley, H. Ochman, P. W. O’Toole, C. Quince, D. A. Relman, F. Shanahan, S. Sunagawa, J. Wang, G. M. Weinstock, G. D. Wu, G. Zeller, L. Zhao, J. Raes, R. Knight, and P. Bork. Enterotypes in the landscape of gut microbial community composition. *Nature Microbiology*, 3(1):8–16, 2017. doi: 10.1038/s41564-017-0072-8.
- E. K. Costello, C. L. Lauber, M. Hamady, N. Fierer, J. I. Gordon, and R. Knight. Bacterial community variation in human body habitats across space and time. *Science*, 326(December):1694–1697, 2011. doi: 10.1093/oxfordhob/9780199556137.003.0004.
- J. Cox-Singh, T. M. Davis, K.-S. Lee, S. S. Shamsul, A. Matusop, S. Ratnam, H. A. Rahman, D. J. Conway, and B. Singh. Plasmodium knowlesi malaria in humans

- is widely distributed and potentially life threatening. *Clinical infectious diseases*, 46(2):165–171, 2008. Publisher: The University of Chicago Press.
- Q. Cui, Y. Ren, and H. Xu. The escalating effects of wildlife tourism on human–wildlife conflict. *Animals*, 11(5):1378, 5 2021. ISSN 2076-2615. doi: 10.3390/ani11051378. number: 5 publisher: Multidisciplinary Digital Publishing Institute.
- L. Czech, P. Barbera, and A. Stamatakis. Genesis and gappa: processing, analyzing and visualizing phylogenetic (placement) data. *Bioinformatics*, 36(10): 3263–3265, 5 2020. ISSN 1367-4803. doi: 10.1093/bioinformatics/btaa070.
- B. Das, T. S. Ghosh, S. Kedia, R. Rampal, S. Saxena, S. Bag, R. Mitra, M. Dayal, O. Mehta, A. Surendranath, S. P. L. Travis, P. Tripathi, G. B. Nair, and V. Ahuja. Analysis of the gut microbiome of rural and urban healthy indians living in sea level and high altitude areas. *Scientific Reports*, 8(1):10104, 7 2018. ISSN 2045-2322. doi: 10.1038/s41598-018-28550-3. number: 1 publisher: Nature Publishing Group.
- L. A. David, A. C. Materna, M. C. Blackburn, S. E. Erdman, J. Friedman, E. J. Alm, A. Perrotta, and M. I. C. Baptista. Host lifestyle affects human microbiota on daily timescales. *Genome Biology*, 15(1):1–15, 2014a. doi: 10.1186/s13059-016-0988-y.
- L. A. David, C. F. Maurice, R. N. Carmody, D. B. Gootenberg, J. E. Button, B. E. Wolfe, A. V. Ling, A. S. Devlin, Y. Varma, M. A. Fischbach, S. B. Biddinger, R. J. Dutton, and P. J. Turnbaugh. Diet rapidly and reproducibly alters the human gut microbiome. *Nature*, 505(7484):559–563, 1 2014b. ISSN 1476-4687. doi: 10.1038/nature12820. number: 7484 publisher: Nature Publishing Group.
- C. De Filippo, D. Cavalieri, M. Di Paola, M. Ramazzotti, J. B. Poullet, S. Massart, S. Collini, G. Pieraccini, and P. Lionetti. Impact of diet in shaping gut microbiota revealed by a comparative study in children from europe and rural africa. *Proceedings of the National Academy of Sciences*, 107(33):14691–14696, 2010.
- F. De Vadder, E. Grasset, L. Mannerås Holm, G. Karsenty, A. J. Macpherson, L. E. Olofsson, and F. Bäckhed. Gut microbiota regulates maturation of the adult enteric nervous system via enteric serotonin networks. *Proceedings of the National Academy of Sciences*, 115(25):6458–6463, 2018. ISSN 1720017115. doi: 10.1073/pnas.1720017115.
- P. H. Degnan, A. E. Pusey, E. V. Lonsdorf, J. Goodall, E. E. Wroblewski, M. L. Wilson, R. S. Rudicell, B. H. Hahn, and H. Ochman. Factors associated with the diversification of the gut microbial communities within chimpanzees from gombe national park. *Proceedings of the National Academy of Sciences*, 109(32):13034–13039, 8 2012.

- L. Deng, L. Wojciech, N. R. J. Gascoigne, G. Peng, and K. S. W. Tan. New insights into the interactions between blastocystis, the gut microbiota, and host immunity. *PLOS Pathogens*, 17(2):e1009253, 2 2021. ISSN 1553-7374. doi: 10.1371/journal.ppat.1009253. publisher: Public Library of Science.
- F. E. Dewhirst, T. Chen, J. Izard, B. J. Paster, A. C. R. Tanner, W.-H. Yu, A. Lakshmanan, and W. G. Wade. The human oral microbiome. *Journal of Bacteriology*, 192(19):5002–5017, 10 2010. ISSN 0021-9193, 1098-5530. doi: 10.1128/JB.00542-10. PMID: 20656903.
- G. M. Douglas, V. J. Maffei, J. Zaneveld, S. N. Yurgel, J. R. Brown, C. M. Taylor, C. Huttenhower, and M. G. Langille. Picrust2: An improved and customizable approach for metagenome inference. *BioRxiv*, page 672295, 2020.
- S. Dugasani, M. R. Pichika, V. D. Nadarajah, M. K. Balijepalli, S. Tandra, and J. N. Korlakunta. Comparative antioxidant and anti-inflammatory effects of [6]-gingerol, [8]-gingerol, [10]-gingerol and [6]-shogaol. *Journal of Ethnopharmacology*, 127(2):515–520, 2010. ISSN 03788741. doi: 10.1016/j.jep.2009.10.004.
- L. A. Dunn, P. F. L. Boreham, and D. J. Stenzel. Ultrastructural variation of blastocystis hominis stocks in culture. *International Journal for Parasitology*, 19(1): 43–56, 2 1989. ISSN 0020-7519. doi: 10.1016/0020-7519(89)90020-9.
- R. C. Edgar. Search and clustering orders of magnitude faster than blast. *Bioinformatics (Oxford, England)*, 26(19):2460–1, 10 2010. doi: 10.1093/bioinformatics/btq461.
- K. Eschweiler, J. B. Clayton, A. Moresco, E. A. McKenney, L. J. Minter, M. J. Suhr Van Haute, W. Gasper, S. S. Hayer, L. Zhu, K. Cooper, and K. Ange-van Heugten. Host identity and geographic location significantly affect gastrointestinal microbial richness and diversity in western lowland gorillas (gorilla gorilla gorilla) under human care. *Animals*, 11(12):3399, 12 2021. ISSN 2076-2615. doi: 10.3390/ani11123399. number: 12 publisher: Multidisciplinary Digital Publishing Institute.
- G. Even, A. Lokmer, J. Rodrigues, C. Audebert, E. Viscogliosi, L. Gurel, M. Chabé, L. Xiao, H. Mirjalali, A. D. Tsaousis, and E. Gentekaki. Changes in the human gut microbiota associated with colonization by blastocystis sp. and entamoeba spp. in non-industrialized populations. *Frontiers in Cellular and Infection Microbiology*, 11, 3 2021. doi: 10.3389/fcimb.2021.533528.
- G. Falony, M. Joossens, S. Vieira-Silva, J. Wang, Y. Darzi, K. Faust, A. Kurilshikov, M. J. Bonder, M. Valles-Colomer, D. Vandeputte, R. Y. Tito, S. Chaffron, L. Rymenans, C. Verspecht, D. L. Sutter, G. Lima-Mendez, K. D’hoë, K. Jonckheere, D. Homola, R. Garcia, E. F. Tigchelaar, L. Eeckhaut, J. Fu, L. Henckaerts, A. Zhernakova, C. Wijmenga, and J. Raes. Population-level analysis of gut microbiome variation. *Science*, 352(6285):560–564, 2016. ISSN 10959203. doi: 10.1126/science.aad3503.

- R. S. Feldberg, S. Chang, A. Kotik, M. Nadler, Z. Neuwirth, D. Sundstrom, and N. Thompson. In vitro mechanism of inhibition of bacterial cell growth by allicin. *Antimicrobial agents and chemotherapy*, 32(12):1763–1768, 1988.
- J. A. Fellows Yates, I. M. Velsko, F. Aron, C. Posth, C. A. Hofman, R. M. Austin, C. E. Parker, A. E. Mann, K. Nägele, K. W. Arthur, J. W. Arthur, C. C. Bauer, I. Crevecoeur, C. Cupillard, M. C. Curtis, L. Dalén, M. Díaz-Zorita Bonilla, J. C. Díez Fernández-Lomana, D. G. Drucker, E. Escribano Esquivá, M. Francken, V. E. Gibbon, M. R. González Morales, A. Grande Mateu, K. Harvati, A. G. Henry, L. Humphrey, M. Menéndez, D. Mihailović, M. Peresani, S. Rodríguez Moroder, M. Roksandic, H. Rougier, S. Sázelová, J. T. Stock, L. G. Straus, J. Svoboda, B. Teßmann, M. J. Walker, R. C. Power, C. M. Lewis, K. Sankaranarayanan, K. Guschanski, R. W. Wrangham, F. E. Dewhirst, D. C. Salazar-García, J. Krause, A. Herbig, and C. Warinner. The evolution and changing ecology of the african hominid oral microbiome. *Proceedings of the National Academy of Sciences*, 118(20):e2021655118, 5 2021. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.2021655118.
- A. D. Fernandes, J. M. Macklaim, T. G. Linn, G. Reid, and G. B. Gloor. Anova-like differential expression (aldex) analysis for mixed population rna-seq. *PLOS ONE*, 8(7):e67019, 7 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0067019. publisher: Public Library of Science.
- A. D. Fernandes, J. N. Reid, J. M. Macklaim, T. A. McMurrough, D. R. Edgell, and G. B. Gloor. 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(1):15, 5 2014. ISSN 2049-2618. doi: 10.1186/2049-2618-2-15.
- H. J. Flint, S. H. Duncan, K. P. Scott, and P. Louis. Interactions and competition within the microbial community of the human colon: links between diet and health. *Environmental Microbiology*, 9(5):1101–1111, 2007. ISSN 1462-2920. doi: <https://doi.org/10.1111/j.1462-2920.2007.01281.x>.
- A. T. Fogel. The gut microbiome of wild lemurs: A comparison of sympatric **lemur catta** and **propithecus verreauxi**. *Folia Primatologica*, 86(1-2):85–95, 2015. doi: 10.1159/000369971.
- J. Fooden. Comparative review of fascicularis-group species of macaques (primates: Macaca). [http://dx.doi.org/10.3158/0015-0754\(2006\)107\[1:CROFSM\]2.0.CO;2](http://dx.doi.org/10.3158/0015-0754(2006)107[1:CROFSM]2.0.CO;2), 5 2006. doi: 10.3158/0015-0754(2006)107[1:CROFSM]2.0.CO;2.
- J. Forsell, J. Bengtsson-Palme, M. Angelin, A. Johansson, B. Evengård, and M. Granlund. The relation between blastocystis and the intestinal microbiota in swedish travellers. *BMC Microbiology*, 17(1):1–9, 12 2017. ISSN 1471-2180. doi: 10.1186/s12866-017-1139-7. number: 1 publisher: BioMed Central.

- J. S. Frankel, E. K. Mallott, L. M. Hopper, S. R. Ross, and K. R. Amato. The effect of captivity on the primate gut microbiome varies with host dietary niche. *American Journal of Primatology*, 81(12):e23061, 2019. ISSN 1098-2345. doi: 10.1002/ajp.23061.
- A. Fuentes. *Monkeys on the Edge: Ecology and Management of Long-Tailed Macaques and their Interface with Humans*. Cambridge University Press, Apr. 2011. ISBN 978-1-139-50041-8. Google-Books-ID: XCSwctd20LwC.
- A. Fuentes, S. Kalchik, L. Gettler, A. Kwiatt, M. Konecki, and L. Jones-Engel. Characterizing human–macaque interactions in singapore. *American Journal of Primatology: Official Journal of the American Society of Primatologists*, 70(9): 879–883, 2008.
- A. Fuentes, A. L. Rompis, I. Putra, N. L. Watiniasih, I. N. Suartha, I. Soma, I. N. Wandia, I. Putra, R. Stephenson, and W. Selamet. Macaque behavior at the human–monkey interface: the activity and demography of semi-free-ranging macaca fascicularis at padangtegal, bali, indonesia. *Monkeys on the edge: ecology and management of long-tailed macaques and their interface with humans*, page 159–179, 2011. publisher: Cambridge University Press New York.
- M. B. Gardner and P. A. Luciw. Macaque models of human infectious disease. *ILAR journal*, 49(2):220–255, 2008.
- G. B. Gloor and G. Reid. Compositional analysis: a valid approach to analyze microbiome high-throughput sequencing data. *Canadian journal of microbiology*, 62(8):692–703, 2016. publisher: NRC research press.
- G. B. Gloor, J. M. Macklaim, and A. D. Fernandes. Displaying variation in large datasets: plotting a visual summary of effect sizes. *Journal of Computational and Graphical Statistics*, 25(3):971–979, 2016a.
- G. B. Gloor, J. R. Wu, V. Pawlowsky-Glahn, and J. J. Egozcue. It's all relative: analyzing microbiome data as compositions. *Annals of Epidemiology*, 26(5): 322–329, 5 2016b. ISSN 1047-2797. doi: 10.1016/j.annepidem.2016.03.003.
- G. B. Gloor, J. M. Macklaim, V. Pawlowsky-Glahn, and J. J. Egozcue. Microbiome datasets are compositional: And this is not optional. *Frontiers in Microbiology*, 8, 2017. ISSN 1664-302X. doi: 10.3389/fmicb.2017.02224. URL <https://www.frontiersin.org/articles/10.3389/fmicb.2017.02224/full?report=reader>. publisher: Frontiers.
- A. Gomez, K. Petrzekova, C. J. Yeoman, K. Vlckova, J. Mrázek, I. Koppova, F. Carbonero, A. Ulanov, D. Modry, A. Todd, M. Torralba, K. E. Nelson, H. R. Gaskins, B. Wilson, R. M. Stumpf, B. A. White, and S. R. Leigh. Gut microbiome composition and metabolomic profiles of wild western lowland gorillas (*Gorilla gorilla gorilla*) reflect host ecology. *Molecular Ecology*, 24(10):2551–2565, 5 2015. ISSN 09621083. doi: 10.1111/mec.13181.

- I. GRECH-MORA, M.-L. FARDEAU, B. K. C. PATEL, B. OLLIVIER, A. RIMBAULT, G. PRENSIER, J.-L. GARCIA, and E. . GARNIER-SILLAM. Isolation and characterization of sporobacter termitidis gen. nov., sp. nov., from the digestive tract of the wood-feeding termite nasutitermes lujae. *International Journal of Systematic and Evolutionary Microbiology*, 46(2):512–518, 1996. ISSN 1466-5034,. doi: 10.1099/00207713-46-2-512. publisher: Microbiology Society,.
- L. K. Greene, E. A. McKenney, T. M. O’Connell, and C. M. Drea. The critical role of dietary foliage in maintaining the gut microbiome and metabolome of folivorous sifakas. *Scientific Reports*, 8(1):1–13, 2018. doi: 10.1038/s41598-018-32759-7.
- L. K. Greene, E. Rambeloson, H. A. Rasoanaivo, E. D. Foss, A. D. Yoder, C. M. Drea, and M. B. Blanco. Gut microbial diversity and ecological specialization in four sympatric lemur species under lean conditions. *International Journal of Primatology*, 11 2021. ISSN 1573-8604. doi: 10.1007/s10764-021-00257-9. URL <https://doi.org/10.1007/s10764-021-00257-9>. [Online; accessed 2021-11-09].
- L. Grieneisen, M. Dasari, T. J. Gould, J. R. Björk, J.-C. Grenier, V. Yotova, D. Jansen, N. Gottel, J. B. Gordon, N. H. Learn, L. R. Gesquiere, T. L. Wango, R. S. Mututua, J. K. Warutere, L. Siodi, J. A. Gilbert, L. B. Barreiro, S. C. Alberts, J. Tung, E. A. Archie, and R. Blekhman. Gut microbiome heritability is nearly universal but environmentally contingent. *Science*, 373(6551):181–186, 7 2021. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.aba5483. publisher: American Association for the Advancement of Science section: Research Article PMID: 34244407.
- L. E. Grieneisen, M. J. Charpentier, S. C. Alberts, R. Blekhman, G. Bradburd, J. Tung, and E. A. Archie. Genes, geology and germs: gut microbiota across a primate hybrid zone are explained by site soil properties, not host species. *Proceedings of the Royal Society B*, 286(1901):20190431, 2019.
- M. D. Gumert. The common monkey of southeast asia: Longtailed macaque populations, ethnophoresy, and their occurrence in human environments. *Monkeys on the Edge: Ecology and Management of Long-tailed Macaques and their Interface with Humans*, page 3–44, 2011. publisher: Cambridge: Cambridge University Press.
- V. L. Hale, C. L. Tan, K. Niu, Y. Yang, R. Knight, Q. Zhang, D. Cui, and K. R. Amato. Diet versus phylogeny: a comparison of gut microbiota in captive colobine monkey species. *Microbial Ecology*, 75(2):515–527, 2 2018. ISSN 1432-184X. doi: 10.1007/s00248-017-1041-8.
- C. S. Han, M. A. Martin, A. E. Dichosa, A. R. Daughton, S. Fietze, H. Kaplan, M. D. Gurven, and J. Alcock. Salivary microbiomes of indigenous tsimane mothers and infants are distinct despite frequent premastication. *PeerJ*, 4:e2660–e2660, 2016. doi: 10.7717/peerj.2660.

- M. F. Hansen, M. Gill, V. A. Nawangsari, K. L. Sanchez, S. M. Cheyne, V. Nijman, and A. Fuentes. Conservation of long-tailed macaques: implications of the updated IUCN status and the CoVID-19 pandemic. *Primate Conservation*, 35: 1–11, 2021.
- A. L. Hicks, K. J. Lee, M. Couto-Rodriguez, J. Patel, R. Sinha, C. Guo, S. H. Olson, A. Seimon, T. A. Seimon, A. U. Ondzie, W. B. Karesh, P. Reed, K. N. Cameron, W. I. Lipkin, and B. L. Williams. Gut microbiomes of wild great apes fluctuate seasonally in response to diet. *Nature Communications*, 9(1), 2018. doi: 10.1038/s41467-018-04204-w.
- D. A. Hogan, Å. Vik, and R. Kolter. A pseudomonas aeruginosa quorum-sensing molecule influences candida albicans morphology. *Molecular Microbiology*, 54(5):1212–1223, 2004. ISSN 0950382X. doi: 10.1111/j.1365-2958.2004.04349.x.
- L. V. Hooper, D. R. Littman, and A. J. Macpherson. Interactions between the microbiota and the immune system. *science*, 336(6086):1268–1273, 2012.
- S. Hosoda, S. Nishijima, T. Fukunaga, M. Hattori, and M. Hamada. Revealing the microbial assemblage structure in the human gut microbiome using latent dirichlet allocation. *Microbiome*, 8(1):1–12, 2020.
- L. Hubert and P. Arabie. Comparing partitions. *Journal of Classification*, 2(1): 193–218, 12 1985. ISSN 1432-1343. doi: 10.1007/BF01908075.
- M. A. Huffman. Current evidence for self-medication in primates: A multidisciplinary perspective. *American Journal of Physical Anthropology*, 104(S25):171–200, 1997. ISSN 1096-8644. doi: 10.1002/(SICI)1096-8644(1997)25+{171::AID-AJPA7}3.0.CO;2-7. URL <http://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291096-8644%281997%2925%20%3C171%3A%3AAID-AJPA7%3E3.0.CO%3B2-7>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291096-8644%281997%2925%2B%3C171%3A%3AAID-AJPA7%3E3.0.CO%3B2-7>.
- M. A. Huffman, J. E. Page, M. V. K. Sukhdeo, S. Gotoh, M. S. Kalunde, T. Chandrasiri, and G. H. N. Towers. Leaf-swallowing by chimpanzees: A behavioral adaptation for the control of strongyle nematode infections. *International Journal of Primatology*, 17(4):475–503, Aug. 1996. ISSN 1573-8604. doi: 10.1007/BF02735188. URL <https://doi.org/10.1007/BF02735188>.
- R. A. Hussein and A. A. El-Anssary. Plants secondary metabolites: the key drivers of the pharmacological actions of medicinal plants. *Herbal medicine*, 1:13, 2019. Publisher: IntechOpen London, UK.
- V. Iebba, F. Santangelo, V. Totino, F. Pantanella, A. Monsia, V. Di Cristanziano, D. Di Cave, S. Schippa, F. Berrilli, and R. D’Alfonso. Gut microbiota related to giardia duodenalis, entamoeba spp. and blastocystis hominis infections in

- humans from côte d'ivoire. *The Journal of Infection in Developing Countries*, 10 (09):1035–1041, 9 2016. ISSN 1972-2680. doi: 10.3855/jidc.8179.
- K. Ilham, Rizaldi, J. Nurdin, and Y. Tsuji. Status of urban populations of the long-tailed macaque (*macaca fascicularis*) in west sumatra, indonesia. *Primates*, 58 (2):295–305, 4 2017. ISSN 1610-7365. doi: 10.1007/s10329-016-0588-1.
- K. Ilham, Rizaldi, J. Nurdin, and Y. Tsuji. Effect of provisioning on the temporal variation in the activity budget of urban long-tailed macaques (*macaca fascicularis*) in west sumatra, indonesia. *Folia Primatologica*, 89(5):347–356, 2018. ISSN 0015-5713, 1421-9980. doi: 10.1159/000491790. publisher: Karger Publishers PMID: 30212833.
- M. Imwong, W. Madmanee, K. Suwannasin, C. Kunasol, T. J. Peto, R. Tripura, L. von Seidlein, C. Nguon, C. Davoeung, N. P. Day, and others. Asymptomatic natural human infections with the simian malaria parasites *Plasmodium cynomolgi* and *Plasmodium knowlesi*. *The Journal of infectious diseases*, 219 (5):695–702, 2019. Publisher: Oxford University Press US.
- M. C. Janiak, M. J. Montague, C. I. Villamil, M. K. Stock, A. E. Trujillo, A. N. DePasquale, J. D. Orkin, S. E. Bauman Surratt, O. Gonzalez, M. L. Platt, M. I. Martínez, S. C. Antón, M. G. Dominguez-Bello, A. D. Melin, and J. P. Higham. Age and sex-associated variation in the multi-site microbiome of an entire social group of free-ranging rhesus macaques. *Microbiome*, 9(1):68, 3 2021. ISSN 2049-2618. doi: 10.1186/s40168-021-01009-w.
- P. A. Jiménez, J. E. Jaimes, and J. D. Ramírez. A summary of blastocystis subtypes in north and south america. *Parasites & Vectors*, 12(1):1–9, 2019.
- S. Jongwutiwes, C. Putaporntip, T. Iwasaki, T. Sata, and H. Kanbara. Naturally acquired *Plasmodium knowlesi* malaria in human, Thailand. *Emerging infectious diseases*, 10(12):2211, 2004. Publisher: Centers for Disease Control and Prevention.
- B. Keijser, E. Zaura, S. Huse, J. van der Vossen, F. Schuren, R. Montijn, J. ten Cate, and W. Crielaard. Pyrosequencing analysis of the oral microflora of healthy adults. *Journal of Dental Research*, 87(11):1016–1020, 11 2008. ISSN 0022-0345. doi: 10.1177/154405910808701104.
- D. Kim, B. Langmead, and S. L. Salzberg. Hisat: A fast spliced aligner with low memory requirements. *Nature Methods*, 12(4):357–360, 2015. doi: 10.1038/nmeth.3317.
- A. R. Klegarth. *Landscape genetic structure, ranging patterns, and management of urban primates*. University of Notre Dame, 2015.
- A. R. Klegarth, S. A. Sanders, A. D. Gloss, K. E. Lane-deGraaf, L. Jones-Engel, A. Fuentes, and H. Hollocher. Investigating biogeographic boundaries of the

- sunda shelf: A phylogenetic analysis of two island populations of *Macaca fascicularis*. *American Journal of Physical Anthropology*, 163(4):658–670, 2017. doi: 10.1002/ajpa.23235.
- A. Klindworth, E. Pruesse, T. Schweer, J. Peplies, C. Quast, M. Horn, and F. O. Glöckner. Evaluation of general 16s ribosomal rna gene pcr primers for classical and next-generation sequencing-based diversity studies. *Nucleic acids research*, 41(1):e1–e1, 2013. doi: 10.1093/nar/gks808.
- D. Knights, T. L. Ward, C. E. McKinlay, H. Miller, A. Gonzalez, D. McDonald, and R. Knight. Rethinking “enterotypes”. *Cell host & microbe*, 16(4):433–437, 2014.
- A. Kodio, D. Coulibaly, A. K. Koné, S. Konaté, S. Doumbo, A. Guindo, F. Bittar, F. Gouriet, D. Raoult, M. A. Thera, and S. Ranque. Blastocystis colonization is associated with increased diversity and altered gut bacterial communities in healthy malian children. *Microorganisms*, 7(12):649, 12 2019. ISSN 2076-2607. doi: 10.3390/microorganisms7120649. number: 12 publisher: Multidisciplinary Digital Publishing Institute.
- A. Koliada, G. Syzenko, V. Moseiko, L. Budovska, K. Puchkov, V. Perederiy, Y. Gavalko, A. Dorofeyev, M. Romanenko, S. Tkach, L. Sineok, O. Lushchak, and A. Vaiserman. Association between body mass index and firmicutes/bacteroidetes ratio in an adult ukrainian population. *BMC Microbiology*, 17(1):120, 5 2017. ISSN 1471-2180. doi: 10.1186/s12866-017-1027-1.
- O. Kolodny, M. Weinberg, L. Reshef, L. Harten, A. Hefetz, U. Gophna, M. W. Feldman, and Y. Yovel. Coordinated change at the colony level in fruit bat fur microbiomes through time. *Nature Ecology & Evolution*, 3(1):116–124, 2019.
- I. Laforest-Lapointe and M.-C. Arrieta. Microbial eukaryotes: a missing link in gut microbiome studies. *MSystems*, 3(2):e00201–17, 2018.
- L. Lahti and S. Shetty. microbiome r package, 2012.
- K. E. Lane, C. Holley, H. Hollocher, and A. Fuentes. The anthropogenic environment lessens the intensity and prevalence of gastrointestinal parasites in balinese long-tailed macaques (*macaca fascicularis*). *Primates*, 52(2):117–128, 2011. doi: 10.1007/s10329-010-0230-6.
- K. E. Lane-deGraaf, A. Fuentes, and H. Hollocher. Landscape genetics reveal fine-scale boundaries in island populations of indonesian long-tailed macaques. *Landscape Ecology*, 29(9):1505–1519, 2014a. ISSN 1098001400690. doi: 10.1007/s10980-014-0069-0.
- K. E. Lane-deGraaf, A. Fuentes, and H. Hollocher. Landscape genetics reveal fine-scale boundaries in island populations of Indonesian long-tailed macaques. *Landscape Ecology*, 29(9):1505–1519, 2014b. ISSN 1098001400690. doi: 10.1007/s10980-014-0069-0.

- S. Lax, D. P. Smith, J. Hampton-Marcell, S. M. Owens, K. M. Handley, N. M. Scott, S. M. Gibbons, P. Larsen, B. D. Shogan, S. Weiss, et al. Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science*, 345(6200):1048–1052, 2014.
- J.-B. Leca, N. Gunst, and M. A. Huffman. Food provisioning and stone handling tradition in japanese macaques: a comparative study of ten troops. *American Journal of Primatology: Official Journal of the American Society of Primatologists*, 70(8):803–813, 2008. publisher: Wiley Online Library.
- K.-S. Lee, P. C. S. Divis, S. K. Zakaria, A. Matusop, R. A. Julin, D. J. Conway, J. Cox-Singh, and B. Singh. Plasmodium knowlesi: Reservoir Hosts and Tracking the Emergence in Humans and Macaques. *PLOS Pathogens*, 7(4):e1002015, Apr. 2011. ISSN 1553-7374. doi: 10.1371/journal.ppat.1002015. URL <https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1002015>. Publisher: Public Library of Science.
- W. Lee, T. Hayakawa, M. Kiyono, N. Yamabata, and G. Hanya. Gut microbiota composition of japanese macaques associates with extent of human encroachment. *American Journal of Primatology*, 81(12):e23072, 2019. ISSN 1098-2345. doi: 10.1002/ajp.23072. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/ajp.23072>.
- P. Legendre and E. D. Gallagher. Ecologically meaningful transformations for ordination of species data. *Oecologia*, 129(2):271–280, 10 2001. ISSN 1432-1939. doi: 10.1007/s004420100716. Company: Springer Distributor: Springer Institution: Springer Label: Springer number: 2 publisher: Springer Berlin Heidelberg.
- R. Leinonen, H. Sugawara, and M. Shumway. The sequence read archive. *Nucleic Acids Research*, 39(Database issue):D19, 1 2011. doi: 10.1093/nar/gkq1019. publisher: Oxford University Press PMID: 21062823.
- M. E. P. Lempang, F. K. Dewayanti, L. Syahrani, D. H. Permana, R. Malaka, P. B. S. Asih, and D. Syafruddin. Primate malaria: An emerging challenge of zoonotic malaria in Indonesia. *One Health*, 14:100389, June 2022. ISSN 2352-7714. doi: 10.1016/j.onehlt.2022.100389. URL <https://www.sciencedirect.com/science/article/pii/S2352771422000210>.
- M. Lepczyńska, J. Białkowska, E. Dzika, K. Piskorz-Ogórek, and J. Korycińska. Blastocystis: how do specific diets and human gut microbiota affect its development and pathogenicity? *European journal of clinical microbiology & infectious diseases*, 36(9):1531–1540, 2017.
- J. M. Leung, A. L. Graham, and S. C. Knowles. Parasite-microbiota interactions with the vertebrate gut: Synthesis through an ecological lens. *Frontiers in Microbiology*, 9(MAY):1–20, 2018. doi: 10.3389/fmicb.2018.00843.

- R. E. Ley, P. J. Turnbaugh, S. Klein, and J. I. Gordon. Human gut microbes associated with obesity. *nature*, 444(7122):1022–1023, 2006.
- R. E. Ley, M. Hamady, C. Lozupone, P. J. Turnbaugh, R. R. Ramey, J. S. Bircher, M. L. Schlegel, T. A. Tucker, M. D. Schrenzel, R. Knight, and J. I. Gordon. Evolution of mammals and their gut microbes. *Science*, 320(5883):1647–1651, 6 2008. doi: 10.1126/science.1155725. publisher: American Association for the Advancement of Science.
- R. W. Li, S. Wu, W. Li, K. Navarro, R. D. Couch, D. Hill, and J. F. Urban Jr. Alterations in the porcine colon microbiota induced by the gastrointestinal nematode *trichuris suis*. *Infection and immunity*, 80(6):2150–2157, 2012. publisher: Am Soc Microbiol.
- L. Lindheim, M. Bashir, J. Münzker, C. Trummer, V. Zachhuber, B. Leber, A. Horvath, T. R. Pieber, G. Gorkiewicz, V. Stadlbauer, and B. Obermayer-Pietsch. Alterations in gut microbiome composition and barrier function are associated with reproductive and metabolic defects in women with polycystic ovary syndrome (pcos): A pilot study. *PLoS ONE*, 12(1):1–20, 2017. ISSN 19326203. doi: 10.1371/journal.pone.0168390.
- J. Lloyd-Price, A. Mahurkar, G. Rahnavard, J. Crabtree, J. Orvis, A. B. Hall, A. Brady, H. H. Creasy, C. McCracken, M. G. Giglio, D. McDonald, E. A. Franzosa, R. Knight, O. White, and C. Huttenhower. Strains, functions and dynamics in the expanded human microbiome project. *Nature*, 550(7674):61–66, 10 2017. ISSN 1476-4687. doi: 10.1038/nature23889.
- S. Louca and M. Doebeli. Efficient comparative phylogenetics on large trees. *Bioinformatics*, 34(6):1053–1055, 3 2018. ISSN 1367-4803. doi: 10.1093/bioinformatics/btx701.
- J. E. Loudon, M. E. Howells, and A. Fuentes. The importance of integrative anthropology: A preliminary investigation employing primatological and cultural anthropological data collection methods in assessing human-monkey co-existence in bali, indonesia. *Ecological and Environmental Anthropology (University of Georgia)*, 1 2006. URL <https://digitalcommons.unl.edu/icwdmeea/26>.
- J. Ma, A. L. Prince, D. Bader, M. Hu, R. Ganu, K. Baquero, P. Blundell, R. Alan Harris, A. E. Frias, K. L. Grove, and K. M. Aagaard. High-fat maternal diet during pregnancy persistently alters the offspring microbiome in a primate model. *Nature Communications*, 5(May):1–11, 2014. doi: 10.1038/ncomms4889.
- R. Mahmud, J. Ibrahim, N. Moktar, and T.-S. Anuar. *Entamoeba histolytica* in southeast asia. In Y. A. L. Lim and I. Vythilingam, editors, *Parasites and their vectors: A special focus on Southeast Asia*, pages 103–129. Springer, Vienna, 2013. ISBN 978-3-7091-1553-4. URL https://doi.org/10.1007/978-3-7091-1553-4_6.

- E. K. Mallott and K. R. Amato. The microbial reproductive ecology of white-faced capuchins (*cebus capucinus*). *American Journal of Primatology*, 80(8):1–11, 2018. doi: 10.1002/ajp.22896.
- A. E. Mann, F. Mazel, M. A. Lemay, E. Morien, V. Billy, M. Kowalewski, A. Di Fiore, A. Link, T. L. Goldberg, S. Tecot, et al. Biodiversity of protists and nematodes in the wild nonhuman primate gut. *The ISME journal*, 14(2):609–622, 2020.
- D. Mariat, O. Firmesse, F. Levenez, V. Guimarães, H. Sokol, J. Doré, G. Corthier, and J.-P. Furet. The firmicutes/bacteroidetes ratio of the human microbiota changes with age. *BMC Microbiology*, 9(1):123, 6 2009. ISSN 1471-2180. doi: 10.1186/1471-2180-9-123.
- I. Mashima, C. F. Theodorea, B. Thaweboon, S. Thaweboon, F. A. Scannapieco, and F. Nakazawa. Exploring the salivary microbiome of children stratified by the oral hygiene index. *PLoS ONE*, 12(9):1–15, 2017. ISSN 1111111111. doi: 10.1371/journal.pone.0185274.
- G. McAlester, F. O’Gara, and J. P. Morrissey. Signal-mediated interactions between *pseudomonas aeruginosa* and *candida albicans*. *Journal of Medical Microbiology*, 57(5):563–569, 2008. ISSN 00222615. doi: 10.1099/jmm.0.47705-0.
- A. I. McCord, C. A. Chapman, G. Weny, A. Tumukunde, D. Hyeroba, K. Klotz, A. S. Koblings, D. N. M. Mborra, M. Cregger, B. A. White, S. R. Leigh, and T. L. Goldberg. Fecal microbiomes of non-human primates in western uganda reveal species-specific communities largely resistant to habitat perturbation. *American Journal of Primatology*, 76(4):347–354, 2014. ISSN 1098-2345. doi: 10.1002/ajp.22238.
- P. McKenna, C. Hoffmann, N. Minkah, P. P. Aye, A. Lackner, Z. Liu, C. A. Lozupone, M. Hamady, R. Knight, and F. D. Bushman. The macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis. *PLoS Pathogens*, 4(2), 2008. doi: 10.1371/journal.ppat.0040020.
- M. R. McLennan and M. A. Huffman. High Frequency of Leaf Swallowing and its Relationship to Intestinal Parasite Expulsion in “Village” Chimpanzees at Bulindi, Uganda. *American Journal of Primatology*, 74(7):642–650, 2012. ISSN 1098-2345. doi: 10.1002/ajp.22017. URL <http://onlinelibrary.wiley.com/doi/abs/10.1002/ajp.22017>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/ajp.22017>.
- P. McMurdie, J. and S. Holmes. phyloseq: An r package for reproducible interactive analysis and graphics of microbiome census data. *PLOS ONE*, 8(4), 4 2013. URL <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0061217>. [Online; accessed 2020-02-05].
- S. Menke, M. Meier, J. K. E. Mfunne, J. Melzheimer, B. Wachter, and S. Sommer. Effects of host traits and land-use changes on the gut microbiota of the namibian

- black-backed jackal (*canis mesomelas*). *FEMS Microbiology Ecology*, 93(11), 2017. doi: 10.1093/femsec/fix123.
- S. Mirarab, N. Nguyen, and T. Warnow. Sepp: Saté-enabled phylogenetic placement. In *Biocomputing 2012*, pages 247–258. WORLD SCIENTIFIC, 11 2011. ISBN 978-981-4596-37-4. URL https://www.worldscientific.com/doi/abs/10.1142/9789814366496_0024. DOI: 10.1142/9789814366496_0024.
- A. H. Moeller and H. Ochman. Factors that drive variation among gut microbial communities. *Gut Microbes*, 4(5):403–408, 2013.
- A. H. Moeller, P. H. Degnan, A. E. Pusey, M. L. Wilson, B. H. Hahn, and H. Ochman. Chimpanzees and humans harbour compositionally similar gut enterotypes. *Nature Communications*, 3:1175–1179, 2012. doi: 10.1038/ncomms2159.
- A. H. Moeller, M. Peeters, J. S. Ndjanga, Y. Li, B. H. Hahn, and H. Ochman. Sympatric chimpanzees and gorillas harbor convergent gut microbial communities. *Genome Research*, 23:1715–1720, 2013. doi: 10.1101/gr.154773.113.23.
- A. H. Moeller, M. Peeters, A. Ayoub, E. M. Ngole, A. Esteban, B. H. Hahn, and H. Ochman. Stability of the gorilla microbiome despite simian immunodeficiency virus infection. *Molecular Ecology*, 24(3):690–697, 2015. doi: 10.1111/mec.13057.
- B. K. Montero, N. Schwensow, M. A. Gillingham, Y. R. Ratovonamana, S. J. Rakotonjanary, V. Corman, C. Drosten, J. U. Ganzhorn, and S. Sommer. Evidence of mhc class i and ii influencing viral and helminth infection via the microbiome in a non-human primate. *PLoS pathogens*, 17(11):e1009675, 2021.
- D. Morales, N. Grahl, C. Okegbe, L. Dietrich, N. Jacobs, and D. Hogan. Control of candida albicans metabolism and biofilm formation by pseudomonas aeruginosa phenazines. *mBio*, 4(1):e00526–12, 2013. doi: 10.1128/mBio.00526-12.Editor.
- B. D. Muegge, J. Kuczynski, D. Knights, J. C. Clemente, A. González, L. Fontana, B. Henrissat, R. Knight, and J. I. Gordon. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. 1883(1998), 2010.
- B. D. Muegge, J. Kuczynski, D. Knights, J. C. Clemente, A. González, L. Fontana, B. Henrissat, R. Knight, and J. I. Gordon. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science*, 332(6032):970–974, 2011. publisher: American Association for the Advancement of Science.
- O. T. Ng, E. E. Ooi, C. C. Lee, P. J. Lee, L. C. Ng, S. W. Pei, T. M. Tu, J. P. Loh, and Y. S. Leo. Naturally acquired human Plasmodium knowlesi infection, Singapore. *Emerging infectious diseases*, 14(5):814, 2008. Publisher: Centers for Disease Control and Prevention.

- M. Nieves-Ramírez, O. Partida-Rodríguez, I. Laforest-Lapointe, L. Reynolds, E. Brown, A. Valdez-Salazar, P. Morán-Silva, L. Rojas-Velázquez, E. Morien, L. Parfrey, et al. Asymptomatic intestinal colonization with protist blastocystis is strongly associated with distinct microbiome ecological patterns. *Msystems*, 3 (3):e00007–18, 2018.
- M. C. Noverr and G. B. Huffnagle. Regulation of candida albicans morphogenesis by fatty acid metabolites. *Infection and immunity*, 72(11):6206–6210, 2004.
- H. Ochman, M. Worobey, C. H. Kuo, J. B. N. Ndjanga, M. Peeters, B. H. Hahn, and P. Hugenholtz. Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *PLoS Biology*, 8(11):3–10, 2010. doi: 10.1371/journal.pbio.1000546.
- A. J. Oksanen, R. Kindt, P. Legendre, B. O. Hara, G. L. Simpson, M. H. H. Stevens, H. Wagner, J. Oksanen, B. O'Hara, and P. M. Solymos. The vegan package. *The vegan package*, pages 190–190, 2008.
- J. Oksanen, G. L. Simpson, F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. B. O'Hara, P. Solymos, M. H. H. Stevens, E. Szoecs, H. Wagner, M. Barbour, M. Bedward, B. Bolker, D. Borcard, G. Carvalho, M. Chirico, M. D. Caceres, S. Durand, H. B. A. Evangelista, R. FitzJohn, M. Friendly, B. Furneaux, G. Hannon, M. O. Hill, L. Lahti, D. McGlinn, M.-H. Ouellette, E. R. Cunha, T. Smith, A. Stier, C. J. F. T. Braak, and J. Weedon. *vegan: Community Ecology Package*. 2022. URL <https://CRAN.R-project.org/package=vegan>.
- K. Oliphant and E. Allen-Vercoe. Macronutrient metabolism by the human gut microbiome: major fermentation by-products and their impact on host health. *Microbiome*, 7(1):91, 6 2019. ISSN 2049-2618. doi: 10.1186/s40168-019-0704-8.
- J. D. Orkin, F. A. Campos, M. S. Myers, S. E. Cheves Hernandez, A. Guadamuz, and A. D. Melin. Seasonality of the gut microbiota of free-ranging white-faced capuchins in a tropical dry forest. *ISME Journal*, 13(1):183–196, 2019a. ISSN 17517370. doi: 10.1038/s41396-018-0256-0.
- J. D. Orkin, S. E. Webb, and A. D. Melin. Small to modest impact of social group on the gut microbiome of wild costa rican capuchins in a seasonal forest. *American Journal of Primatology*, 81(10-11):e22985, 2019b.
- L. W. Parfrey, W. A. Walters, C. L. Lauber, J. C. Clemente, D. Berg-Lyons, C. Teiling, C. Kodira, M. Mohiuddin, J. Brunelle, M. Driscoll, N. Fierer, J. A. Gilbert, and R. Knight. Communities of microbial eukaryotes in the mammalian gut within the context of environmental eukaryotic diversity. *Frontiers in Microbiology*, 5: 298, 6 2014. ISSN 1664-302X. doi: 10.3389/fmicb.2014.00298.
- A. B. Pedersen, S. Altizer, M. Poss, A. A. Cunningham, and C. L. Nunn. Patterns of host specificity and transmission among parasites of wild primates. *International journal for parasitology*, 35(6):647–657, 2005. Publisher: Elsevier.

- A. C. Poole, J. K. Goodrich, N. D. Youngblut, G. G. Luque, A. Ruaud, J. L. Sutter, J. L. Waters, Q. Shi, M. El-Hadidi, L. M. Johnson, H. Y. Bar, D. H. Huson, J. G. Booth, and R. E. Ley. Human salivary amylase gene copy number impacts oral and gut microbiomes. *Cell Host & Microbe*, 25(4):553–564.e7, 4 2019. ISSN 1931-3128. doi: 10.1016/j.chom.2019.03.001.
- S. E. Pryde, S. H. Duncan, G. L. Hold, C. S. Stewart, and H. J. Flint. The microbiology of butyrate formation in the human colon. *FEMS Microbiology Letters*, 217(2):133–139, 12 2002. ISSN 0378-1097. doi: 10.1111/j.1574-6968.2002.tb11467.x.
- C. Quast, E. Pruesse, P. Yilmaz, J. Gerken, T. Schweer, P. Yarza, J. Peplies, and F. O. Glöckner. The silva ribosomal rna gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 41(D1):D590–D596, 11 2012. doi: 10.1093/nar/gks1219.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL <https://www.R-project.org/>.
- T. Ren, L. E. Grieneisen, S. C. Alberts, E. A. Archie, and M. Wu. Development, diet and dynamism: longitudinal and cross-sectional predictors of gut microbial communities in wild baboons. *Environmental microbiology*, 18(5):1312–1325, 2016.
- C. M. Riley, S. L. Jayasri, and M. D. Gumert. Results of a nationwide census of the long-tailed macaque (*macaca fascicularis*) population of singapore. *Raffles Bulletin of Zoology*, 63(December):503–515, 2015.
- C. M. Riley, A. S. DuVall-Lash, S. L. Jayasri, B. L. Koenig, A. R. Klegarth, and M. D. Gumert. How Living Near Humans Affects Singapore’s Urban Macaques. In M. T. Waller, editor, *Ethnoprimatology: Primate Conservation in the 21st Century*, Developments in Primatology: Progress and Prospects, pages 283–300. Springer International Publishing, Cham, 2016. ISBN 978-3-319-30469-4. doi: 10.1007/978-3-319-30469-4_16. URL https://doi.org/10.1007/978-3-319-30469-4_16.
- T. Roberts, D. Stark, J. Harkness, and J. Ellis. Subtype distribution of blastocystis isolates identified in a sydney population and pathogenic potential of blastocystis. *European journal of clinical microbiology & infectious diseases*, 32(3):335–343, 2013.
- T. Rognes, T. Flouri, B. Nichols, C. Quince, and F. Mahé. Vsearch: a versatile open source tool for metagenomics. *PeerJ*, 4:e2584, 10 2016. ISSN 2167-8359. doi: 10.7717/peerj.2584. PMID: 27781170 PMCID: PMC5075697.

- B. A. Rosa, T. Supali, L. Gankpala, Y. Djuardi, E. Sartono, Y. Zhou, K. Fischer, J. Martin, R. Tyagi, F. K. Bolay, et al. Differential human gut microbiome assemblages during soil-transmitted helminth infections in indonesia and liberia. *Microbiome*, 6(1):1–19, 2018.
- D. Rothschild, O. Weissbrod, E. Barkan, A. Kurilshikov, T. Korem, D. Zeevi, P. I. Costea, A. Godneva, I. N. Kalka, N. Bar, S. Shilo, D. Lador, A. V. Vila, N. Zmora, M. Pevsner-Fischer, D. Israeli, N. Kosower, G. Malka, B. C. Wolf, T. Avnit-Sagi, M. Lotan-Pompan, A. Weinberger, Z. Halpern, S. Carmi, J. Fu, C. Wijmenga, A. Zhernakova, E. Elinav, and E. Segal. Environment dominates over host genetics in shaping human gut microbiota. *Nature*, 555(7695):210–215, 2018. doi: 10.1038/nature25973.
- R. B. Sartor. Microbial influences in inflammatory bowel diseases. *Gastroenterology*, 134(2):577–594, 2 2008. doi: 10.1053/J.GASTRO.2007.11.059.
- V. Sawaswong, K. Praianantathavorn, P. Chanchaem, A. Khamwut, T. Kemthong, Y. Hamada, S. Malaivijitnond, and S. Payungporn. Comparative analysis of oral-gut microbiota between captive and wild long-tailed macaque in thailand. *Scientific Reports*, 11(1):1–13, 2021.
- P. D. Scanlan, C. R. Stensvold, M. Rajilić-Stojanović, H. G. Heilig, W. M. De Vos, P. W. O’Toole, and P. D. Cotter. The microbial eukaryote blastocystis is a prevalent and diverse member of the healthy human gut microbiota. *FEMS Microbiology Ecology*, 90(1):326–330, 2014. doi: 10.1111/1574-6941.12396.
- J. M. Schurer, V. Ramirez, P. Kyes, T. Tanee, N. Patarapadungkit, P. Thamsenunpapp, S. Trufan, E. T. Grant, G. Garland-Lewis, S. Kelley, H. Nueaitong, R. C. Kyes, and P. Rabinowitz. Long-Tailed Macaques (*Macaca fascicularis*) in Urban Landscapes: Gastrointestinal Parasitism and Barriers for Healthy Coexistence in Northeast Thailand. *The American Journal of Tropical Medicine and Hygiene*, 100(2):357–364, Feb. 2019. ISSN 0002-9637. doi: 10.4269/ajtmh.18-0241. URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6367618/>.
- A. M. Seekatz, A. Panda, D. A. Rasko, F. R. Toapanta, E. A. Elloe-Fadrosh, A. Q. Khan, Z. Liu, S. T. Shipley, L. J. DeTolla, M. B. Sztein, and C. M. Fraser. Differential response of the cynomolgus macaque gut microbiota to shigella infection. *PLOS ONE*, 8(6):e64212, 6 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0064212. publisher: Public Library of Science.
- N. Segata, J. Izard, L. Waldron, D. Gevers, L. Miropolsky, W. S. Garrett, and C. Huttenhower. Metagenomic biomarker discovery and explanation. *Genome Biology*, 12(6), 2011. doi: 10.1186/gb-2011-12-6-r60.
- A. Septembre-Malaterre, F. Remize, and P. Poucheret. Fruits and vegetables, as a source of nutritional compounds and phytochemicals: Changes in bioactive

- compounds during lactic fermentation. *Food Research International*, 104:86–99, Feb. 2018. ISSN 0963-9969. doi: 10.1016/j.foodres.2017.09.031. URL <https://www.sciencedirect.com/science/article/pii/S0963996917305987>.
- J. Sha, M. Gumert, B. Lee, A. Fuentes, S. Rajathurai, S. Chan, and L. Jones-Engel. Status of the long-tailed macaque *macaca fascicularis* in singapore and implications for management. *Biodiversity and Conservation*, 18(11):2909–2926, 2009a.
- J. C. M. Sha and G. Hanya. Diet, activity, habitat use, and ranging of two neighboring groups of food-enhanced long-tailed macaques (*macaca fascicularis*). *American Journal of Primatology*, 75(6):581–592, 2013.
- J. C. M. Sha, M. D. Gumert, B. P.-H. Lee, L. Jones-Engel, S. Chan, and A. Fuentes. Macaque–human interactions and the societal perceptions of macaques in singapore. *American Journal of Primatology: Official Journal of the American Society of Primatologists*, 71(10):825–839, 2009b.
- J. C. M. Sha, M. D. Gumert, B. P. Y.-H. Lee, A. Fuentes, S. Rajathurai, S. Chan, and L. Jones-Engel. Status of the long-tailed macaque *Macaca fascicularis* in Singapore and implications for management. *Biodiversity and Conservation*, 18(11):2909–2926, Oct. 2009c. ISSN 1572-9710. doi: 10.1007/s10531-009-9616-4. URL <https://doi.org/10.1007/s10531-009-9616-4>.
- S. Shano, A. Islam, E. Hagan, M. K. Rostal, S. Martinez, A. Al Shakil, M. Hasan, L. Francisco, M. M. Husain, M. Rahman, and others. Environmental Change and Zoonotic Disease Risk at Human-Macaque Interfaces in Bangladesh. *Eco-Health*, 18(4):487–499, 2021. Publisher: Springer.
- G. Shukla, R. Bhatia, and A. Sharma. Prebiotic inulin supplementation modulates the immune response and restores gut morphology in giardia duodenalis-infected malnourished mice. *Parasitology Research*, 115(11):4189–4198, 11 2016. ISSN 1432-1955. doi: 10.1007/s00436-016-5196-x.
- M. S. Singapore. Climate of Singapore. Technical report. URL <http://www.weather.gov.sg/climate-climate-of-singapore/>.
- B. Singh, L. K. Sung, A. Matusop, A. Radhakrishnan, S. S. Shamsul, J. Cox-Singh, A. Thomas, and D. J. Conway. A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings. *The Lancet*, 363(9414):1017–1024, 2004. Publisher: Elsevier.
- M. Singh, K. Suresh, L. C. Ho, G. C. Ng, and E. H. Yap. Elucidation of the life cycle of the intestinal protozoan *blastocystis hominis*. *Parasitology Research*, 81(5):446–450, 1995. doi: 10.1007/BF00931510.

- R. K. Singh, H.-W. Chang, D. Yan, K. M. Lee, D. Ucmak, K. Wong, M. Abrouk, B. Farahnik, M. Nakamura, T. H. Zhu, et al. Influence of diet on the gut microbiome and implications for human health. *Journal of translational medicine*, 15 (1):1–17, 2017.
- J. Šlapeta, S. E. Dowd, A. D. Alanazi, M. E. Westman, and G. K. Brown. Differences in the faecal microbiome of non-diarrhoeic clinically healthy dogs and cats associated with giardia duodenalis infection: impact of hookworms and coccidia. *International Journal for Parasitology*, 45(9-10):585–594, 2015.
- H. Y. Sohn, K. H. Son, C. S. Kwon, G. S. Kwon, and S. S. Kang. Antimicrobial and cytotoxic activity of 18 prenylated flavonoids isolated from medicinal plants: *Morus alba* L., *Morus mongolica* Schneider, *Broussnetia papyrifera* (L.) Vent, *Sophora flavescens* Ait and *Echinosophora koreensis* Nakai. *Phytomedicine*, 11(7):666–672, Nov. 2004. ISSN 0944-7113. doi: 10.1016/j.phymed.2003.09.005. URL <https://www.sciencedirect.com/science/article/pii/S094471130400090X>.
- H. Sokol, B. Pigneur, L. Watterlot, O. Lakhdari, L. G. Bermúdez-Humarán, J.-J. Gratadoux, S. Blugeon, C. Bridonneau, J.-P. Furet, G. Corthier, C. Grangette, N. Vasquez, P. Pochart, G. Trugnan, G. Thomas, H. M. Blottière, J. Doré, P. Marteau, P. Seksik, and P. Langella. *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 (43):16731–16736, 10 2008. doi: 10.1073/pnas.0804812105. publisher: Proceedings of the National Academy of Sciences.
- F. Sommer, N. Adam, M. E. Johansson, L. Xia, G. C. Hansson, and F. Bäckhed. Altered mucus glycosylation in core 1 o-glycan-deficient mice affects microbiota composition and intestinal architecture. *PloS one*, 9(1):e85254, 2014.
- D. Steinley. Properties of the hubert-arable adjusted rand index. *Psychological methods*, 9(3):386, 2004. publisher: American Psychological Association.
- C. R. Stensvold and C. G. Clark. Current status of blastocystis: A personal view. *Parasitology International*, 65(6):763–771, 2016. doi: 10.1016/j.parint.2016.05.015.
- C. R. Stensvold and M. van der Giezen. Associations between gut microbiota and common luminal intestinal parasites. *Trends in Parasitology*, 34(5):369–377, 5 2018. ISSN 1471-4922. doi: 10.1016/j.pt.2018.02.004.
- C. R. Stensvold, B. A. Sørland, R. P. Berg, L. O. Andersen, M. van der Giezen, J. L. Bowtell, A. A. El-Badry, S. Belkessa, Ö. Kurt, and H. V. Nielsen. Stool microbiota diversity analysis of blastocystis-positive and blastocystis-negative individuals. *Microorganisms*, 10(2):326, 2022.

- I. Strużycka. The oral microbiome in dental caries. *Polish journal of microbiology*, 63(2):127, 2014.
- L. E. Tailford, E. H. Crost, D. Kavanaugh, and N. Juge. Mucin glycan foraging in the human gut microbiome. *Frontiers in genetics*, 6:81, 2015.
- K. S. Tan, H. Mirza, J. D. Teo, B. Wu, and P. A. MacAry. Current views on the clinical relevance of blastocystis spp. *Current infectious disease reports*, 12(1): 28–35, 2010.
- T. C. Tan and K. G. Suresh. Amoeboid form of blastocystis hominis—a detailed ultrastructural insight. *Parasitology Research*, 99(6):737–742, 11 2006. ISSN 1432-1955. doi: 10.1007/s00436-006-0214-z.
- D. Tasdemir, A. J. MacIntosh, P. Stergiou, M. Kaiser, N. R. Mansour, Q. Bickle, and M. A. Huffman. Antiprotozoal and antihelminthic properties of plants ingested by wild japanese macaques (macaca fuscata yakui) in yakushima island. *Journal of ethnopharmacology*, 247:112270, 2020.
- B. Thierry. The macaques: a double-layered social organization. *Primates in perspective*, page 224–240, 2007. publisher: Oxford University Press.
- L. R. Thompson, J. G. Sanders, D. McDonald, A. Amir, J. Ladau, K. J. Locey, R. J. Prill, A. Tripathi, S. M. Gibbons, G. Ackermann, et al. A communal catalogue reveals earth’s multiscale microbial diversity. *Nature*, 551(7681):457–463, 2017.
- R. Y. Tito, S. Chaffron, C. Caenepeel, G. Lima-Mendez, J. Wang, S. Vieira-Silva, G. Falony, F. Hildebrand, Y. Darzi, L. Rymenans, C. Verspecht, P. Bork, S. Vermeire, M. Joossens, and J. Raes. Population-level analysis of blastocystis subtype prevalence and variation in the human gut microbiota. *Gut*, pages 1–10, 2018. ISSN 14683288. doi: 10.1136/gutjnl-2018-316106.
- P. Trosvik, E. J. de Muinck, E. K. Rueness, P. J. Fashing, E. C. Beierschmitt, K. R. Callingham, J. B. Kraus, T. H. Trew, A. Moges, A. Mekonnen, V. V. Venkataraman, and N. Nguyen. Multilevel social structure and diet shape the gut microbiota of the gelada monkey, the only grazing primate. *Microbiome*, 6(1):84, 5 2018. ISSN 2049-2618. doi: 10.1186/s40168-018-0468-6.
- J. Tung, L. B. Barreiro, M. B. Burns, J.-C. Grenier, J. Lynch, L. E. Grieneisen, J. Altmann, S. C. Alberts, R. Blekhman, and E. A. Archie. Social networks predict gut microbiome composition in wild baboons. *elife*, 4:e05224, 2015.
- P. J. Turnbaugh, R. E. Ley, M. A. Mahowald, V. Magrini, E. R. Mardis, and J. I. Gordon. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*, 444(7122):1027–1031, 12 2006. doi: 10.1038/nature05414.
- K. Vaisusuk, W. Saijuntha, S. Sedlak, T. Thanchomnang, W. Pilap, W. Suksavate, C. R. Stensvold, and C. Tantrawatpan. Blastocystis subtypes detected in long-tailed macaques in thailand—further evidence of cryptic host specificity. *Acta*

- Tropica*, 184:78–82, 8 2018. ISSN 0001-706X. doi: 10.1016/j.actatropica.2017.09.002.
- A. Valentini, P. Taberlet, C. Miaud, R. Civade, J. Herder, P. F. Thomsen, E. Bellemain, A. Besnard, E. Coissac, F. Boyer, C. Gaboriaud, P. Jean, N. Poulet, N. Roset, G. H. Copp, P. Geniez, D. Pont, C. Argillier, J.-M. Baudoin, T. Peroux, A. J. Crivelli, A. Olivier, M. Acqueberge, M. Le Brun, P. R. Møller, E. Willerslev, and T. Dejean. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, 25(4):929–942, 2016. ISSN 1365-294X. doi: 10.1111/mec.13428. URL <http://onlinelibrary.wiley.com/doi/abs/10.1111/mec.13428>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/mec.13428>.
- J. D. Van Bleijswijk, C. Whalen, G. C. Duineveld, M. S. Lavaleye, H. J. Witte, and F. Mienis. Microbial assemblages on a cold-water coral mound at the se rock-all bank (ne atlantic): Interactions with hydrography and topography. *Biogeosciences*, 12(14):4483–4496, 2015. doi: 10.5194/bg-12-4483-2015.
- M. J. Vavrek. Fossil: palaeoecological and palaeogeographical analysis tools. *Palaeontologia electronica*, 14(1):16, 2011.
- W. Wang, Y. Zu, Y. Fu, and T. Efferth. In Vitro Antioxidant and Antimicrobial Activity of Extracts from *Morus alba* L. Leaves, Stems and Fruits. *The American Journal of Chinese Medicine*, 40(02):349–356, Jan. 2012. ISSN 0192-415X. doi: 10.1142/S0192415X12500279. URL <https://www.worldscientific.com/doi/abs/10.1142/S0192415X12500279>. Publisher: World Scientific Publishing Co.
- I. Wawrzyniak, P. Poirier, E. Viscogliosi, M. Dionigia, C. Texier, F. Delbac, and H. E. Alaoui. Blastocystis, an unrecognized parasite: an overview of pathogenesis and diagnosis. *Therapeutic advances in infectious disease*, 1(5):167–178, 2013.
- B. Wheatley, R. Stephenson, H. Kurashina, and K. Marsh-Kautz. A cultural primatological study of macaca fascicularis on ngeaur island, republic of palau. *CAMBRIDGE STUDIES IN BIOLOGICAL AND EVOLUTIONARY ANTHROPOLOGY*, page 240–253, 2002. publisher: Cambridge University Press.
- B. P. Wheatley. *Sacred monkeys of Bali*. Waveland Press, c., 1999.
- J. J. Wilcox and H. Hollocher. Unprecedented symbiont eukaryote diversity is governed by internal trophic webs in a wild non-human primate. *Protist*, 169(3): 307–320, 2018. doi: 10.1016/j.protis.2018.03.001.
- J. J. Wilcox, K. E. Lane-Degraaf, A. Fuentes, and H. Hollocher. Comparative community-level associations of helminth infections and microparasite shedding in wild long-tailed macaques in bali, indonesia. *Parasitology*, 760(2015):480–489, 2014. doi: 10.1017/S0031182014001462.

- G. D. Wu, J. Chen, C. Hoffmann, K. Bittinger, Y.-y. Chen, S. A. Keilbaugh, M. Bewtra, D. Knights, W. A. Walters, R. Knight, R. Sinha, E. Gilroy, K. Gupta, R. Baldassano, L. Nessel, and H. Li. Linking long-term dietary patterns with gut microbial enterotypes. *Science*, 334(October):105–109, 2011. ISSN 0036-8075. doi: 10.1126/science.1208344.
- Y. Xu, Y. Zhu, X. Li, and B. Sun. Dynamic balancing of intestinal short-chain fatty acids: The crucial role of bacterial metabolism. *Trends in Food Science & Technology*, 100:118–130, 6 2020. ISSN 0924-2244. doi: 10.1016/j.tifs.2020.02.026.
- Y. Ye and T. G. Doak. A parsimony approach to biological pathway reconstruction/inference for genomes and metagenomes. *PLOS Computational Biology*, 5(8):e1000465, 8 2009. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1000465. publisher: Public Library of Science.
- C. P. Yeager. Feeding ecology of the long-tailed macaque(*macaca fascicularis*) in kalimantan tengah, indonesia. *International Journal of Primatology*, 17(1): 51–62, 2 1996. ISSN 1573-8604. doi: 10.1007/BF02696158.
- S. Yildirim, C. J. Yeoman, M. Sipos, M. Torralba, B. A. Wilson, T. L. Goldberg, R. M. Stumpf, S. R. Leigh, B. A. White, and K. E. Nelson. Characterization of the fecal microbiome from non-human wild primates reveals species specific microbial communities. *PLOS ONE*, 5(11):e13963, 11 2010. ISSN 1932-6203. doi: 10.1371/journal.pone.0013963.
- P. Yilmaz, L. W. Parfrey, P. Yarza, J. Gerken, E. Pruesse, C. Quast, T. Schweer, J. Peplies, W. Ludwig, and F. O. Glöckner. The silva and “all-species living tree project (ltp)” taxonomic frameworks. *Nucleic Acids Research*, 42(D1):D643–D648, 1 2014. ISSN 0305-1048. doi: 10.1093/nar/gkt1209.
- S. Yooseph, E. F. Kirkness, T. M. Tran, D. M. Harkins, M. B. Jones, M. G. Torralba, E. O’Connell, T. B. Nutman, S. Doumbo, O. K. Doumbo, B. Traore, P. D. Crompton, and K. E. Nelson. Stool microbiota composition is associated with the prospective risk of plasmodium falciparum infection. *BMC Genomics*, 16(1): 631, 8 2015. ISSN 1471-2164. doi: 10.1186/s12864-015-1819-3.
- E. Zaura, B. J. Keijser, S. M. Huse, and W. Crielaard. Defining the healthy” core microbiome” of oral microbial communities. *BMC microbiology*, 9(1):1–12, 2009.
- X. Zhang, K. A. Kadir, L. F. Quintanilla-Zariñan, J. Villano, P. Houghton, H. Du, B. Singh, and D. G. Smith. Distribution and prevalence of malaria parasites among long-tailed macaques (*Macaca fascicularis*) in regional populations across Southeast Asia. *Malaria Journal*, 15(1):450, Sept. 2016. ISSN 1475-2875. doi: 10.1186/s12936-016-1494-0. URL <https://doi.org/10.1186/s12936-016-1494-0>.

N. Zmora, J. Suez, and E. Elinav. You are what you eat: diet, health and the gut microbiota. *Nature reviews Gastroenterology & hepatology*, 16(1):35–56, 2019.