UNIVERSITY OF CALIFORNIA

Santa Barbara

Microbial Relationships in Lignocellulolysis by Herbivore Gut Consortia

A dissertation submitted in partial satisfaction of the

requirements for the degree

Doctor of Philosophy

in

Molecular, Cellular, and Developmental Biology

by

Katharine Leah Dickson

Committee in charge:

Professor Michelle O'Malley, Co-Chair

Professor David Low, Co-Chair

Professor Christopher Hayes

Professor Christopher Lawson, University of Toronto

September 2023

The dissertation of Katharine Leah Dickson is approved.

_____

Professor Christopher Hayes

_____

Professor Christopher Lawson, University of Toronto

_____

Professor David Low, Committee Co-Chair

_____

Professor Michelle O'Malley, Committee Co-Chair

July 2023

Microbial Relationships in Lignocellulolysis by Herbivore Gut Consortia

DEDICATION

To those who can't see the light at the end of the tunnel yet: keep moving.

ACKNOWLEDGMENTS

My time at UCSB was nothing less than a rebirth. Those of you who knew me well are aware of the fact that what I went through in the past seven years was not just a professional and scientific, but personal and spiritual journey of transformation, healing, growth, and self-actualization. I would not have finished this degree without the support of a small village worth of people. All of them have my eternal gratitude. A piece of my heart will always remain in Santa Barbara.

First and foremost I thank my doctoral advisor, Professor Michelle O'Malley. I came to her laboratory after a particularly rough first three years of graduate school, at a time when I deeply doubted the viability of my future as a scientist and had reached the lowest point in my graduate career. Enrolling in Omics-Enabled Biotechnology and meeting her was one of the best decisions I made in my graduate career. How fortunate I was to find such a brilliant, kind, innovative mentor. Over the course of four years, she gave me both the space and the guidance to find my footing again, and to grow, thrive, and truly come into my own as a scientist. Most graduate students do not generally get to claim that their advisors saved their lives (because without going into too much detail, she did), or that they wrote a sizable chunk of their dissertation while briefly living with their advisor after being displaced in an apartment fire, but fortunately, I do.

I thank my committee co-chair, Prof. David Low, and my other committee members, Prof. Christopher Hayes and Prof. Christopher Lawson, as well as all of my labmates. First I thank Profs. Xuefeng "Nick" Peng and Christopher Lawson, who taught me my craft when they were postdoctoral researchers in the O'Malley lab. I thank Dr. Susanna Seppala, our

lab's project scientist, for being an important source of both scientific advice and moral support for me and every other graduate student and postdoc who comes through the O'Malley lab. I thank Dr. Tom Lankiewicz, Dr. St. Elmo Wilken, Dr. Candice Swift, and Dr. Jennifer Brown (and Drs. Sean Gilmore and John Henske, who left before I joined the lab), those of my labmates who performed the initial collection of the data from which Chapter 2 of this dissertation is derived; I also thank Elaina Blair, who collected HPLC data for this chapter. I thank Shiyan "Shirley" Jin, who assisted with the attempts to isolate anaerobic fungi from primate feces in Chapter 3, as well as Colleen Ahern, who provided code tips. I thank Taylor Gierke, one of our lab's undergraduates, for giving me the opportunity to pass my skills on to the next generation of graduate students before I defend my PhD and move on to bigger things.

I thank the staff at the UCSB California NanoSystems Institute's Center for Scientific Computing with whom I worked most closely during my PhD: Dr. Paul Weakliem, Nathan "Fuzzy" Rogers, and Yu-Chieh "Jay" Chi. I harassed Paul and Fuz multiple times for computing help and bribed them with Greek shortbread cookies so they would continue to humor me, and with their help for several years, I turned mountains of letters into scientific insights. Jay joined them in the end to help me finish strong.

I thank our collaborators at the Santa Barbara Zoo, Cincinnati Zoo, and Joint Genome Institute. At the Cincinnati Zoo, I thank Dr. Erin Curry, Dr. Victoria McGee, Barbara Henry, and Mary Noell. At the Santa Barbara Zoo, I thank Jennifer Kennedy, Sara Neumann, Amy Heath, Julie Barnes, and Trent Barnhart (and I hope, in some small way, that I helped them fix what ails Nzinga). At the Joint Genome Institute, I thank Dr. Asaf Salamov, Dr. Kerrie Barry, and Dr. Igor Grigoriev.

I thank the animals whose lives and microbiota supported my PhD: Anza, Octavius, Maximus, Phoebe, Valentina, and Tiberius the colobus monkeys at the Cincinnati Zoo, Nzinga and Bangori the gorillas at the Santa Barbara Zoo, and Elway the San Clemente goat at the Santa Barbara Zoo. May their contributions to science, as well as other zoo animals' contributions to science, be repaid with better futures for their species.

I have accumulated a small legion of friends and other allies whose steadfast care buoyed me through the last seven years. My roommates through most of graduate school, Matthew Fritzler, Dr. Sharad Shankar, and Mychael Gomez, were steadfast presences for me through both good times and some of the darkest periods of my graduate career, including, but not limited to, the pandemic and the fire that ultimately lost us our apartment. In no particular order, I also thank the following friends and other comrades: Nolan Anderson, Jacqueline Clark, Roman Aguilera, Harrison Tasoff, Dr. Jared Goldberg, Dr. Jakkarin Limwongyut, Mark Arildsen, Adolfo Holguin, Dr. Abe Pressman, Dr. Angela Zhang, Dr. Carol Tsai-Bochkov, Dr. Elizabeth Decolvenaere, Dr. Meghna Soni, Dr. Scott Fenton, Dr. Pratik Raghu, Sean Pierce, Alon Avidor, Charlotte Mountain, Skyler Johnson, Dr. Cliff Hardy, Dr. Amy Eisenberg, Zoe Welch, Payton DeMarzo, Becky Martin, April Amante, Baker, Kristen Thomas McGill, An Bui, Misa Nguyen, Janna Haider, Steven Rizzie, and Raj Chaklashiya.

I thank Professor O'Malley's family, Professor Matthew Helgeson and their children Elliott, Lucas, and Mia Helgeson (as well as their cat Leo, my dissertation-writing companion and unofficial fifth committee member – I have the honor of having the third dissertation this cat has supervised), who joined her in generously opening their home to me for three weeks after I lost mine. I thank my family in Kansas City: Joan Kelly, Philip

Benton, Dr. Michaela Steinmetz-Benton, Finn Benton, and many others. Others I would also like to thank are Keith Higginbotham, Jacqueline Henretig, Hope Andreason, Prof. Dan Morse, and Prof. Kathy Foltz.

This dissertation owes a debt in a much broader sense to the microbiota of the world, from which I have taken a variety of personal and spiritual lessons, and a great deal of inspiration and comfort, in addition to the scientific and professional ones I detail in the rest of this work. Fungi, nature's great decomposers, are powerful symbols of persistence, transmutation, and renewal. The fungi our lab studies take dead matter and turn it into something life-giving and usable that sustains the community around them. Microbiomes represent one of nature's purest expressions of the power of diversity and community, and host-associated microbiota represent how the fate of the biggest things is ultimately intertwined with that of the smallest. Incredible things come from the cooperation of a community of folks with many different capabilities. I have found healing in considering microbes and microbiota as symbols of community, connection, power, and thriving. I am stronger for being part of a flourishing, diverse scientific community, and I am who I am with the help and kindness of all those who have nurtured me as a scientist and as a person through the last seven years. Together, we have produced so much that we never could alone.

VITA OF KATHARINE LEAH DICKSON
July 2023

EDUCATION

Doctor of Philosophy in Molecular, Cellular, and Developmental Biology, University of California, Santa Barbara, July 2023
Bachelor of Science in Biology, George Mason University, June 2014 (cum laude)

PROFESSIONAL EMPLOYMENT

2016-2019: Teaching Assistant, Department of Molecular, Cellular, and Developmental Biology, University of California, Santa Barbara

PUBLICATIONS

E.M. Blair, **K.L. Dickson**, M.A. O'Malley, "Microbial communities and their enzymes facilitate degradation of recalcitrant polymers in anaerobic digestion," *Current Opinion in Microbiology* 64, 100-108, 2021.
R.A. Crook, **K. Dickson**, R.T. Hanlon, E.T. Walters, "Nociceptive sensitization reduces predation risk," *Current Biology* 24(10), 1121-1125, 2014.

PLANNED PUBLICATIONS

K.L. Dickson, E.M. Blair, X. Peng, S.P. Gilmore, J.K. Henske, C.L. Swift, K. Barry, I.V. Grigoriev, M.K. Theodorou, M.A. O'Malley, " Lignocellulose complexity shapes the transcriptional activity and community function of goat gut microbiota," In progress.
K.L. Dickson, T. Gierke, S. Jin, M.A. O'Malley, " Enrichment of fecal samples of captive *Colobus guereza* and *Gorilla gorilla gorilla* reveals a diverse community of lignocellulolytic prokaryotes and fungi with potential key roles in herbivory and health," In progress.
X. Peng, **K.L. Dickson**, E.M. Blair, S.P. Gilmore, J.K. Henske, C.L. Swift, K. Barry, I.V. Grigoriev, M.K. Theodorou, D.L. Valentine, M.A. O'Malley, "Genome-centric metatranscriptomic analyses of fungal and bacterial consortia that enable lignocellulose breakdown in goat gut microbiome", In progress.
C.L. Swift, L.V. Butkovich, **K.L. Dickson**, C.E.Lawson, X. Peng, K. Barry, M.A. O'Malley, "Fungal and bacterial secondary metabolites shape consortia membership in an enriched goat fecal microbiome," In progress.

PRESENTATIONS

K.L. Dickson, E.M. Blair, X. Peng, S.P. Gilmore, J.K. Henske, C.L. Swift, K. Barry, I.V. Grigoriev, M.K. Theodorou, M.A. O'Malley, " Genome-centric metatranscriptomics reveals limits to the capacity of substrate to shape lignocellulolytic gut enrichment consortia," US Dept. of Energy Joint Genome Institute User Meeting 2022, Berkeley, CA, August 29-31.
K.L. Dickson, M.A. O'Malley, " Characterizing the role of anaerobic fungi in lignocellulolytic microbial communities and the gut mycobiome of herbivorous non-human primates," Fungal Genetics 2022, Pacific Grove, CA, March 15-20.

FIELDS OF STUDY

Microbial Ecology, Microbiome Engineering, Herbivore Gut Microbiota
Supervisor: Professor Michelle O'Malley

ABSTRACT

Microbial Relationships in Lignocellulolysis by Herbivore Gut Consortia

by

Katharine Leah Dickson

Microbiota in the digestive tracts of herbivores are complex, highly-structured communities that degrade lignocellulosic biomass into simple sugars and then metabolize it into fatty acids, gases, and other small molecules, approximately three times as efficiently as the most advanced human-engineered processes in use (Godon et al., 2013). The degradative power of these communities arises from partnerships between organisms carrying out complementary metabolic tasks, which together constitute the processes of anaerobic digestion. Cooperative strategies arising from these partnerships that microbiota deploy during anaerobic digestion of complex substrates are not well understood. A fuller understanding of these strategies will shed light on underlying factors driving the division of labor at each stage of lignocellulolysis and aid in identifying genomic and transcriptional mechanisms that predict patterns of labor division in a given consortium. Applying this knowledge will enable the development of approaches to modulate the activity of existing microbiota and construct synthetic consortia to optimally transform complex substrates into desired products.

In this body of work, we examine microbial relationships and consequent digestive strategies in the degradation of complex biomass as they manifest through both composition and activity, through employing enrichment to study the gut microbiota of captive

individuals of three herbivore species – the San Clemente Island goat (*Capra hircus*), a foregut-fermenting hoofed mammal, the Eastern black and white colobus monkey (*Colobus guereza*), a foregut-fermenting primate, and the Western lowland gorilla (*Gorilla gorilla gorilla*), a hindgut-fermenting primate. In the first of two studies, we enriched a goat fecal microbiome in parallel cultures on three lignocellulosic substrates and hemicellulose, and then sequenced the actively transcribed genes from the microbes enriched on each substrate. We demonstrated that parallel enrichment on complex biomass subsrates and purified hemicellulose enriches consortia whose transcriptional activity is shaped by the substrate. However, this activity was likely limited by genetically-encoded factors inherent to the consortium. We also observed a previously-uncharacterized strategy for division of carbohydrate-active enzyme expression labor exhibited across all lignocellulosic substrates. We then investigated the composition and structure of lignocellulolytic communities in *C. guereza* and *G. gorilla gorilla* fecal microbiota, and attempted to enrich anaerobic fungi of the phylum Neocallimastigomycota from fecal samples of captive *C. guereza* and *G. gorilla gorilla.* This attempt was unsuccessful at enriching anaerobic fungi, but instead enriched a diverse community of other fungal taxa with diverse roles in lignocellulolysis, metabolism, and regulation of the microbial community and its host. We found that the *C. guereza* and *G. gorilla gorilla* gut microbiome contain rare bacterial and fungal taxa that may have important roles in lignocellulolysis in the gut, and that enrichment on complex substrates with the aid of antibiotics is an underutilized but crucial method for fully characterizing the diversity of herbivore gut microbiota by enabling the growth and sequencing of rare taxa in the microbiome. Taken together, these studies investigate relationships between microbes in

lignocellulolytic gut consortia, through the lenses of community composition and activity, to yield new insights about how these consortia work together.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

1. Introduction

1.1. Motivation

The digestive tract is nature's most exquisite bioreactor. Microbiomes in the digestive tracts of herbivores degrade and metabolize plant biomass in the process of anaerobic digestion to supply nutrients to their animal hosts. Microbes in these communities employ mechanical and chemical tools to break down biomass, which, together, they do much more efficiently than human-engineered methods of degradation (Bayané & Guiot, 2011; Deublein & Steinhauser, 2011; Godon et al., 2013). This incredible degradative activity makes these microbiomes, and isolates and consortia enriched from them, attractive targets for the development of biotechnological applications. Central to this activity are the relationships between microbes involved in anaerobic digestion in the herbivore gut. Insight into how these relationships are shaped in nature reveals additional avenues to pursue in designing and controlling the formation and maintenance of engineered consortia, and in leveraging the unique metabolic capabilities of the many residents of the herbivore gut.

We understand little about how complex substrates shape enriched consortia. Animal studies cannot remove the effects of the host upon both substrate and microbiome, and small, defined consortia and simple substrates fail to capture more complex organization of metabolic activity that may happen when organisms work in concert to consume substrates like those they consume in nature. The discovery and characterization of potential emergent phenomena of organized microbial activity demands an approach accommodating complexity. Moreover, the development of a variety of effective and sustainable solutions to open problems in technological sectors as diverse as energy, pharmaceuticals, and

bioproduction demands that we look to what nature has already developed for the purpose; nature's oldest bioreactors still do these jobs better than anything humans have invented.

Attempts to understand these relationships within the gut microbiota of herbivorous non-human primates (NHPs), particularly relationships between prokaryotes and gut fungal populations that play a key role in biomass degradation in many herbivores (Liggenstoffer et al., 2010; Murphy et al., 2019; Wang et al., 2019), bring the study of these complex communities evolutionarily closer to home. The human gut microbiome evolved alongside us as we diverged from our fellow primates, who consume far more fiber on average than we do (Milton, 1999). Exploring the biomass-degrading capacity of NHP gut microbiomes opens up new research avenues in pursuit of understanding the evolutionary trajectory of herbivory and other dietary strategies in primates and the emergence of humans' unique microbiome and dietary preferences, as well as development of strategies to modulate human microbiota. These studies also play an important role in conservation. Many NHP species are endangered or vulnerable due to invasion and fragmentation of their habitat by humans, and there is evidence that their gut microbiota are disrupted as a result (Amato et al., 2013; Barelli et al., 2020; Estrada et al., 2017; IUCN, 2022). NHPs are also vulnerable to morbidity and mortality from gastric issues in captivity, and evidence suggests captivity de-diversifies and humanizes NHP gut microbiota (Amato et al., 2016; Clayton et al., 2016; Hale et al., 2019; McKenzie et al., 2017; Strong et al., 2016; Sun et al., 2021).

1.2. Degradation of lignocellulose by gut consortia from terrestrial herbivores

1.815 billion tons of plant biomass, one of Earth's most abundant renewable natural resources, is produced every year (Bilal et al., 2021). **Lignocellulose** makes up most of the plant secondary cell wall, and approximately 30-50% of whole plant dry mass (Zhang et al.,

2019). Lignocellulose is made of three polymers: **lignin**, a highly crosslinked heteropolymer of aromatic alcohols; **cellulose**, a linear homopolymer of $\beta(1\rightarrow4)$-linked D-glucose; and **hemicellulose**, a branched heteropolymer of five- and six-carbon sugars (Pérez et al., 2002). The lignin, cellulose, and hemicellulose molecules in a secondary cell wall are in turn linked to each other with covalent and non-covalent bonds (Nishimura et al., 2018). The complex structure of lignocellulose requires the concerted activity of a diverse, complex community of microbes in the herbivore gut to fully degrade, and to liberate enough short-chain fatty acids to nourish the animal host, in a process called **anaerobic digestion** (Bayané & Guiot, 2011).

1.2.1. Anaerobic digestion

Anaerobic digestion is a multi-stage degradation of biomacromolecules into carboxylic acids, alcohols, gases, and other small molecules in the absence of oxygen, involving a complex network of metabolic pathways (Blair et al., 2021; Yadav et al., 2022). A variety of auxiliary processes, such as sulfate and nitrate reduction, occur alongside anaerobic digestion to dispose of other metabolites the process generates (Novaes, 1986). In addition to the gut lumen, this process takes place in a variety of other environments on Earth, such as some soil and marine sediments, as well as industrial anaerobic digesters that exploit this natural process to perform services such as municipal waste degradation and wastewater treatment. Given this dissertation's focus on lignocellulose degradation, we will describe the process of anaerobic digestion of polysaccharides, but these stages occur for all biomacromolecules.

In the first stage, **hydrolysis**, a polysaccharide is degraded into monosaccharides and other small-molecule compounds. Hydroxyl positioning, anomeric configurations, and

different ring sizes allow for **1.05 x 10^{12} possible structures** for a molecule containing **only six linked monosaccharides** (Laine, 1994). Industrial lignocellulose degradation methods that do not employ a microbiome instead hydrolyze biomass with high pressure and temperature, sometimes after pre-treatment, but this generates numerous side reactions and offers little control over the products of the process (Yu et al., 2021). In environments with a microbiome, polysaccharides are degraded in a much more controlled fashion by **carbohydrate-active enzymes (CAZymes)**. CAZymes are structurally and functionally diverse, reflecting the myriad of possible complex polysaccharides they have evolved to degrade. The CAZyme classification system, created by Henrissat *et al*. in 1989 and expanded into today's CAZy Database, sorts CAZymes into six modules by catalytic activity, with the number of currently recognized families thereof included next to the module type in parentheses: glycoside hydrolase (173), glycosyltransferase (116), polysaccharide lyase (42), carbohydrate esterase (20), auxiliary activity (17), and carbohydrate binding module (94) (Drula et al., 2021; Henrissat et al., 1989). These enzymes may be secreted as free enzymes, or they may be deployed as part of cellulosomes, modular structures attached to the organism cell wall on which CAZymes are mounted to enable co-location of hydrolysis and carbohydrate transport into the organism (Gilmore et al., 2015; Terry et al., 2019).

Fermentation takes place after microbiome members absorb molecules liberated by hydrolysis. **Acidogenesis**, the stage immediately following hydrolysis, occurs when organisms ferment carbohydrates into carboxylic acids and ethanol. Organisms taking part in acidogenesis may be divided into primary fermenters, who ferment sugar, and secondary fermenters, who ferment the carboxylic acid products of primary fermentation (Heyer et al.,

2015). Most primary fermenters also take part in the previous hydrolysis stage, as they employ their CAZyme repertoire to cleave mono- and oligosaccharides for their own and others' consumption (Peng et al., 2021). Bacteria, fungi, and protozoa may carry out primary fermentative processes in anaerobic consortia (Peng et al., 2021; Williams et al., 2020). Secondary fermentation is the exclusive province of bacteria, most of which are obligate syntrophs and ferment longer short-chain fatty acids (SCFAs) into smaller SCFAs for metabolism by methanogens (Heyer et al., 2015). The range of products generated by this process includes 1- to 6-carbon SCFAs, H_2, and CO_2. Additional metabolism by reverse beta-oxidizers may produce medium- and long-chain fatty acids of 6 or more carbons (Liao et al., 2016). **Acetogenesis** is performed by acetogenic bacteria and generates acetate via the reduction of organic acids or CO_2 (Angelidaki et al., 2011; Jin et al., 2021). In anaerobic consortia, this may occur via reduction of CO_2 by the Wood-Ljungdahl pathway, autotrophy from CO_2 and H_2, or fermentation of organic acids (Ljungdahl, 1986).

The final stage, **methanogenesis**, occurs solely in **methanogens**, members of the Archaea (Evans et al., 2019). Methanogens employ one or more pathways from three classes of metabolic routes to generate methane: **hydrogenotrophic** methanogenesis, from H_2 or formate; **aceticlastic** methanogenesis, from acetate; or **methylotrophic** methanogenesis, from methanol and methylated compounds. All three pathways are united by a final step in which methyl-coenzyme M is reduced to methane by methyl-coenzyme M reductase (mcr).

1.2.2. Microbial interactions in anaerobic digestion

Anaerobic digestion requires a community of microbes equipped with complementary metabolic pathways that together carry out this multi-stage process (Leng et al., 2017), and as such, the ability of a microbiome to stably carry out anaerobic digestion depends on the

regulation of ecological interactions between them. These interactions are governed by a delicate balance between cooperation and competition, as well as species diversity. More specifically, they are governed by limiting positive feedback loops arising from cooperation and consequent mutual dependency (which may lead to a cascading series of decreases in fitness in other strains if just one strain's fitness is depleted), as well as a weakening of exploitative ecological interactions (that are both non-competitive and non-cooperative), which secondarily affect the interactions between productivity and stability – what Coyte *et al*. termed "ecological network theory" (2015).

Mutualistic interactions link all four stages of anaerobic digestion and are required for the complete degradation of substrate into gaseous and soluble end products. The most straightforward instance of mutualism in anaerobic digestion is the cross-feeding which takes place between consortium members. This promotes stability because the strains can grow best at intermediate densities, while they do not grow as well at lower densities (Vet et al., 2020). Participants in a mutualistic interaction may also alter the gene expression and metabolic output of one another. For example, anaerobic fungi and methanogens live together in the rumen; in coculture experiments, some anaerobic fungi exhibit increases in CAZyme expression and changes to the end products of lignocellulose degradation (Gilmore et al., 2019; Li et al., 2020; Swift et al., 2019). A special case of obligate mutualism unique to microbiota, syntrophy, occurs when otherwise thermodynamically unfavorable metabolic reactions are made favorable only in a partnership between two or more microbes (Morris et al., 2013). For example, syntrophic secondary fermenters that oxidize 3+-carbon fatty acids into acetate and produce hydrogen in an endergonic reaction rely on hydrogen-consuming microbes, hydrogenotrophic methanogens, to remove H_2 and thus make the coupled

metabolic reactions by these microbes thermodynamically favorable (Kouzuma et al., 2015). Since the methanogens use H_2 from the syntrophs to make methane, this results in a mutually beneficial positive feedback mechanism that enables and enhances acetogenesis and methanogenesis.

Non-mutualistic interactions regulate the diversity and stability of the consortium. Commensal strains in the human gut, such as the yeast *Candida albicans*, promote resistance to pathogenic microbes by occupying space in the gut and triggering protective immune responses; however, a lapse in regulation of the microbiome by its host can lead to opportunistic infection by these microbes, and the presence of such an infection may be a sign of compromised immune function (Kumamoto et al., 2020; Miceli et al., 2011). Antagonistic interactions, such as protozoal predation on bacteria and fungi in lignocellulolytic consortia (Solomon et al., 2021; Williams et al., 2020), function as potent population regulators and drivers of coevolution (Mall et al., 2022; Nair et al., 2019). Antagonistic interactions within lignocellulolytic microbiota are attenuated to some extent by the digestion of complex carbohydrates, and this inhibition is not uniform among bacterial groups (Deng & Wang, 2017).

Competition plays a central role alongside cooperation in regulating the stability of a microbiome (Coyte et al., 2015). On simple substrates, the law of competitive exclusion (Gause's law) stipulates that out of two species occupying the same niche, one will eventually either go extinct or adopt a new non-competitive niche as a result of outcompetition by the other species (Elton, 1927). Competition, like other antagonistic interactions, is reduced by the provision of a complex lignocellulosic substrate due to the inability of any one microbe to occupy all possible polysaccharide-utilizing niches on any

substrate with a large diversity of disaccharide bonds (Deng & Wang, 2017; Laine, 1994; Pérez & Tvaroška, 2014; Rakoff-Nahoum et al., 2016; Solden et al., 2018). Competition can also promote cooperation: for example, it fosters cooperative behaviors within groups of microbes employing antagonistic strategies that align in some fashion with each other (D'Souza et al., 2018). Competition may additionally foster cooperation within a strain by limiting population density and depleting public goods produced by cooperating individuals of the same strain, if those individuals have preferential access to those public goods (Celiker & Gore, 2012). The stabilizing role of competition in an environment where non-competition is fostered by a multitude of metabolic niches arises from its ability to dampen positive metabolic feedback loops and mutual dependencies generated by cooperation. Coyte et al. (2015) showed that an increase in cooperation nearly always decreases the probability of stabilization after a perturbation, and additionally slows the rate of return to stability, and showed that it does so through the mutual dependence created by cooperation: decreases in density by one species are followed by decreases in density of cooperative species. Coyte et al. also found that the destabilizing effect of high biodiversity in a consortium is mitigated if additional species introduce competitive interactions, and not cooperative interactions, and this has the effect of forcing tradeoffs between metabolic efficiency and stability.

1.2.3. Degradation of lignocellulose in the herbivore gut

Human-engineered solutions to dispose of lignocellulosic biomass fall far short of nature's most efficient plant-degrading bioreactor: the herbivore digestive tract. The highest loading rates of industrial reactors treating organic solids via high-solids technologies are around 10 kg of organic dry matter per m^3 of reactor volume in a day, whereas the rumen of

cattle, in comparison, processes about three times as much per unit of volume per day as the most efficient reactors, and some insects can process up to 400 kg of organic dry matter per m^3 of gut volume in a day, though they cannot be scaled up to the capacity of a modern-day industrial bioreactor due to physiological mechanisms that maximize the efficiency of the transport of reactants and the continuous absorption of chemicals by the intestine (Godon et al., 2013).

The vertebrate herbivore gut is a hypoxic environment separated into chambers, each with functions oriented toward the degradation and absorption of plant biomass: mechanical and chemical pretreatment (the mouth), chemical degradation (an acidic chamber in the stomach), a fermentation chamber (one or more microbial chambers in the fore- or hindgut), and an absorption chamber (the intestine) (Karasov & Douglas, 2013). Plant biomass in all vertebrates proceeds through mechanical and enzymatic pretreatment in the mouth, through acid and protease degradation in all or part of the stomach, to absorption and fermentation in the small intestine, large intestine, and (if present) the cecum. Vertebrate herbivores adhere to one of two broad types of chamber configuration, defined by the presence (in foregut fermenters) or absence (in hindgut fermenters) of an additional large fermentation chamber before the acid chamber, with further division of hindgut into cecal fermenters and colon fermenters.

Vertebrate herbivory depends on the activity of an associated microbiome, primarily housed in the dedicated fermentation chambers of the host's gut (Hume & Warner, 1979). Animal genomes contain few CAZymes of their own. Cellulases have only been reported from the genomes of arthropods, nematodes, and mollusks, all of which are invertebrate taxa (Watanabe & Tokuda, 2001). Vertebrate genomes contain amylases, deployed in different

parts of the digestive tract (Crerar & Rooks, 1987; da Lage et al., 2011). The host and its gut microbiome regulate each other (Read & Holmes, 2017). The specific composition of a given mammalian herbivore gut microbiome varies depending on several factors, which can be broadly categorized into environmental influences and input, host phylogeny, and gut physiology (Amato et al., 2019; Ley et al., 2008; Meili et al., 2022; Nishida & Ochman, 2018b; Reese & Dunn, 2018), but generally contains a combination of the following: bacteria which perform cellulolytic, fermentative, and auxiliary functions, methanogenic archaea, cellulolytic and fermentative fungi, cellulolytic and fermentative protozoa, and viruses (largely phages) (Hobson & Stewart, 1997; Lobo & Faciola, 2021).

1.2.3.1. Degradation of lignocellulose by anaerobic fungi (Neocallimastigomycota)

Fungi of the early-diverging fungal phylum Neocallimastigomycota live in the digestive tracts of a phylogenetically diverse sample of herbivores (Liggenstoffer et al., 2010). These fungi field the largest, most diverse arsenal of CAZymes of any organism, and they deploy them as part of modular protein complexes called cellulosomes on the cell wall surface (Haitjema et al., 2017). Evidence from parallel enrichments where cellulolytic bacteria and fungi have been separately enriched suggests that the CAZymes deployed by anaerobic fungi complement those deployed by cellulolytic bacteria (Peng et al., 2021).

Besides liberating excess fermentable sugars for the rest of the consortium, including sugars they cannot consume such as arabinose and galactose (Henske et al., 2018), anaerobic fungi consume and ferment a diverse set of sugars themselves. Seppälä *et al*. (2016) mapped the membrane proteome of *Neocallimastix californiae*, *Anaeromyces robustus*, and *Piromyces finnis* and found that they expressed a large repertoire of putative sugar transporters representing all families of sugar transporters and putatively transporting a

diverse set of mono- and oligosaccharides. The primary fermentation products produced by anaerobic fungi are H_2, CO_2, ethanol, formate, lactate, and acetate (Bauchop & Mountfort, 1981).

Anaerobic fungi engage in metabolic interactions with other organisms in their environment. Anaerobic fungi and cellulolytic bacteria have a complex and dynamic relationship with each other in the herbivore gut. Joblin and Naylor found that bacterial fermentation products inhibited cellulose degradation by *Neocallimastix frontalis* (1993).Wilken *et al*. combined existing data measuring excess fermentable sugars liberated by fungi on lignocellulose with flux balance analysis to computationally screen a panel of six microbes of biotechnological interest as potential metabolic partners for anaerobic fungi, and found that anaerobic fungi were most suited to pair with *Clostridia ljungdahlii* and *Methanosarcina barkeri*, both also found in rumen microbiota (Wilken et al., 2018). Swift *et al*. established that anaerobic fungi exhibit antagonistic behavior in coculture with *Fibrobacter succinogenes* UWB7, a cellulolytic rumen bacterium, and the presence of *F. succinogenes* UWB7 induced a stress response in anaerobic fungi and activated secondary metabolic pathways (Swift et al., 2021). In addition, a study of relationships between anaerobic fungi and bacteria in the gut of cattle with and without bloat (a build-up of gaseous fermentation products) suggested that core fungal species' abundances were negatively correlated with those of several bacterial species, suggesting competition among these organisms, and that this pattern of negative correlation was disrupted by bloat. (Azad et al., 2020). However, anaerobic fungi and bacteria exhibit complementary cellulolytic strategies in the gut (Peng et al., 2021). In addition, coculture of anaerobic fungi with the cellulolytic bacterium *Clostridium acetobutylicum* appears to result in stable (over the 29-

day growth period) cocultures that promote the production of butyrate by *C. acetobutylicum* (Brown et al., 2022). Rumen protozoal genomes contain chitinases, which indicate that they may graze on anaerobic fungi (Solomon et al., 2022). Out of all the other organisms with whom anaerobic fungi may have a relationship in the herbivore gut, the most data has been generated from studying methanogens' relationships with these fungi. Anaerobic fungi and methanogens associate with other in biofilms (Mason & Stuckey, 2016). It is still unclear whether coculturing fungi with methanogens results in actual changes in biomass degradation relative to fungal monocultures, and non-destructive methods of quantifying these organisms in coculture have only recently been developed by Leggieri *et al*. (2021). However, changes in CAZyme, carbohydrate-binding module, dockerin, and pyruvate formate-lyase expression, as well as absolute and relative medium metabolite concentration have been observed between fungal monoculture and coculture with methanogens (Bauchop & Mountfort, 1981; Brown et al., 2021; Joblin & Williams, 1991; Swift et al., 2019). Finally, the presence of fungi outside the Neocallimastigomycota, or fungal extracts, appears to increase the metabolic activities of anaerobic fungi. Coculture with the yeast *Saccharomyces cerevisiae* stimulates zoospore production, increases cellulose degradation, and increases the production of H_2, formate, lactate, and acetate (Chaucheyras et al., 1995). Yeast extract is a component of complex media used to culture anaerobic fungi (Peng et al., 2018; Theodorou et al., 2005). Dietary amendment with cultures, extracts, or fermentation products of *Aspergillus oryzae*, a fungus not native to herbivore gut microbiota but present in soils and on plants (Elkhateeb, 2005), appears to promote the growth and metabolism of anaerobic fungi in the rumen and in culture, including upregulating CAZyme expression and

secretion and increasing and altering gas production (Harper et al., 1996; H. Sun et al., 2014; Schmidt et al., 2004).

1.3. Division of metabolic labor in lignocellulose degradation

Lignocellulose is degraded by consortia that divide the labor required to metabolize the substrate among many microbes. Division of labor is an adaptive feature of consortia that degrade complex molecules, and arises due to evolutionary and metabolic tradeoffs that are resolved by specialization in different parts of a metabolic process (Lindemann, 2020). These tradeoffs arise from physiological, and physiochemical, limitations, including but not limited to the following: distribution of metabolic fluxes between intracellular metabolic networks, energetic limits to the number of active metabolic pathways in a single organism, incompatibilities between metabolic reactions within an intracellular environment, and mutation accumulation and antagonistic pleiotropy gained as a result of adaptation to one environment in favor of another (de Groot et al., 2019; Noor et al., 2016; reviewed in Giri et al., 2019). Division of labor in lignocellulolytic consortia flows from the inability of any one microbe to completely degrade a complex lignocellulosic substrate due to the increasing metabolic burden on an organism with each expressed enzyme, and the dependence of some microbes on others to produce metabolizable substrates for them (Wu et al., 2016). Communities that engage in division of labor are assembled and stabilized in such a way that benefits are evenly allocated among community members, with a tendency for strains involved in the last steps of metabolism to dominate the community (M. Wang et al., 2022). Tsoi *et al*. analyzed 24 common metabolic pathway structures potentially subject to division of labor and derived general criteria under which this phenomenon is favored over metabolism by a single strain in consortia: given a specific metabolic pathway, division of

labor is favored under conditions of high enzymatic burden, intermediate and product toxicity, pathway complexity, and number of extracellular steps – and in general, if cell density overcomes inefficiency in resource exchange (Tsoi et al., 2018).

Four criteria have been defined by Giri *et al.* that must be met for a consortium to have divided its metabolic labor, and lignocellulolytic consortia likely meet all four of these criteria (Giri et al., 2019; Lindemann, 2020). First, microbes must provide functional complementarity to each other and trade some sort of commodity, such as a metabolite or enzymatic function. Lignocellulolytic consortia display complementarity in their hydrolytic and metabolic machinery (Peng et al., 2021). Second, this division of labor must confer a synergistic advantage: the interaction between strains must increase fitness over strain fitness in isolation from each other. Many relationships in lignocellulolytic consortia increase each partner's fitness and, in the case of syntrophies, allow metabolic processes to take place that would not take place if partners were separated. Third, fitness and relative abundance must exhibit an inverse relationship (negative frequency-dependent selection), which leads to stable coexistence over long periods of time. Lindemann postulated that lignocellulolytic consortia meet this criterion through the dependence of the cellulolytic activities of strains on the activity of other organisms (2020). Fourth, mechanisms of positive assortment must lead an individual to be more likely to encounter their interaction partners relative to meeting other microbes not participating in said interaction. Complex substrates themselves provide both spatial and chemical scaffolding for the organization of a consortium, and so too do the tissue lining the gut lumen and the biofilms formed by some microbes. Lindemann postulated that competition for resources and intermicrobial communication sets up conditions favorable to the colocation of interaction partners.

The complexity of lignocellulosic substrates reduces competition in a lignocellulolytic consortium through requiring a diverse repertoire of enzymes for degradation and metabolism, as well as providing a spatial scaffold for microbial colonization (Lindemann, 2020). The distribution of CAZyme deployment by different strains of cellulolytic microbes, explored in Chapter 2, is not well understood. Carbon catabolite repression forces the preferential consumption of one mono- or disaccharide over another, which limits the capacity of single isolates to degrade substrate, but allows consortia who differentially prioritize sugar consumption to cooperate in more efficiently consuming substrate (L. Wang et al., 2019). The consequent resource partitioning stabilizes the consortium, and enhances the growth rate, productivity, and yield of each strain (Chappell & Nair, 2017; Shahab et al., 2020). Metabolic dissimilarities establish cross-feeding interactions, and in lignocellulolytic consortia, this enables organisms with other functions (*e.g.* secondary fermenters and fatty acid chain elongators) to carry out metabolism to provision the host with other nutrients even though they cannot degrade the complex substrate (Giri et al., 2021; Lindemann, 2020).

1.4. Substrate effects on lignocellulolytic microbial communities

1.4.1. In vivo *studies*

The consumption of lignocellulose by an animal host shapes the composition and diversity of the gut microbiome. Across a diverse sampling of animals both vertebrate and invertebrate, an increase in the percentage of calories derived from plants is associated with an increase in both Shannon diversity and richness in the gut microbiome (Reese & Dunn, 2018). Interaction with humans generates shifts in the herbivore gut microbiome. For example, human encroachment on animal habitat may have differential effects on primate

species feeding in the same forest (Barelli et al., 2020). Captivity has a variable effect on alpha diversity depending on animal taxon (McKenzie et al., 2017). Frequent exposure to humans promotes the humanization of captive primates' and lizards' microbiota (Clayton et al., 2016; Tang et al., 2020).

In vivo studies of substrate effects on the gut microbiota of non-gnotobiotic herbivores permit the study of the effect of substrate on a defined microbiome along with effects from and on the host. Studies of feed composition in ruminant livestock typically involve observing microbiome composition and activity in response to one of two kinds of dietary manipulations: either change of forage and grain composition, or amendment with concentrates and other additives (Gruninger et al., 2019). The overall carbohydrate content of feed given to livestock, as well as the forage source and type of preservation, shape the microbiome, both in its overall composition and structure and in how organisms associate with each other. For example, Kumar *et al*. found that a switch from a high-forage to a high-grain diet produced a loss of interactions between anaerobic fungi and other microbes (Kumar et al., 2015). Goats switched from a forage to a concentrate diet exhibited shifts in the genus-level composition of their rumen anaerobic fungal community (Fliegerova et al., 2021). Dietary additive studies provide some insight into how specific substrate components may drive microbial activity: for example, manipulating protein produces shifts in ammonia and volatile fatty acids, manipulating lipids alters fatty acid hydrogenation, and manipulating carbohydrates produces shifts in methane and C1-C4 fatty acids (Belanche et al., 2012; Carreño et al., 2019; Newbold & Ramos-Morales, 2020).

Gnotobiotic animal models allow manipulation of a gut microbiome with known composition directly within the environment of the gut. Many investigations assessing the

impacts of diet on microbiome composition and activity in gnotobiotic models have been carried out in rodents to model human gut microbiota (Martín et al., 2016). For example, Turnbaugh *et al*. found that a switch from a low-fat, polysaccharide-rich diet to a high-fat, high-sugar diet changed the structure of the humanized gut microbiota of gnotobiotic mice (Turnbaugh et al., 2009). Gnotobiotic sheep models have been used to evaluate the effects of defined microbiota on the growth and rumen function of lambs fed on high-starch diets, as well as dissect the molecular interactions between *Fibrobacter succinogenes, Ruminococcus flavefaciens,* and *Ruminococcus albus* (Yeoman et al., 2021).

Studies performed in samples directly taken from an herbivore's digestive tract cannot factor out the effects of the animal host on microbiome composition and activity. In mammals, multiple organ systems regulate, and are regulated by, the gut microbiome (Stappenbeck & Virgin, 2016). Gnotobiotic animals allow control over microbiome composition and are ideal for studies in which both the contributions of host factors and external manipulations (such as substrate change) of the microbiome are assessed (Martín et al., 2016). Due to the need to control their microbiota, gnotobiotic models require special procedures and facilities for rearing and housing that are difficult to access for many researchers. The generation of gnotobiotic ruminant models poses an additional challenge due to the critical role of the gut microbiome in proper development of their digestive tract, which is anatomically underdeveloped and metabolically non-functional at birth (Govil et al., 2017). However, these studies most faithfully represent the activities of a microbiome in its natural setting. There are limits to the extent to which features of the herbivore gut that enable consortium stability can be replicated with a defined consortium; Martin, Bermudez-Humaran, and Langella (2016) advocate for the use of gnotobiotic animals as a compromise

between the complex natural environment of the animal gut and the controlled and simplified, but loss-prone, environment of enrichment culture, even if the use of gnotobiotic animals is accompanied by significant challenges.

1.4.2. Enrichment consortia permit the isolation and study of subsets of microbiota

Enrichment consortia remove the effects of the animal host and facilitate the dissection of microbially-driven mechanisms of microbiome interaction and activity. Enrichment of any consortium produces a simplified, more experimentally tractable subset of organisms – never the full source microbiome – with microbial interactions that are carried over to some extent from the source microbiota (Hug & Co, 2018). Enrichment has been used to successfully select more efficient lignocellulolytic consortia (Cheng et al., 2009; Feng et al., 2011; Lazuka et al., 2015). Lignocellulolytic consortia have been enriched from a wide taxonomic spread of herbivores – for example, from goats (Peng et al., 2021), beavers (*Castor americanus*), moose (*Alces alces*) (Wong et al., 2016), buffalo (Callaghan et al., 2015), and termites (Lazuka et al., 2018).

Defined synthetic consortia raised on simple substrates are already highly tractable model systems for the study of microbial ecology and the production of desired dynamics of interaction and metabolism, and enable a variety of applications (Deter & Lu, 2022). Much of what we know about the ecology of microbial communities has been learned from rearing consortia on simple substrates (Lindemann, 2020). Larger, more complex consortia growing on complex, chemically undefined substrates are difficult to investigate in the same degree of detail. The standard approach to these consortia typically entails one or more types of "meta-omics" investigation, and the enrichment usually follows either a sequential or parallel experimental design. Sequential enrichment may be used to investigate the effects of

substrate shift on a relatively stabilized enrichment consortium, or to shape an enrichment consortium in stages using substrate. For example, Carlos *et al*. observed shifts in community composition, including functional convergence with the same consortium initially reared on lignin, after shifting a microbial consortium reared on xylan to lignin (Carlos et al., 2018). Parallel enrichment of a single source microbiome allows for the simultaneous comparison of different substrates' effects on a consortium. The panel of substrates on which these consortia are enriched is generally reflected in which taxa are enriched, as well as in genomic markers such as the CAZyme and metabolic enzyme repertoires of the taxa enriched. Peng *et al*. (2021) enriched a source microbiome from goat fecal pellets on xylan and three structurally diverse plant biomass substrates (alfalfa, bagasse, and reed canary grass) and showed that more complex substrates enriched a highly functionally redundant lignocellulolytic community whereas xylan, a pure hemicellulose substrate, enriched a small community dominated by a few specialists. Chapter 2 of this dissertation describes the follow-up study by Dickson *et al*. (in preparation), in which this same parallel enrichment was found to exhibit substrate-based differences in which CAZymes were expressed and to what degree, but in which fundamental patterns of division of CAZyme labor held relatively constant between substrates.

Other substrate-related effects, rather than the original panel of substrates selected, impact consortia. Previous use of a substrate may have a significant effect over and above substrate type; Jimenez *et al*. inoculated consortia of fresh and once-used wheat straw, switchgrass, and corn stover with a soil consortium and observed that the relative abundances of taxa in used substrates were more similar to each other than to their fresh equivalents (Jiménez et al., 2016). Enriching source microbiota from separate sources on the

same substrate may result in convergence between the microbiota in one or more characteristics. Wong *et al.* (2016) enriched source microbiota from the feces of beavers (*Castor canadensis*) and moose (*Alces alces*) in parallel on poplar hydrolysate and three different cellulose preparations for three years, and found that communities enriched on the same carbon source converged in composition, while communities from the same source enriched on different carbon sources did not share core species.

Repetition of the same enrichment experiment multiple times, by itself, may yield different community composition at the species level, even if family and higher levels of taxonomy are highly similar across replicates (Estrela et al., 2021). Substrate sterility at inoculation affects enrichment results, with non-sterile substrate producing decreases in species richness and Shannon and Simpson diversity indices, as well as decreasing degradative activity, hydrolytic enzyme activity, and volatile fatty acid production (Lazuka et al., 2018).

1.5. Variation in gut microbiome structure and composition among herbivore species

In addition to the carbon substrate consumed by the herbivore host, two other factors, host phylogeny (*i.e.* the species of the host and the evolutionary relationships between hosts) and gut physiology, serve as major drivers of variation in gut microbiome structure and composition among herbivore species. Here, we will focus primarily on the qualitative effects of each of these factors in shaping the gut microbiome. It is important to note that substrate (and other environmental inputs), gut physiology, and host phylogeny are all interrelated to some extent, and are themselves broad categories of more specific factors that may influence the gut microbiome.

1.5.1. The effects of gut physiology

At the level of the subphylum Vertebrata, gut physiology is the strongest driver of microbiome composition, and this is also the case when considering the Mammalia by themselves (Reese & Dunn, 2018). Mammalian herbivores ferment ingested plant biomass in either a sacculated foregut, as in the case of ruminants, or in their hindgut, as in the case of animals like horses and humans (**Figure 1**)(Godon et al., 2013). The location of the primary fermentation chamber, relative to other areas of the gut, impacts microbiome composition (Ley et al., 2008; Liggenstoffer et al., 2010). Ingested material undergoes both mechanical and chemical pre-processing, dependent on gut configuration, prior to reaching the microbiome: it is chewed and attacked by salivary enzymes in the mouth, and in hindgut fermenters is pre-processed by treatment with hydrochloric acid and gastric enzymes in the stomach, and by pancreatic enzymes and additional enzymes, which primarily attack lipids, proteins, starch, and oligosaccharides such as lactose and sucrose, in the small intestine prior to fermentation in the colon (Bellmann et al., 2015; Bradley et al., 2013; Godon et al., 2013; Hooton et al., 2015; Isenman et al., 1999; Janiak, 2016; Whitcomb and Lowe, 2007). Fermentation in a chambered foregut is a digestive adaptation almost exclusively found among animals that consume high proportions of complex polysaccharide biomass (with the exception of the toothed whales, described in Section 1.6.2). Secondary fermentation takes place in all foregut fermenters, but only in some hindgut fermenters (Hume, 2013). Gut volume shapes microbial community diversity; Godon *et al*. analyzed the relationship between body mass and bacterial diversity and found that the latter increased with the former following a power law, regardless of phylogeny, diet, or gut anatomy (Godon et al., 2016). Increased transit time decreases diversity, as communities of microbes are less likely

to be able to establish themselves in a digestive tract through which substrate travels more quickly (McKenney et al., 2018; Roberfroid et al., 2010).

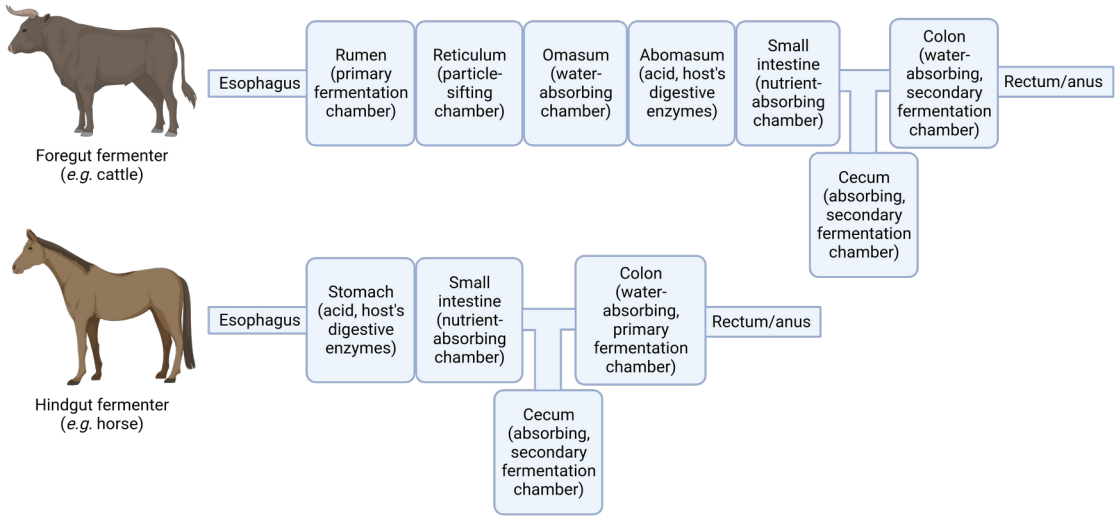


Figure 1. Diagram of the digestive tracts of foregut-fermenting and hindgut-fermenting herbivores. Made with Biorender.

One potential factor in shaping gut microbiota across herbivores that intersects with host phylogeny is the host's genome and those of its products that participate in digestion. The emerging field of hologenomics addresses related questions, considering the host and its microbiome as a system called a holobiont (Bordenstein & Theis, 2015; Margulis, 1993). Herbivores have undergone gene changes associated with the adaptation to a lignocellulose diet. For example, a genetic screen across 31 placental mammals with varying diets found that herbivores repeatedly lost the triglyceride lipase inhibitor PNLIPRP1, and within this group, foregut fermenters retained syncollin (SYCN), which participates in exocytosis in the pancreas (Hecker et al., 2019). In addition, mammals and some of their gut microbes have evolved in parallel, and a review by Groussin, Mazel, and Aim suggested that individual gut bacteria diverge in patterns that recapitulate host phylogeny (Groussin et al., 2020). Also, the vertebrate immune system interacts in a reciprocal fashion with the gut microbiome and

the epithelium of the digestive tract, and studies in humans show that the gut microbiome does not simply evade the immune system but instead works with the host to achieve and maintain homeostasis (Lee et al., 2022); in herbivores, these interactions are not well studied, but regulation is likely to follow similar patterns in humans and other animals. For example, in addition to shaping the gut microbiota, diet composition also has been found to shape secretion of salivary immunoglobulin IgA and regulation of rumen epithelial Toll-like receptors (Fouhse et al., 2017; Liu et al., 2015).

1.5.2. The effects of host phylogeny

Herbivory has independently evolved multiple times within the Mammalia, and most extant mammals are herbivores (Ley et al., 2008). Hindgut fermentation is a basal trait among herbivores in the Mammalia, and foregut fermentation has evolved independently twice within this class, in the artiodactyls and primates, and once outside the mammals, in the hoatzin, a bird (Grajal et al., 1989; Stewart et al., 1987). As mentioned above, gut physiology is the strongest overall driver of microbiome composition and diversity when viewing the Mammalia as a whole, but within individual mammalian taxa, the importance of host phylogeny varies. A comparison of microbiome compositions across 112 species representing 14 mammal orders by Nishida and Ochman (2018) found that, over the 75 million years following the diversification of the Mammalia, most of the studied lineages diverged from each other at the same rate. However, Cetartiodactyla, the order containing ruminants and whales, evolved much more quickly than the other orders studied. The investigators also found substantial variation across orders in the degree of association between divergence rate and changes in diet and environment. Notably, divergence in the Cetartiodactyla was associated with changes in diet and environment, whereas in the

Primates, differences accumulated regardless of diet. Other studies have shown that among the primates, host phylogeny is the strongest driver of the composition of the prokaryotic fraction of the microbiome (McCord et al., 2014; Ochman et al., 2010; Yildirim et al., 2010). Each of these factors may have differential effects on different fractions of the microbiome; for example, diet and environment play a larger role in driving the composition of the fungal portion of the microbiome (Zoelzer, Burger, and Dierkes, 2021). Little data exists accounting for the eukaryotic portion of the microbiome in studies of the relative impact of host phylogeny on gut microbiota. Host phylogeny was the most significant driver of anaerobic fungal community composition and diversity in a study of known hosts encompassing the green iguana (*Iguana iguana*) and 32 diverse foregut and hindgut fermenters, encompassing members of the Marsupialia, Cetartiodactyla, and Perissodactyla (Liggenstoffer et al., 2010).

A comparison of two extreme examples of microbiome variation within mammalian taxa illustrates the variation across mammals in how differently host phylogeny shapes the gut microbiome. Cetaceans are the sister clade to the artiodactyls, among whom are included the ruminants (Graur & Higgins, 1994). They are descended from terrestrial herbivores and have no significant cellulosic biomass available to them in the ocean, either consuming large nektonic prey (toothed whales) or massive volumes of small fish and planktonic crustaceans (baleen whales) to satisfy the energy requirements of a large, highly insulated body and an active predatory lifestyle (Beier & Bertilsson, 2013; Blanco et al., 2001; Gatesy et al., 2013; Langer, 2001). Investigation of toothed whale microbiota suggests they have a gut microbiome more similar to that of terrestrial carnivores (Soverini et al., 2016). Baleen-enabled filter feeding is a derived trait in cetaceans, and baleen whale microbiota exhibit

more similarities with those of terrestrial herbivores, with respect to functional repertoire and high-level taxonomy, reflecting their increased reliance on the polysaccharide chitin (Beier & Bertilsson, 2013; Sanders et al., 2015). On the other side of the scale, the giant panda (*Ailuropoda melanoleuca*) is related to other bears, which consume meat-rich diets, and has a digestive tract identical in shape and relative volume, but consumes mostly highly-fibrous bamboo; its fecal microbiome still more closely resembles that of carnivores than other herbivores (Guo et al., 2019; Xia et al., 2022).

1.5.2.1. The primate gut

With the exception of monkeys of the subfamily Colobinae, primates are hindgut fermenters (Wolfensohn, 2004). Primates occupy a diverse set of dietary niches. Even among herbivorous non-human primates (NHPs), diets vary greatly in their species and plant part composition, with intraspecific variation by season, age, and a variety of other factors, and the range of gut morphologies herbivorous NHPs exhibit demonstrates this diversity. Compared to the extensive body of research on the human gut microbiome, much less attention has been paid to NHP gut microbiota (Clayton et al., 2018). Among NHPs, most culture-independent studies have been published on the microbiota of the Haplorrhini ("Old World" monkeys and great apes) (Clayton, 2015). Of these investigations, most have focused on macaques (subfamily Cercopithecinae), and most macaque studies have taken samples from captive monkeys. Few species of the Catarrhini ("New World" monkeys) and Strepsirrhini (lemurs and lorises) are represented in extant studies.

Prokaryotic populations of the wild NHP gut vary chiefly by host phylogeny. Among strepsirrhines, the microbiome of wild lemur populations is generally dominated by the Firmicutes, then Bacteroidetes, with smaller populations of Proteobacteria; among wild

lorises, which are specialist gumivores, Proteobacteria are more abundant than in the lemurs, and the relative abundance of Firmicutes, Bacteroidetes, and Actinobacteria as the most abundant other phyla is more variable (Clayton, 2015; Clayton et al., 2018). Catarrhines, represented by howler monkeys (*Alouatta* spp.) and spider monkeys (*Ateles* spp.), have microbiota dominated by the Firmicutes and Bacteroidetes (in wild howler monkeys, the Firmicutes are more dominant), with smaller populations of Proteobacteria. "Old World" non-colobine monkeys (colobine microbiota are explored below) exhibit a broad range of microbiome structures reflective of their varied diets, with varying prominence of the phyla Firmicutes, Bacteroidetes, Proteobacteria, and Actinobacteria. Among the apes, gorillas, the largest primates, exhibit large colons and host prokaryotic populations with high fiber degradation and fermentation capacities, whereas humans' closest relatives, the chimpanzees (*Pan* spp.) have microbiota exhibiting similarities to those of humans consuming non-Westernized diets.

Less is known about the fungal population of the NHP gut. Most herbivorous NHP microbiome studies measuring the relative abundance of fungi by ITS sequences reveal a fungal population most often having Ascomycota as the majority phylum, with smaller populations of Basidiomycota and Zygomycota, and sizable populations of yeasts across these phyla, but large proportions of fungi often remain unidentified in studies of herbivorous primate mycobiota (Barelli et al., 2020; Borruso et al., 2021; James et al., 2022; Mann et al., 2019; Schulz et al., 2018; B. Sun et al., 2021).

Fungi of the Neocallimastigomycota may be critical members of the gut microbiota of some herbivorous NHPs. In 2018, Schulz *et al*. isolated ITS1 sequences of Neocallimastigomycota from gorilla feces, the first identification of this clade in primates

(Schulz et al., 2018). The amplified sequences clustered most closely with lineage AL3, previously isolated from horses, who exhibit relatively similar digestive physiology, particularly relative colon size. In 2022, Houtkamp *et al*. released a preprint suggesting that transcriptional activity of the Neocallimastigomycota comprised nearly all eukaryotic transcripts recovered from their samples (Houtkamp et al., 2022). Neocallimastigomycota have previously been found in human samples, but there is no indication as to whether these fungi were not simply transient through the gut (Mar Rodríguez et al., 2015).

Monkeys of the subfamily Colobinae consume highly folivorous diets, which are fermented in a foregut with three or four chambers, uniquely so among the primates (Bauchop & Martucci, 1968; Chivers, 1995). One species, the proboscis monkey (*Nasalis larvatus*), exhibits behavior similar to rumination (but not sufficiently identical to be classed as such), which is known to impact the structure of the microbiome (Matsuda et al., 2011, 2015). A recent study by Amato *et al*. indicated that the prokaryotic portion of their microbiome exhibits convergence with that of ruminant herbivores (2020). Three studies exist that include investigations of the gut mycobiota from one or more colobine species. Barelli *et al*. surveyed the mycobiome of red colobus (*Procolobus gordonorum*) as part of a study assessing the impact of habitat disturbance on wild arboreal and ground-feeding primates, and found that the mycobiome of the red colobus population surveyed was dominated at the phylum level by Ascomycota (>50% of relative abundance), with very small populations of Basidiomycota and Zygomycota, and about 30% of relative abundance was composed of fungi unidentified at the phylum level (Barelli et al., 2020). Xu *et al*. analyzed the fecal microbiome of the Yunnan snub-nosed monkey (*Rhinopithecus bieti*) to profile the lignocellulolytic capacity of its microbiome, and found that Ascomycota was the

dominant phylum, primarily *Aspergillus* spp. (particularly *A. fumigatus*), *Gibberella* spp., *Magnaporthe* spp., and *Neurospora* spp. (Xu et al., 2015). Two members of the Basidiomycota were also detected, *Schizophyllum commune* and *Cryptococcus neoformans*. Anaerobic fungi were not detected, but Xu *et al*. nonetheless speculated that they were likely to be present. Mann *et al*.'s study of the biodiversity of protists and nematodes in the NHP microbiome took samples from wild *Colobus guereza* and *Piliocolobus badius*, and found a mycobiome dominated mostly by Ascomycota, with a much smaller proportion of other fungi (Mann et al., 2019). The same study asserted that fungal OTUs detected in that study were predominantly specific to individuals with very few OTUs shared across hosts and that most were derived from food or the environment. A majority of Ascomycota in the recovered mycobiome is consistent across all three studies; in Chapter 3, we investigate the captive *C. guereza* mycobiome in detail.

1.6. Techniques

1.6.1. Enrichment techniques for gut-derived consortia

Selection of appropriate substrate and medium depends on the study objective, but may be done to mimic the source microbiome's natural conditions or to enrich specific taxa under specific pressures. Mineral medium (MM) has been used to enrich bacterial samples from beaver and moose feces, and MM supplemented with vitamins and trace elements has successfully enriched bacterial consortia from termite feces (Lazuka et al., 2018). M63 medium, a MM supplemented with glycerol, has been used to enrich bacterial consortia from chicken feces (Carlos et al., 2018). The metabolome of rumen fluid contains a myriad of amino acids, other organic acids, trace elements, phospholipids, and other trace compounds difficult to replicate in defined media, and rumen fluid is a component of

complex media used to enrich gut-derived consortia containing microbes with unknown nutritional requirements (Saleem et al., 2013). Rumen-fluid-containing complex medium is frequently used in the enrichment of gut consortia containing anaerobic fungi, though is not strictly necessary for the survival of fungi, as demonstrated by Wilken *et al*.'s use of the defined medium M2 (containing cellobiose, L-cysteine, trace elements, and hemin) in the construction of a genome-scale metabolic model for *Neocallimastix lanati* (Wilken et al., 2021). Enrichment on different biomass substrates may enrich different fungi from a sample (for example, differential enrichment of anaerobic fungal genera in Peng et al., 2021, explored further in Section 1.7.2). Substrate matching to the animal host's diet may be an appropriate approach for optimal enrichment of fungi from a gut sample, and we employ this technique in Chapter 3 by using alfalfa, a pectin-rich dicotyledonous plant, to enrich anaerobic fungi from the foregut of the Eastern black-and-white colobus, *Colobus guereza*, a primate that subsists mainly on pectin-rich dicots.

Selective enrichment of desired components of the herbivore gut microbiome may be achieved through the application of either specific carbon substrates or antibiotics. Chloramphenicol is frequently used to enrich fungi alone, and a combination of penicillin and streptomycin has been used to enrich fungus and methanogen pairs from feces in studies of their metabolic relationship (Bauchop & Mountfort, 1981; Gilmore et al., 2019; W. Jin et al., 2011; Li et al., 2020; Peng et al., 2021). Parallel enrichment may be easily combined with selective enrichment; in combination with the selective enrichment described above, Peng *et al*. (2021) enriched the same goat fecal microbiome on four different substrates to simultaneously examine how substrate and antibiotic selection shaped its composition, stability, and metabolic activities.

Herbivore microbiota are spatially organized in the gut. No previous studies have attempted to impose spatial organization on enriched gut consortia, but some inferences may be made from studies involving other consortia. Shahab *et al*. set up an oxygen gradient allowing the coculture of the aerobic fungus *Trichoderma reesei* with facultative and obligate anaerobic bacteria (Shahab et al., 2020). The colonization of a bioreactor with *T. reesei*, allowing it to form a biofilm, prior to inoculation with a rumen microbiome, increased cellulolytic activity over that of a bioreactor without a biofilm-forming fungus (Xiros et al., 2019).

1.6.2. Isolation and phylogenetic barcoding of anaerobic fungi of the Neocallimastigomycota

The standard procedures for culturing anaerobic fungi from gut samples or feces largely follow the methods initially published by Hungate and Macy (1973), with subsequent modifications by Balch and Wolfe (1976), Bryant (Bryant, 1972), and Miller and Wolin (1974): inoculation of colonized material into liquid medium containing substrate under a headspace of 100% CO2, followed by a series of dilutions resulting in 1000-10000-fold dilution of the original inoculated medium. As detailed in section 1.7.1, broad-spectrum antiprokaryotics such as chloramphenicol may be added to enrich fungi alone, or more targeted antibacterials such as penicillin/streptomycin may be employed to enrich fungi along with methanogens. Cultures containing anaerobic fungi may be detected by a measured increase in the pressure of headspace gas or by the formation of a floating mat of substrate (Theodorou et al., 2005). Isolation of individual strains further follows the roll-tube

method published by Theodorou, Brookman, and Trinci, in which zoospores colonize agar medium over a series of 10-fold serial dilutions and individual colonies may be picked.

As with anaerobic fungi-dominated consortia, complex medium containing rumen fluid is the medium of choice for isolating anaerobic fungi (Saye et al., 2021). The choice of substrate has an impact on the fungal species enriched. Peng *et al*. enriched a goat fecal source microbiome dominated by *Piromyces* spp. and *Caecomyces* spp. with smaller populations of *Neocallimastix* spp. (Peng et al., 2021). Enrichment on alfalfa, bagasse, and reed canary grass produced a consortium dominated by *Neocallimastix* spp. and to a lesser extent *Piromyces* spp. and *Caecomyces* spp., whereas enrichment on xylan enriched a consortium wholly composed of *Caecomyces* spp; *Neocallimastix* spp. and *Piromyces* spp. are rhizoidal fungi capable of penetrating plant biomass mechanically, whereas *Caecomyces* spp. are non-rhizoidal fungi.

Historically, the internal transcribed spacer (ITS) regions of the *rrn* operon, encoding the ribosomal RNA (rRNA), have been used as the diagnostic regions for taxonomic assignment of fungi (Schoch et al., 2012). The ITS1 and ITS2 sequences are 130-200 bp regions separating the 18S small subunit (SSU), 5.8S, and 28S large subunit (LSU) rRNA (**Figure 2**). However, several characteristics of this region hinder its use in the Neocallimastigomycota. The ITS1 region is highly polymorphic in length and secondary structure, and may exhibit significant sequence divergence between copies in a single strain, exceeding cutoffs used for species- and sometimes genus-level distinction (Callaghan et al., 2015; Edwards et al., 2019; Koetschan et al., 2014). The region encoding the hypervariable domains 1 and 2 (hereafter referred to as the D1/D2 region) of the gene encoding the 28S large subunit (LSU) ribosomal RNA has been proposed as a more robust barcoding

alternative (reviewed in more depth in Section 1.7.3.1). Hanafy *et al*. sought to develop the use of the 28S D1/D2 as a taxonomic barcode for the Neocallimastigomycota by correlating the ITS1 regions of existing cultured genera and most existing candidate genera to the corresponding D1/D2 region, as well as generating a database of D1/D2 regions from these results (Hanafy et al., 2020). They found high variation in sequence length (141-250 bp, with 75% of sequences between 182-208 bp) and divergence cutoffs (0.4-21%, with 75% of pairwise divergence values between 1.7-6%) in the ITS1 region, and much lower variation in sequence length (740-767 bp) and divergence cutoffs (0.1-9.2%) in the 28S D1/D2. Within-strain length variability and sequence divergence were lower for the 28S D1/D2 than the ITS1. Using the D1/D2 region, Hanafy *et al*. identified multiple novel candidate genera, suggesting that the use of the ITS1 as a taxonomic marker may obscure the true diversity of this taxon.

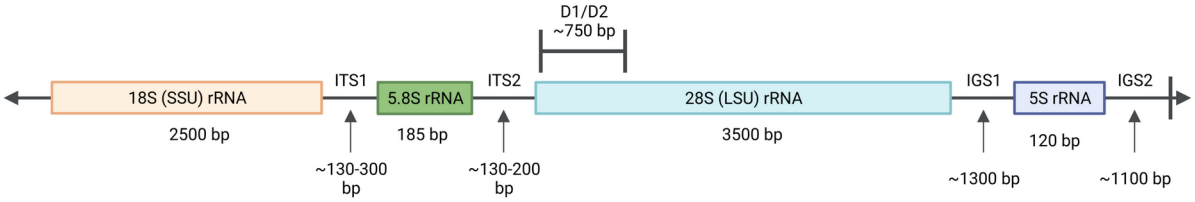


Figure 2. Diagram of a fungal genomic region encoding ribosomal RNA (rRNA). SSU = small ribosomal subunit, LSU = large ribosomal subunit, ITS = internal transcribed spacer, IGS = intergenic spacer. Sizes of regions shown are based on Liggenstoffer *et* al. (2010), Genbank reference sequence AJ864475, and Morrison et al. (2020). Adapted from Edwards *et al.* (2017). Made with Biorender.

1.6.3. Assessing the composition and activity of herbivore gut microbiota

Meta-omics studies enable systems-level characterization of the herbivore gut microbiome. Combining meta-omics tools enables researchers to capture snapshots of which microbes are present, their transcriptional and metabolic capabilities and activities, and the metabolites they consume, produce, and tansform (Leggieri et al., 2021). Culture-

independent approaches to investigating the complex community of the herbivore gut microbiome require an investigator to account for its multi-domain composition and the presence of rare and novel strains, as well as eukaryotic populations. Metagenome-assembled genome (MAG) reference databases such as the Genomic Encyclopedia of Bacteria and Archaea (GEBA) and the Hungate1000 collection have facilitated the study of herbivore gut microbiota (Mukherjee et al., 2017; Whitman et al., 2015; D. Wu et al., 2009). However, simultaneous MAG assembly and alignment of the transcriptome to the reconstructed MAGs captures the activity of isolates not entered into these databases. In Peng *et al.*'s assembly and metabolic investigations of MAGs from the goat gut microbiome, 94% of the 719 assembled prokaryotic MAGs were previously undescribed at the species level, with many of these being unidentified below the family, order, or class level, and as described in Chapter 2, many of these were transcriptionally active members of their enrichment consortia (Peng et al., 2021). Rare strains require large sample volumes and sufficient sequencing depth to capture sufficient nucleic acid material to characterize their identity and activity (H. Jin et al., 2022). The study of gut eukaryotic populations poses special technical challenges. Those specific to fungi are covered in Section 1.7.3.1. Ciliate protists (Entodiniomorpha and Holotricha), the other major eukaryotic group in the herbivore gut, have a high AT genome content, have two nuclei (the macro- and micronucleus), and suffer from a lack of reference sequences (Williams et al., 2020).

Combining meta-omic and culturomic techniques permits the more comprehensive identification of strains that cannot be identified from homologous sequences alone, allows the in-depth study of isolates from a microbiome, and aids in identifying the source of metagenomic sequences of unknown origin (Greub, 2012). Seshadri *et* al.'s sequencing and

cataloging of the Hungate1000 collection of gut bacteria and archaea from herbivore gut microbiota is the largest herbivore culturomics effort to date; in the process of guiding the selection and isolation efforts of gut prokaryotes for this collection, Creevey *et al*. found that bacteria of the Actinobacteria, Firmicutes, and Proteobacteria are overrepresented, whereas bacteria of the Bacteroidetes are underrepresented (Creevey et al., 2014; Seshadri et al., 2018). This collection is not freely available, however, to researchers, which motivates isolation and enrichment efforts on a laboratory basis. Without knowledge of appropriate growth conditions, many herbivore gut microbes remain uncultivable (S. Liu et al., 2022).

Bioinformatics pipelines to study the herbivore gut microbiome should account for the taxonomically diverse nature of the gut microbiota, including rare and novel taxa. The choice between rDNA amplicon sequence variant (ASV) -based analysis or operational taxonomic unit (OTU)-based analysis during the process of denoising, or distinguishing sequencing errors from actual nucleotide variation, affects alpha diversity estimations, with some studies suggesting that ASV-based analysis provides finer taxonomic resolution, which is crucial in studies of complex microbiota containing rare taxa (Bharti & Grimm, 2021). Illumina short-read sequencing is a mainstay of metagenome sequencing due to the lower cost, higher accuracies of shorter reads, and greater sequencing depth, while long-read sequencing is critical for sequencing otherwise inaccessible genomic regions (especially in eukaryotic MAGs), though both of these eventually rely on the quality of downstream read processing and genome assembly. A hybrid sequencing approach has been successfully employed to characterize low-abundance strains in the human gut (H. Jin et al., 2022). Several widely used assemblers exist dedicated to metagenome assembly, such as metaSPAdes, MegaHit, and SOAPdenovo2 (D. Li et al., 2015; Luo et al., 2012, 2015; Nurk

et al., 2017). Coassembly enables assembly of all the reads from a project together instead of multiple separate assemblies (multiassembly), and has the advantage of enabling the recovery of low-abundance MAGs and genes as well as reducing duplication error rates, but requires extensive computational resources and may result in gene fragmentation (Delgado & Andersson, 2022). Tools such as MetaHipMer have been developed specifically for coassembly in metagenome-size datasets to attempt to resolve these tradeoffs (Hofmeyr et al., 2020). Aligning metatranscriptomes to a concatenated file of assembled MAGs from the same sample, using a splice-aware aligner to account for eukaryotic RNA splicing, accounts for activity from novel taxa. Metaproteomic and metametabolomic pipelines require other meta-omics analyses to link metabolites to organisms, particularly primary metabolites (Leggieri et al., 2021; X. Zhang & Figeys, 2019). Network-based statistical methods have been developed to analyze these datasets in ways that integrate multiple types of omics data (Jiang et al., 2019).

Genome-scale metabolic (GMM) models and transcriptional regulation network (TRN) models, and multiscale GMMs, represent methods to link metagenomic, metatranscriptomic, metaproteomic, and metametabolomic information together, via combining them with mathematical descriptions of activity via fluxomics (centered around flux balance analysis of metabolic pathways in an organism) and intracellular interactomics (Bi et al., 2022). Nearly all extant GMMs fail to account for gene expression’s impact on metabolism, and integration of TRNs with GMMs allows a metabolic model to represent metabolite flux under a variety of physiological states in an organism (Cruz et al., 2020). Integrating multiple GMMs into consortium-wide models has been employed to model consortia of

organisms to study microbial interactions, and is readily scalable to study the exchange of multiple metabolites between different species (Basile et al., 2020; Garza et al., 2018).

1.6.3.1. Meta-omics of the herbivore gut mycobiome

Technical challenges inherent to fungal genomes and physiology pose the most significant barriers to studying gut mycobiota. Fungi are generally present in herbivore gut microbiota in small proportions (Druzhinina & Kubicek, 2012; Orpin & Ho, 1991; Trinci et al., 1994). The site of collection – foregut samples, different locations along the hindgut, or feces – may result in different relative abundances of fungal taxa; for example, anaerobic fungal community structure varies along the hindgut in horses (Mura et al., 2019). Obtaining their DNA requires penetrating their chitinous cell wall, which is highly resistant to microbial degradation and conventional lysis, but may be penetrated by cautious application of mechanical disruption methods such as freeze-drying, liquid nitrogen, or bead-beating (Haitjema et al., 2014). DNA should be extracted from cultures in mid-log to late-log phase, and a combination of cetyltrimethylammonium bromide (CTAB) extraction and the commercially available QIAGEN PowerPlant® Pro DNA extraction kit has been shown to deliver adequate results for *Piromyces* spp., *Neocallimastix* spp., and *Anaeromyces* spp., but these methods may nonetheless present difficulties in isolating high-quality and high-molecular weight DNA from other fungi, particularly bulbous isolates ((Edwards et al., 2017)).

As explored briefly in Section 1.7.2 with respect to the Neocallimastigomycota, there is a robust case for replacement of the ITS as a taxonomic barcode for fungi as a whole. It has significant disadvantages stemming from its high length, structural, and intra-strain

sequence polymorphism, and the 28S LSU D1/D2 region has emerged as a more favorable region for taxonomic diagnosis among some taxa. The use of the 28S D1/D2 as a taxonomic barcode has been validated previously in Dikarya (Ascomycota and Basidiomycota) and Mucoromycota, with a long history of use as a barcoding region for yeasts, and a database for this region has been developed for the Neocallimastigomycota (Fell et al., 2000; Gade et al., 2017; Hanafy et al., 2020; Kwiatkowski et al., 2012; Scorzetti et al., 2002). Outside the Dikarya, few fungal barcoding studies exist, and early-diverging fungi are generally overlooked or ignored (Reynolds et al., 2022). Treatment of amplicon sequences with lyticase has been shown to improve resolution of the mycobiome through increasing usable fungal ITS reads and community alpha diversity, as well as increasing the likelihood of detection of rare or difficult-to-detect fungal genera (Pierre et al., 2021).

Shotgun sequencing effectively circumvents problems that arise in rDNA amplicon sequencing from fungi, such as primer bias (Tedersoo et al., 2015, 2018). Many fungal genomes are marked by AT-rich repeats; just as with GC-rich regions, these form areas of low sequence coverage, and the genomes of anaerobic fungi are the most AT-rich among the fungi, from 25-29% GC (Wilken et al., 2020). As of the publication of this dissertation, there are eight public genome assemblies of fungi of the Neocallimastigomycota; the sequencing of *Piromyces* sp. E2 employed a combination of Sanger and Illumina Solexa sequencing to obtain a fragmented assembly with 39.7% of scaffolds in gaps, and all subsequent assemblies published at the Department of Energy Joint Genome Institute's MycoCosm portal employed hybrid sequencing approaches to assemble with sufficient quality (Haitjema et al., 2017). Shotgun sequencing, which employs short reads, has been used once before to assemble high-quality anaerobic fungal MAGs from metagenomes, and

has effectively been used in functional metagenomics of fungal MAGs; however, Peng *et al.*'s assembled fungal MAGs were, on average, 73% complete, with 14% duplicated Benchmarking Universal Single-Copy Orthologs (BUSCOs)(Peng et al., 2021).

1.7. Approaching microbial relationships in lignocellulolytic gut-derived microbiota through multi-omic and multispecies perspectives

Relationships in lignocellulolytic microbiota from herbivores, and in derived consortia, may be understood through a variety of omics approaches, methods of perturbation, and different gut microbiome structures across herbivore species. In Chapter 2, we challenged inocula from a goat fecal source microbiome with three different complex substrates, as well as purified hemicellulose, to probe the effects of complex substrates on the transcriptional activity of enrichments from a single source microbiome. In Chapter 3, we sequenced the 16S and ITS2 regions of fecal microbiota from *Colobus guereza* and *Gorilla gorilla gorilla*, two non-human primates with very different diets and gut physiologies, and employed enrichment on diet-matched substrates in order to identify rare taxa in their gut microbiota with potential key roles in lignocellulolysis. These two studies approach two different facets of how microbial relationships occupy a critical role in biomass breakdown using data from three species with very different approaches to herbivory. Synthesizing multi-level omics datasets across a variety of conditions provides the fullest picture of lignocellulolytic programming in a gut microbiome. Expanding this view to cover many different configurations of lignocellulolytic gut microbiota allows us to survey the many strategies nature has developed to rise to the challenge of degrading Earth's most abundant renewable resource and to unearth deeper commonalities that unite these different strategies. Finally,

engineers can derive inspiration from the strategies nature has developed to solve these problems in order to generate new approaches to the design and manipulation of lignocellulose-degrading microbiota in order to harness this abundant natural resource to meet a variety of potential biotechnological needs.

2. Lignocellulose complexity shapes the transcriptional activity and community function of goat gut microbiota

The prior basis for this work is discussed in Peng *et al*. (*Nature Microbiology*, 2021).

2.1. Introduction

The mammalian herbivore digestive tract exploits the collective action of a complex community of anaerobic microbes from all domains of life to degrade and metabolize plant biomass, delivering short-chain fatty acids (SCFA) and other crucial fermentation products to their dependent animal host as well as facilitating the metabolism of plant toxins (Demment and Van Soest, 1985; Mackie, 2002; Kohl et al., 2014). The efficiency with which these microbiomes degrade lignocellulose, as well as the wealth of transporters, biomass-degrading enzymes, and secondary metabolite production pathways within the genomes of the microbes in these consortia, make herbivore gut microbiomes attractive candidates for biotechnological development (Podolsky et al., 2019). Evidence from both animal studies and enrichment cultures indicates that the choice of carbon substrate shapes the composition and activity of a microbial community in a way that persists over time, and, in industry, altering substrate composition has been employed successfully as a method for greenhouse gas mitigation (Gruninger et al., 2019; Fliegerova et al., 2021; Snelling et al., 2019; Wong et al., 2016; Peng et al., 2021; Haque, 2018). More mechanistic studies of substrate shaping have focused on defined consortia growing on chemically defined substrates (Rakoff-Nahoum et al., 2016; Deng & Wang, 2016; Deng & Wang, 2017). However, the mechanisms by which this occurs are still poorly understood through a metatranscriptomic lens. Additionally, to our knowledge, no such study has investigated the

effects of enrichment on multiple complex lignocellulosic substrates on the shaping of consortium activity. These enrichment conditions remove the influence of the animal host, and faithfully represent the impact of actual plant biomass in shaping the metabolic activity of microbiota that evolved to degrade this biomass, and not simpler, industrially-refined polysaccharides.

To understand how lignocellulosic substrate choice drives metabolic function and partnerships in lignocellulolytic anaerobic microbiomes, we performed parallel enrichment experiments to shape consortia derived from goat feces and identified key microbes within each substrate driving carbon metabolism, consortium stability, and participating in within- and cross-domain partnerships. We challenged four sets of enrichment communities with one of four biomass substrates (alfalfa, bagasse, reed canary grass, and xylan) and three different antibiotic treatment states (untreated to enrich for prokaryotes, penicillin-streptomycin to enrich for methanogens and fungi, and chloramphenicol to enrich for fungi only), then sequenced shotgun metagenomes and marker genes from these communities in order to track taxonomic and functional diversity, and also sequenced both ribosomal RNA-depleted and poly(A) enriched RNA libraries in order to track transcriptional activity as a proxy for consortium metabolism. The first study derived from this dataset recovered 686 novel uncultured metagenome-assembled genomes (MAGs) out of a total of 719 prokaryotic MAGs unique at the species level, as well as 18 MAGs from fungi of the phylum Neocallimastigomycota, and established that community composition of these cultures attained stability after more than ten culture generations, staying unchanged after cryopreservation (Peng et al., 2021).

In this analysis of metatranscriptomic data from the set of antibiotic-untreated consortia enriched by Peng *et al*., we aligned metatranscriptomic reads to the genomes of the 719 prokaryotic MAGs and the reference genome of *Neocallimastix californiae* (K.V. Solomon et al., 2016) to uncover overall patterns in metabolic pathway expression on each substrate, identify key MAGs driving metabolic activity on each substrate, and reconstruct the metabolic activity and interactions of the MAGs in each consortium in conjunction with metabolomic measurements. Differences in community composition and function among the four carbon substrates emerged, with xylan supporting a minimal community metabolically dominated by one MAG and exhibiting sparse methanogen activity, and the three lignocellulosic substrates (alfalfa, bagasse, and reed canary grass) supporting larger, more complex communities of bacteria and methanogens with differences in their hydrolytic, fermentative, and methanogenic strategies that reflected substrate nutritional composition. Relative abundance did not reflect contributions to metabolic activity in lignocellulolytic consortia. Carbohydrate-active enzyme (CAZyme) expression exhibited a division of labor suggesting complementary broad-spectrum and redundancy strategies in lignocellulolysis. This selective enrichment of the same source microbiome on four different substrates, and the resultant differences in community structure and activity between these enrichments, provides novel mechanistic insight into how the chemical composition of the carbon substrate provided to an anaerobic microbiome (such as that of the rumen) drives that microbiome's metabolic responses and suggests a simple strategy for tuning the chemical productivity of a microbial consortium with minimal additional intervention by altering its nutrition source.

2.2. Methods

2.2.1. Fecal sample collection, anaerobic enrichment, and metabolomic data collection

A full description of the methods used for collection, extraction, library preparation, sequencing, and metagenomic assembly and analysis is available in Peng et al. (2021). Peng et al. collected feces from a San Clemente Island goat (*Capra aegagrus hircus*) at the Santa Barbara Zoo (Santa Barbara, CA) was used to inoculate batch cultures in serum bottles, containing complex medium (MC-) with 7% ovine rumen fluid and crude substrate ground to 4 mm in size, under a carbon dioxide headspace. They provided to each culture one of four carbon substrates (alfalfa stems, bagasse, reed canary grass, xylan) and treated it with one of three antibiotic conditions (antibiotic-free to enrich for prokaryotes, 2 mg/ml penicillin and streptomycin to enrich for methanogens and fungi, or 2 mg/ml chloramphenicol to enrich for anaerobic fungi), with three biological replicates per culture, for a total of 36 batch cultures. This manuscript uses only the data derived from antibiotic-free cultures. Peng et al. monitored activity daily by sampling the gaseous headspace of each batch culture to measure concentrations of hydrogen (H_2) and methane (CH_4) and accumulated pressure. Concurrently, they sampled 1 ml of liquid medium from each batch culture to measure the production of short-chain fatty acids and ethanol via high-pressure liquid chromatography (HPLC), and to measure the production of reducing sugars, using both a dinitrosalicylic acid assay (Wood et al. 2012, Miller 1959) and HPLC. After subculturing into fresh media following three days of growth, for a total of ten subcultures ("generations", designated G0 through G10), they harvested the remainder of the culture for nucleic acid extraction.

2.2.2. Nucleic acid extraction and sequencing

Peng et al. (2021) extracted DNA and RNA from the same sample following the QIAGEN AllPrep® DNA/RNA/miRNA Universal handbook using the protocol for “cells” with the following modifications. They thawed frozen biomass pellet samples and centrifuged them at 4°C for ten minutes at 12,000 rpm in a fixed angle rotor (Eppendorf F-34-6-38). They decanted and discarded the supernatant using a pipette. They transferred the pellets into 2-ml bead beating tubes containing 1 ml of 0.5 mm zirconia/silica beads (Biospec product #11079105z) and 700 µl of buffer RLT plus from the QIAGEN AllPrep® DNA/RNA/miRNA Universal kit (Cat. No. 80224). All samples were lysed on a Biospec Mini-BeadBeater-16 for one minute and thirty seconds, followed by cooling on ice for two minutes before a second round of bead beating for one minute and thirty seconds. All bead tubes containing samples were centrifuged at 13,000 rpm for 3 minutes (Eppendorf FA-45-24-11). The supernatant in each sample was transferred into a 1.5-ml microcentrifuge tube and centrifuged again at 13,000 rpm for 3 minutes before nucleic acid extraction. Peng et al. measured the resultant quantity and quality of DNA and RNA using a Qubit (ThermoFisher Scientific) and TapeStation 2200 (Agilent). For antibiotic-free consortia, Peng et al. depleted ribosomal RNA using the Illumina Ribo-Zero rRNA Removal Kit (Yeast) spiked into the Illumina Ribo-Zero Gold rRNA Removal Kit (Epidemiology). Stranded RNA-seq libraries were created by the Joint Genome Institute and quantified by qPCR. They sequenced libraries by paired-end dual-indexed 150 bp reads using the NovaSeq S4 (Illumina, San Diego, CA) and performed quality control following pipeline version 3.4.0 from bbtools (version 38.20).

2.2.3. Bioinformatic analysis

Peng et al. pre-processed raw reads by the Microbial Genome Annotation Pipeline; BBDuk (Bushnell et al., 2017) (version 38.67) was used to remove contaminants, trim reads that contained adapter sequence and right quality trim reads where quality drops to zero. BBDuk was used to remove reads that contained 1 or more 'N' bases, had an average quality score across the read less than 10, or had a minimum length ≤ 51 bp or 33% of the full read length. Peng et al. removed reads mapped with BBMap to masked human, cat, dog, and mouse references, common microbial contaminants, known spike-ins, and ribosomal RNA references.

I concatenated the genomes of 719 dereplicated prokaryotic MAGs identified by Peng et al. (2021) with the reference genome of *Neocallimastix californiae* (henceforth referred to as 719+G1). A table of the 719 prokaryotic MAGs, their relative abundances, metabolic pathway completenesses, and cohesin/dockerin genome content is available in Peng et al. (2021). To evaluate overall transcriptional activity of each MAG, I aligned metatranscriptomic reads to 719+G1 in the splice-aware aligner BBmap (version 38.63) with the parameters "k=13, minid=0.95, ambig=random", and read counts per contig were obtained using BBtools pileup.sh (Bushnell et al., 2017). These genomes were selected for the reference set because the study carried out by Peng et al. indicated they constituted most of the microbial community. I calculated the reads assigned to each MAG by adding the total number of reads mapping to contigs belonging to each MAG, then a reads-per-kilobase-million (RPKM) value was obtained for each MAG. To evaluate transcriptional activity at individual open reading frames (ORFs), 719+G1 was first preprocessed and

annotated in MetaPathways 2 version 2.5.1 using the RefSeq database (Konwar et al., 2013). I then aligned the metatranscriptomic reads to the preprocessed FASTA file in BBmap with the parameters “k=13, minid=0.95, ambig=random”. ORF prediction annotations by MetaPathways were converted to simple annotation format (SAF) and read counts per ORF were obtained with primary read filtering (counting only primary alignments) using featureCounts v2.0.1 (Y. Liao et al., 2014); these were normalized to RPKM, and then further converted to two relative RPKM (relRPKM) values, using the method outlined in Scarborough et al. (2018): one obtained by dividing each ORF RPKM by the median of all the non-zero RPKMs in the sample, then taking the $\log_2$ of that value, and one obtained by dividing each ORF RPKM by the median of all the non-zero RPKMs in that MAG in that sample. I performed metabolic pathway analysis with the assistance of PathwayTools version 23.5 (Karp et al., 2021). Key MAGs for each substrate, with the exception of xylan, were chosen by taking the 20 most active MAGs with nonzero reported relative abundance. I then annotated key MAGs individually in MetaPathways2 using the RefSeq and metacyc_v5_2011 databases (Caspi et al., 2014; Pruitt, Brown, and Tatusova et al., 2012)).

To assess CAZyme expression, I performed quality control of the subset of ORFs annotated as CAZymes by MetaPathways by mapping the CAZyDB annotation for each ORF to its corresponding RefSeq annotation and curating the annotations according to the following algorithm: ORFs annotated by RefSeq as non-CAZymes were removed from the CAZyme annotation set. In cases where CAZyDB and RefSeq assigned two different annotations to the same ORF and both annotation results were CAZymes, I selected the CAZyDB annotation. I reported a given key MAG’s activity for a given CAZyme in a sample as the highest $\log_2$(relRPKM) for any ORF of that CAZyme, where $\log_2$(relRPKM)

was expressed as the $\log_2$ of the RPKM for that ORF, divided by the median of all the non-zero RPKMs in that MAG in that sample. I performed clustering of MAG and CAZyme relRPKM with the Python module Seaborn v0.11.1 (Waskom, 2021), using the unweighted pair group method with arithmetic mean (UPGMA) clustering algorithm and Euclidean distance metric.

To determine which CAZymes and fermentation pathways were being used most by a given key MAG, I reported that MAG's activity for a given enzyme in a sample as the highest $\log_2$(relRPKM) for any ORF of that enzyme, where relRPKM was expressed as the RPKM for that ORF divided by the median of all the non-zero RPKMs in that MAG in that sample. To determine the relative contributions of each methanogen MAG to methanogenesis activity in a sample, I reported the expression of each ORF for all methanogenesis-associated enzymes in a given methanogen as the highest $\log_2$(relRPKM) for any ORF of that transporter/enzyme, where relRPKM was expressed as the RPKM for that ORF, divided by the median of all the non-zero RPKMs across all MAGs in the sample. In the case of an enzyme with multiple subunits, individual subunits of one enzyme were determined by establishing proximity to each other in the MAG, then all subunits of the enzyme were averaged and the highest subunit average for each gene cluster was taken; the subunit with the lowest RPKM for the gene cluster with the highest average RPKM across all subunits was reported as the RPKM for that enzyme.

2.2.4. Data availability

Metatranscriptome sequencing reads may be accessed at the Joint Genome Institute under the JGI Project IDs listed in Supplemental Table 1 of Dickson et al. (manuscript in

preparation). Contigs for each MAG are available at NCBI's Whole Genome Shotgun database under project number PRJNA530070.

2.3. Results

2.3.1. Challenge with a panel of carbon substrates enriches distinct lignocellulolytic communities shaped by their substrate

Peng *et al*. challenged a source gut microbiome, obtained from goat feces, with four different carbon substrates and two antibiotic treatments to selectively enrich viable consortium members, in a total of thirty-six separate consortia; of these, twelve were antibiotic-free consortia raised on each of the four substrates, with three replicates per substrate, containing only prokaryotes (Peng *et al*. 2021). More than 127 Gbp (1.27 x 1011 bp) of metatranscriptome sequencing from the antibiotic-free consortia was aligned to the genomes of 719 prokaryotic MAGs (694 bacteria, 25 archaea) previously assembled by Peng *et al*., as well as the reference genome of *Neocallimastix californiae*, which was chosen as a proxy for all anaerobic fungi in these consortia, due to having the largest genome size of all published anaerobic fungal genome assemblies to date and the dominance of *Neocallimastix spp.* in anaerobic fungal populations across all consortia except for xylan, in which *Caecomyces spp.* were the dominant fungi (Solomon *et al*. 2016, Peng *et al*. 2021). Peng *et al*. compared these MAGs to 8,178 genomes from 3 of the largest ruminant gut metagenomic datasets available at the time, in addition to the Genomic Encyclopedia of Bacteria and Archaea (GEBA) collection (Mukherjee et al., 2017), a collection of 1,520 genomes from human gut bacteria (Zou et al., 2019), and 221 additional reference genomes from the RefSeq database at the National Center for Biotechnology Information (NCBI) (O'Leary et al., 2016), and found that 677 of the 719 prokaryotic MAGs (94%) in this

dataset were unidentified at the species level. 45% (9/20) of alfalfa key MAGs, 35% (7/20) of bagasse key MAGs, 30% (6/20) of reed canary grass key MAGs, and no xylan key MAGs were unidentified at the species level. Key MAGs unidentified at the species level constituted 9.2%, 15.6%, 13.9%, and 0% of relative abundance at generation 10 in alfalfa, bagasse, reed canary grass, and xylan, respectively.

Peng *et al*. employed amplicon sequence variant (ASV)-based marker gene analysis to track the composition and stabilization time of each enrichment consortium, and found that only a small percentage (0-2.2%) of the source microbiota was enriched in these cultures, with community membership differing by substrate. Unexpectedly, alignment of metatranscriptomic reads from these cultures to the genomes of the 719 prokaryotic MAGs and the G1 reference genome revealed that transcriptional activity in these consortia did not reflect the relative abundance of each consortium member, and there was no statistically significant association between relative abundance and overall activity in the alfalfa and bagasse consortia, where there were three replicates (**Figures 3 and 4**). However, these samples were snapshots of metabolic activity collected at the end of a batch cycle, which may not reflect abundance-activity correlation averaged over the batch cycle. In addition, some MAGs' metabolism may have yielded more energy for the MAG and thus resulted in higher population growth rates. The relative abundance and transcriptional contribution by generations 9 and 10 of the key MAGs common to all lignocellulosic consortia differed for each lignocellulosic substrate.

Consortia grown on the lignocellulosic substrates (alfalfa, bagasse, and reed canary grass) supported diverse communities of bacteria and archaea. Key members of each consortium ("key MAGs") were identified by selecting the 20 most active MAGs in each

consortium at generation 9 with nonzero abundance in generation 10, or, in the case of xylan, selecting all MAGs active at generation 9 with nonzero abundance in generation 10.

The most transcriptionally active MAG in the alfalfa consortia was *Streptococcus equinus*, with *Ruminococcus albus* and Paludibacteraceae UBA1723 as the next most active MAGs. After these MAGs, the relative transcriptional activity of the other MAGs to each other was not consistent across replicates. In the third antibiotic-free alfalfa replicate at passage 10, from which DNA data was taken to calculate relative abundance, *S. equinus* was the fifth most abundant MAG, *R. albus* was the third most abundant MAG, and Paludibacteraceae UBA1723 was the thirteenth most abundant MAG. Of the seventeen remaining key MAGs, two were methanogens of the genus *Methanobrevibacter*, two were members of the Paludibacteraceae, seven were members of the cellulolytic family Lachnospiraceae, two belonged to the species *Prevotella ruminicola*, and the remaining MAGs were *Schwartzia succinivorans* and undescribed members of the Bacteroidales, Oscillospiraceae, and Tissierellales.

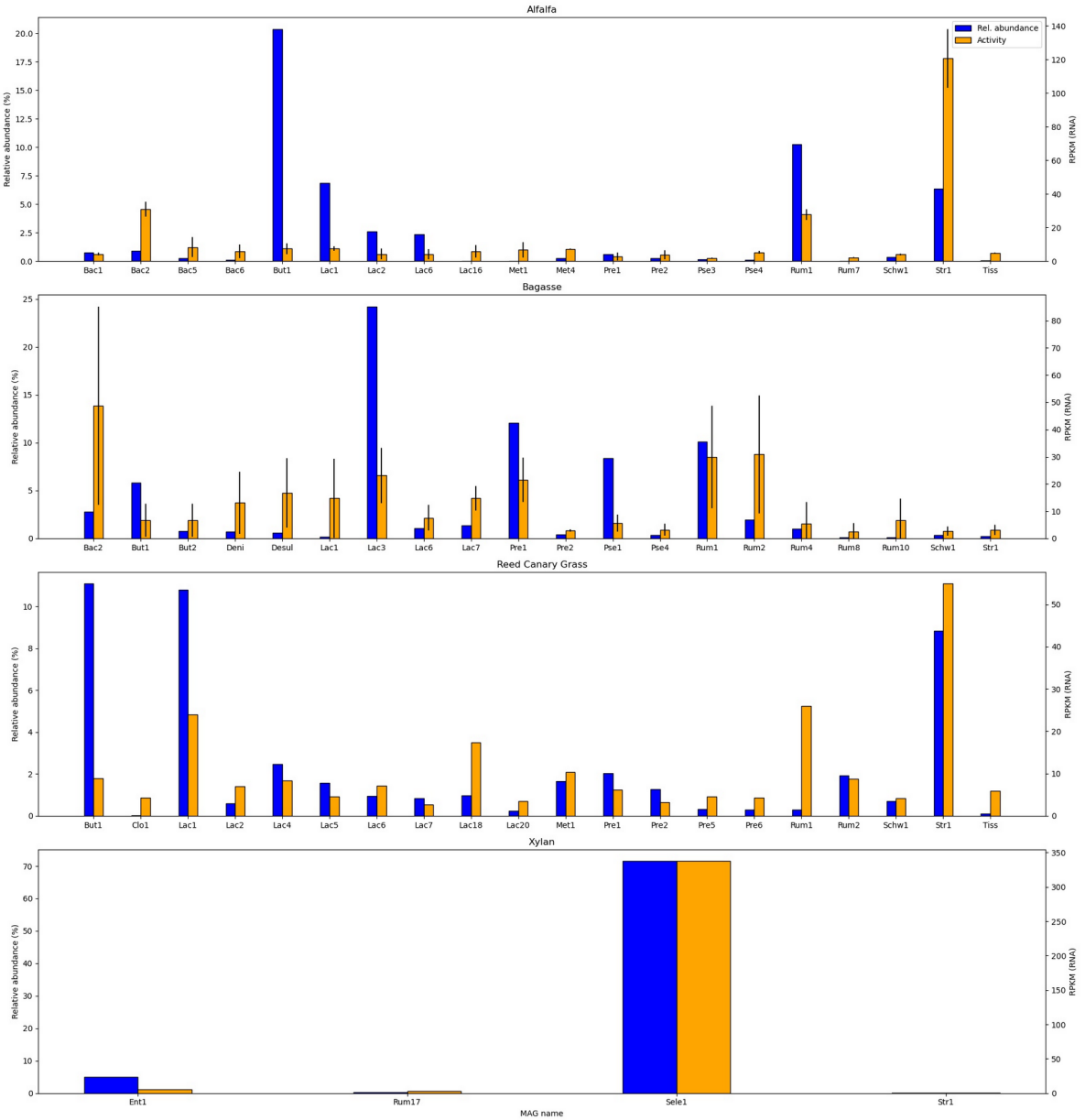


Figure 3. Activity and relative abundance of key MAGs across the carbohydrate substrates alfalfa, bagasse, reed canary grass, and xylan in antibiotic-free consortia. Key MAGs were defined as the top 20 active MAGs in lignocellulosic consortia and as all active MAGs in xylan consortia. Blue bars: relative abundance at generation 10. Orange bars: Overall activity in RPKM at generation 9. Bar heights are means of all replicates at a given generation. For relative abundance, n=1 for alfalfa, bagasse, and reed canary grass, and n=2 for xylan. For overall activity, n=3 for alfalfa and bagasse, n=1 for reed canary grass, and n=2 for xylan.

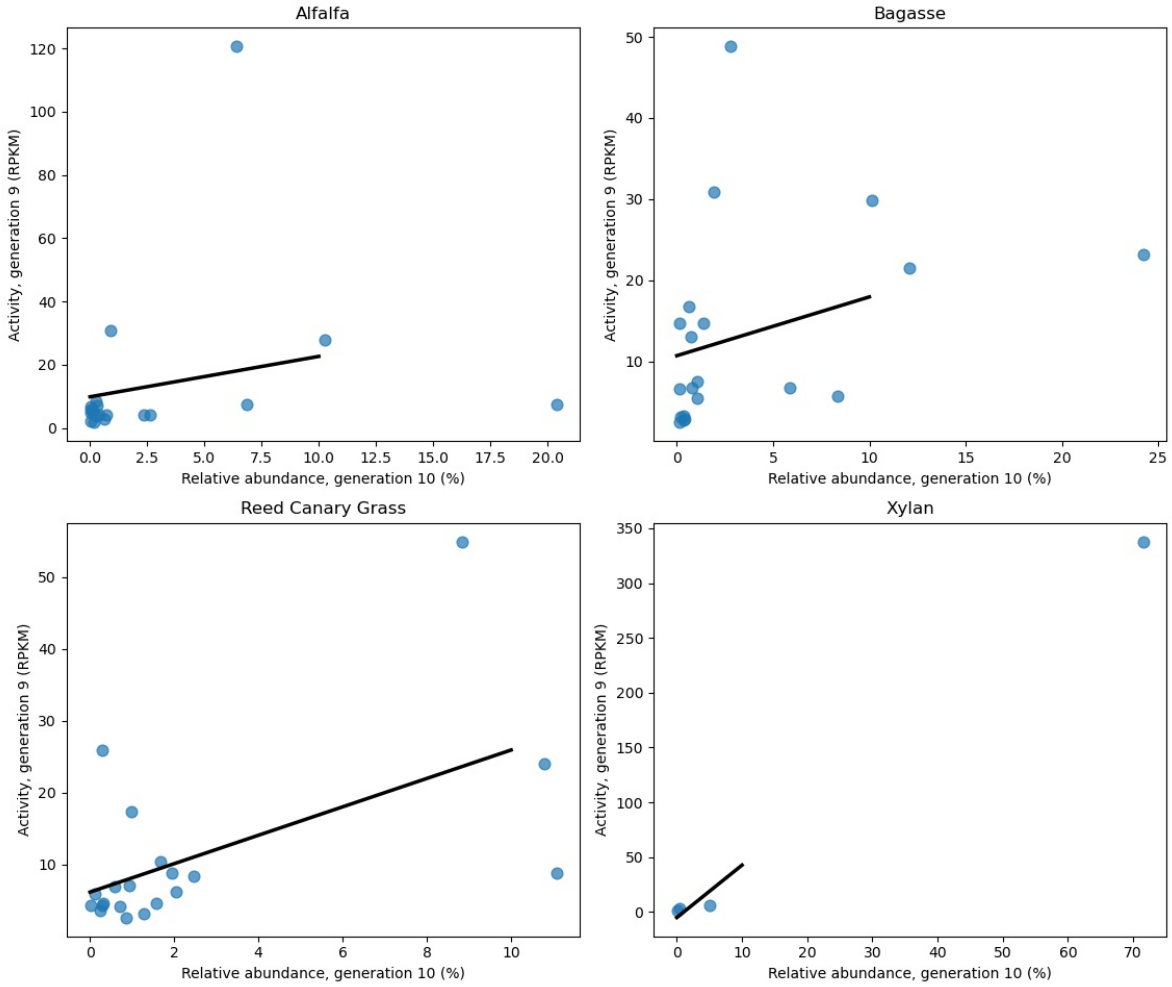


Figure 4. No correlation between relative abundance and overall activity of key MAGs on the substrates alfalfa, bagasse, reed canary grass, and xylan. Scatter plot of mean relative abundance and mean overall activity of key MAGs across the carbohydrate substrates alfalfa, bagasse, reed canary grass, and xylan. Pearson's R^2 and P-value for each substrate: Alfalfa: R^2 = 0.2440971060010743, P = 0.299670393660444. Bagasse: R^2 = 0.3571282414638144, P = 0.12215802451597986. Reed canary grass: R^2 = 0.5617254991481838, P = 0.009953256400154546. Xylan: R^2 = 0.9985532585240124, P = 0.0014467414759875563.

The relative transcriptional activity of MAGs to each other in the microbial community in the bagasse consortia was not consistent across replicates. Paludibacteraceae UBA1723 displayed the largest variance in activity and was the most active MAG in two out of the three replicates at generation 9, and in the second antibiotic-free bagasse replicate at passage 10, Paludibacteraceae UBA1723 was the sixth most abundant MAG. Of the

remaining key MAGs, fourteen belonged to cellulolytic taxa, including eight in the Lachnospiraceae, two in the species *Prevotella ruminicola*, and two in the genus *Ruminococcus* (*R. albus* and *R. flavefaciens*), and two in the family Oscillospiraceae.

The most transcriptionally active MAG in the reed canary grass consortium replicate from which RNA data was available was *Streptococcus equinus*, with *Ruminococcus albus*, Lachnospiraceae 1XD8-76, and *Acetatifactor* sp. as the next most active MAGs. In the third antibiotic-free reed canary grass replicate at generation 10, *S. equinus* was the third most abundant MAG, *R. albus* was the fifth most abundant MAG, Lachnospiraceae 1XD8-76 was the second most abundant MAG, and *Acetatifactor* sp. was the eighteenth most abundant MAG. Of the remaining key MAGs, twelve belonged to cellulolytic taxa, including seven more Lachnospiraceae, three members of the genus *Prevotella* (including two members of the species *Prevotella ruminicola*), and two members of the genus *Ruminococcus* (*R. albus* and *R. flavefaciens*).

The microbial community in the xylan consortia was dominated both in relative abundance and transcriptional activity by *Selenomonas ruminantium*. The remaining members were a bacterium from the family Acutalibacteraceae, *Streptococcus equinus*, and *Enterococcus lactis*. *S. ruminantium* comprised more than 55% of relative abundance in xylan consortia by generation 5 and exhibited more activity than the next most active MAG in each replicate by a factor of at least 36. Str1 never constituted more than 7.5% of relative abundance over all generations sampled.

A core community of eight key MAGs was key in all lignocellulosic enrichment consortia, comprised of six MAGs from the phylum Firmicutes (three Lachnospiraceae, *Ruminococcus albus*, *Schwartzia succinivorans*, and *Streptococcus equinus*) and two MAGs belonging to the phylum Bacteroidetes (both identified as *Prevotella ruminicola*). Of the remaining key MAGs for each substrate, an additional set of eight MAGs was only represented in one or two sets of key MAGs, but present in all three lignocellulosic consortia at generations 9/10: seven members of the Firmicutes (Oscillospiraceae E4, *Pseudobutyrivibrio* sp., Lachnospiraceae bacteria XBD2001 and CAG-194, Tissierellales bacterium PP17-6a, and *Ruminococcus flavefaciens*) and one member of the Actinobacteria (*Denitrobacterium detoxificans*). In addition, each lignocellulosic substrate enriched a subset of key MAGs found either only on that substrate, or that substrate and one other substrate. Alfalfa consortia uniquely enriched a member of Paludibacteraceae genus RF16, a member of Lachnospiraceae genus XBB1006, and *Pseudobutyrivibrio ruminis*. Bagasse enriched a small, but stable and active population of *Desulfovibrio desulfuricans*, as well as one member of the Oscillospiraceae genus CAG-10 and a member of the Butyricoccaceae. Reed canary grass enriched a member of Anaerovoraceae genus RUG099 and one member each of Lachnospiraceae genera UBA2868 and CAG-590. Alfalfa and bagasse shared a member of Paludibacteraceae genus UBA1723. Alfalfa and reed canary grass shared a member of Paludibacteraceae genus UBA4363, a member of Bacteroidales family W3P20-009, *Methanobrevibacter thaueri*, another *Methanobrevibacter* sp., and a member of Lachnospiraceae genus CAG-590. Alfalfa and xylan shared *Selenomonas ruminantium*. Bagasse and reed canary grass shared *Butyrivibrio* sp., a member of Lachnospiraceae genus

G11, *Pseudobutyrivibrio* sp., a member of Oscillospiraceae genus E4, a member of Bacteroidales family UBA2918, and *Prevotella* sp..

2.3.2. Division of cellulolytic labor by substrate

Alfalfa stems, sugarcane bagasse, reed canary grass, and xylan are chemically heterogeneous complex substrates with different proportions of lignocellulose components to each other, and demand degradation by different sets of CAZymes (**Table 5**). I wanted to understand how the chemical composition of each substrate shaped the division of cellulolytic labor. I calculated relRPKMs of ORFs encoding families of CAZymes from a list of cellulases, hemicellulases, pectinases, and esterases (CHPEs) previously compiled by Peng *et al*. from the Carbohydrate-Active Enzymes Database (http://www.cazy.org)(Drula et al., 2021), curated based on the catalytic domains present in each CAZyme. relRPKM was selected over RPKM to report both intra-consortium and inter-consortium variability in expression due to its ability to capture the degree to which a given MAG was expressing an enzyme of interest relative to other genes in its genome and thus serve as a simple way to represent what a MAG was metabolically prioritizing; by contrast, overall enzyme RPKM captures overall transcriptional activity, but does not provide insight into how a given MAG is prioritizing gene expression and is affected by the relative abundance of that MAG. Clustering of RPKM values using the UGPMA algorithm with Euclidean distance metric was carried out in order to discern clusters of similar CAZyme expression. The CAZyme families GH43, GH51, and CE4 were found across all substrates to be members of these clusters and highly redundant in expression across many members of each consortium; GH43, in particular, was expressed by more than half of each substrate's key MAGs (**Figure 6; Supplemental Figure 1A, Appendix 2**). Additionally, among all three lignocellulosic

consortia, GH5 and GH10 were also found to be members of these clusters, and broadly expressed by key MAGs on their respective substrates.

Substrate	Composition	Major Polysaccharide Bond Types	Polysaccharide Degradation Products	Putative Major CAZymes Involved in Degradation	References
Alfalfa stems	**Pectin:** 117-157 g/kg dry matter. 50% homogalacturonan, 50% rhamnogalacturonan I/II. **Cellulose:** 239-348 g/kg dry matter. **Lignin:** 96-200 g/kg dry matter (Klason lignin). **Ash:** 9% dry matter (mean). **Hemicellulose:** 87-128 g/kg dry matter. **Overall monosaccharide composition:** 204-369 g/kg glucose, 64-115 g/kg xylose, 8-30 g/kg arabinose, 14-22 g/kg galactose, 13-24 g/kg mannose, 5-8 g/kg rhamnose, 0.9-2 g/kg fucose, 71-109 g/kg uronic acids.	**Pectin:** Homogalacturonan (HG): α-1,4-linked D-GalA unbranched homopolymer backbone. GalA residues may be methylesterified at the C6 position and/or acetylated at C2 or C3. Rhamnogalacturonan I (RG-I): →2)-α-L-Rha*p*-(1→4)-α-D-Gal*p*A-(1→ backbone with 12% C4-linked Rha*p* forming side chains by L-Ara and D-Gal oligosaccharide . Rhamnogalacturonan II (RG-II): (1→4)-α-D-Gal*p*A unbranched homopolymer backbone with 4 possible side chains: (1) L-Gal*p*-(1→2)-β-D-Glc*p*A-(1→4)-α-L-Fuc*p*-(1→4)-β-L-Rha*p*-(1→3')-β-D-Api*f*-(1→2)-backbone; D-Api*f* units may crosslink with each other by a borate diester bond; 2-O-methyl-α-D-Xyl*p*-(1→2)-sidechain α-L-Fuc*p*; α-D-Gal*p*-(1→2)- and β-D-Gal*p*-(1→3)- sidechain β-L-Rha*p*; (2) 2-O-methyl-α-D-Fucp-(1→2)-β-D-Galp-(1→2)-α-L-Ace*f*A-(1→3)-β-L-Rhap-(1→3')-β-D-Api*f*-(1→2)-backbone; α-L-Rhap-(1→2)-α-L-Ara*p*-(1→4)-sidechain D-Gal*p;* (3) α-D-Gal*p*A-(1→5)-α-D-Kdo-(2→3)-backbone α-D-Gal*p*A; (4) β-L-Ara*f*-(1→5)-β-D-Dha-(2→3)-backbone α-D-Gal*p*A. **Cellulose:** (1→4)-β-D-Glu. **Hemicellulose:** (sole structurally identified hemicellulose) 1,4-linked β-D-Xyl*p* backbone with 4-O-MeGlc*p*A substitution at O-2 position and T-L-Ara*f* substitutions at O-3 position.	**Pectin:** HG: Galacturonate, acetate, methanol. RG-I: Rhamnopyranose, galacturonate, arabinose, galactose. RG-II: Galacturonate, 3-deoxy-D-manno-oct-2-ulosonate, L-arabinofuranose, 3-deoxy-lyxo-heptulosarate, fucopyranose, rhamnopyranose, galactopyranose, apiose, acerate, L-arabinopyranose, glucuronate, methanol. **Cellulose:** Cellobi/triose, D-glucopyranose. **Hemicellulose:** Xylopyranose, methanol, glucuronate, arabinofuranose.	**Pectin: GH: 4, 28. PL: 1-4, 9-11, 26. CE: 8, 12, 14, 16. Cellulose: GH: 1, 3, 5-10, 12, 26, 39, 44-45, 48, 51, 124, 148. Hemicellulose:** GH: 1, 2, 3, 5, 6, 8, 9, 10, 11, 12, 16, 26, 30, 39, 43, 44, 45, 31, 51, 52, 54, 62, 74, 98, 120, 159. CE: 1-7, 12, 15, 16.	Archibald et al., 1962; Bar-Peled et al., 2012; Chen et al., 2015a; Chen et al., 2015b; Drula et al., 2022; Du et al., 2022; Jung & Engels, 2002; Jung & Lamb, 2004; Jung & Lamb, 2006; Mozaffari et al., 2000; Samac et al., 2004
Sugarcane bagasse	**Pectin:** 0.18% dry matter. **Cellulose:** 32-45% dry matter. **Lignin:** 16.1-32% dry matter. **Ash:** 1-9% dry matter. **Hemicellulose:** 30.3% dry matter. Xylose (68.6–76.6%), arabinose (12.8–15.6%), glucose (7.4–13.1%), uronic acid (3.5–4.8%), galactose (1.9–3.0%), rhamnose (trace–1.1%), and mannose (trace–0.3%)	**Pectin:** HG: As alfalfa. RG-I: As alfalfa. RG-II: As alfalfa. **Cellulose:** (1→4)-β-D-Glu. **Hemicellulose:** (1→4)-β-D-Xyl*p* backbone; side chains: α-L-Ara*f* residues linked to C3 of backbone D-Xyl*p* with a ferulic acid group linked at C5; 4-O-methyl-D-GluA units are also linked to the backbone (Kato et al., 1987; Sun, J.X. et al., 2004). It is unclear what proportion of the D-Xyl*p* backbone has either side chain.	**Pectin:** HG: As alfalfa. RG-I: As alfalfa. RG-II: As alfalfa. **Cellulose:** Cellobi/triose, D-glucopyranose. **Hemicellulose:** Xylopyranose, arabinofuranose, ferulic acid, glucuronate, methanol.	**Cellulose:** GH: 1, 3, 5-10, 12, 26, 39, 44-45, 48, 51, 124, 148. **Hemicellulose:** GH: 1, 2, 3, 5, 6, 8, 9, 10, 11, 12, 16, 26, 30, 39, 43, 44, 45, 31, 51, 52, 54, 62, 74, 98, 120, 159. CE: 1-7, 12, 16.	Alokika et al., 2021; Bar-Peled et al., 2012; Drula et al., 2022; Du et al., 2022; Kato et al., 1987; Sun, J.X. et al., 2004
Reed canary grass	**Pectin:** 10-20 g/kg dry matter (estimation as cool-season grass) **Cellulose:** 25.50-45.31% dry matter. **Lignin:** 16.69-22.73% dry matter (Klason lignin) **Ash:**10.1% dry matter (mean). **Hemicellulose:** 23.64-31.27% dry matter.	**Pectin:** HG: As alfalfa. RG-I: As alfalfa. RG-II: As alfalfa. **Cellulose:** (1→4)-β-D-Glu. **Hemicellulose:** Unknown; contains xylan.	**Pectin:** HG: As alfalfa. RG-I: As alfalfa. RG-II: As alfalfa. **Cellulose:** Cellobi/triose, D-glucopyranose. **Hemicellulose:** Xylopyranose; arabinofuranose and/or glucuronate.	**Cellulose:** GH: 1, 3, 5-10, 12, 26, 39, 44, 45, 48, 51, 124, 148. **Hemicellulose:** GH: 1, 2, 3, 5, 6, 8, 9, 10, 11, 12, 16, 26, 30, 39, 43, 44, 45, 31, 51, 52, 54, 62, 74, 98, 120, 159. CE: 1-7, 12, 16.	Allison et al., 2012; Archibald et al., 1962; Bar-Peled et al., 2012; Drula et al., 2022; Du et al., 2022; Feng et al., 2002; Li, J. et al., 2014; Moore & Hatfield, 2015

Corn core xylan	**Pectin:** None. **Cellulose:** None. **Lignin:** None. **Ash:** None. **Hemicellulose:** 50% xylose, 15% arabinose, 20% glucose, 10% galactose, 2.5% mannose, 2.5% glucuronate.	**Hemicellulose:** (1→4)-β-D-Xyl*p* backbone. Contains Ara*f* and/or GluA. Origin and bonds of other components unclear.	**Hemicellulose:** Xylopyranose, arabinofuranose, glucuronate, glucose, galactose, mannose.	**Hemicellulose:** GH: 1, 2, 3, 5, 6, 8, 9, 10, 11, 12, 16, 26, 30, 39, 43, 44, 45, 31, 51, 52, 54, 62, 74, 98, 120, 159. CE: 1-7, 12, 16.	Drula et al., 2022; Melo-Silveira et al., 2012

Table 5. Chemical composition of alfalfa stems, sugarcane bagasse, reed canary grass, and corn core xylan. All monosaccharide abbreviations taken from the Symbol Nomenclature for Glycans (SNFG).

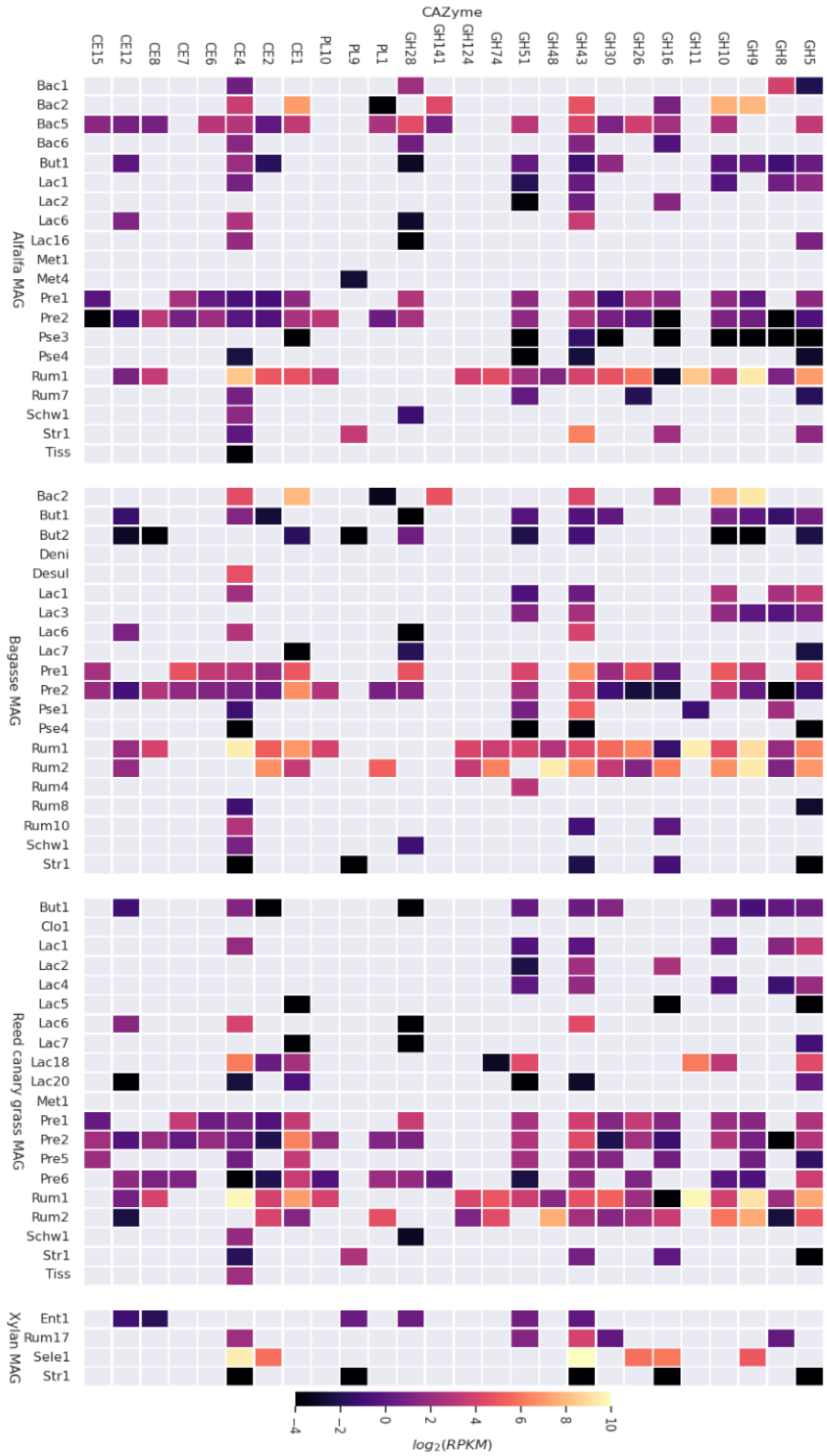


Figure 6. Heatmap of carbohydrate-active enzyme (CAZyme) expression, measured in RPKM, in key MAGs across the carbohydrate substrates alfalfa, bagasse, reed canary grass, and xylan. Cell values are means across all biological samples.

The alfalfa consortia exhibited a pattern of CAZyme secretion reflecting the hydrolysis of a pectinous substrate (**Figure 7**). The predominant enzymes participating in overall biomass breakdown were the glucanase family GH9, the mostly-xylanase (with some glucanase activity) family GH10, and the acetyl xylan esterase family CE1 from Paludibacteraceae UBA1723, the mixed-activity family GH5, GH9, the xylanase/xylosidase family GH11, the acetyl xylan esterase family CE4, and the pectate lyase family PL10 from *Ruminococcus albus*, and the mixed xylanase family GH43 and pectate lyase family PL9 from *Streptococcus equinus* (**Figure 6**). Two clusters of relatively broad CAZyme expression across multiple MAGs emerged after UPGMA clustering; these were GH28 and CE4, and GH5, GH16, GH43, and GH51. *S. equinus*, Paludibacteraceae UBA1723, and *R. albus*, the three most active MAGs in these consortia, expressed high levels of GH43, PL9, GH9, GH10, CE1, GH5, GH11, and CE4. Paludibacteraceae UBA1723 and *R. albus* liberated most acetate from xylan via secretion of CE1 and CE4. Pectin degradation, critical to the metabolism of the high-pectin alfalfa (11.7-15.7% of dry matter, **Table 5**), was accomplished primarily by *S. equinus*, then *R. albus*, with additional input from *P. ruminicola* and Bacteroidales W3P20-009. *S. equinus* secreted most consortium pectinase in the form of PL9 (likely rhamnogalacturonan endolyase, as other activities of this family overlap with PL1 and PL10, which target homogalacturonan, and *S. equinus* was the only MAG filling this niche), *R. albus* and *P. ruminicola* secreted PL10 (pectate lyases, targeting unmethylated homogalacturonan, including exo-pectate lyases), and Bacteroidales W3P20-009 and *P. ruminicola* secreted PL1 (likely pectin lyases targeting methylated homogalacturonan, or if any pectate lyase activity, endo-pectate lyase). Bacteroidales W3P20-009, *P.* ruminicola, *R. albus*, *Butyrivibrio* sp., and *Enteroclostridium*

clostridioformis also secreted smaller amounts of CE8 and CE12, releasing methanol and acetate (respectively) from pectin. The four core producers of the most diverse CAZyme repertoires, made up of both *Prevotella ruminicola* MAGs, *R. flavefaciens,* and Bacteroidales W3P20-009, constituted 0.2-10.3% of relative abundance and together accounted for 11.3% of relative abundance and produced all but one of the CAZymes found in these consortia.

The consortia enriched on bagasse secreted a repertoire of CAZymes reflective of their low-pectin, relatively chemically heterogeneous monocot substrate (**Figure 8**). The predominant enzymes participating in overall biomass breakdown were the families GH9, GH10, and CE1 from Paludibacteraceae UBA1723, the families GH5, GH9, GH11, CE1, and CE4 from *R. albus*, and the families GH5, GH9, and GH10, the glucanase families GH48 and GH74, and the acetyl xylan esterase family CE2 from *R. flavefaciens* (**Figure 6**, **Figure 8**). The first cluster of the most broadly-expressed CAZymes consisted of GH5, GH8, GH43, GH51, and CE4. A second cluster consisted of GH9, GH10, GH16, and CE1. *R. albus* and *R. flavefaciens* were the two most overall active transcribers of CAZymes in this consortium in both magnitude and diversity of transcription. The minimal amount of pectin on this substrate was primarily degraded by *P. ruminicola*, *R. albus*, and *R. flavefaciens*, via CAZymes PL1 and PL10. *R. flavefaciens* secreted most PL1, probably targeting methylated pectin, and *R. albus* secreted most PL10, probably targeting unmethylated pectate. (*P. ruminicola* secreted smaller amounts of both pectinases, but secreted a number of other carbohydrate esterases unsecreted by these other MAGs.) *S. equinus* secreted no pectinases on this substrate. *R. albus* liberated most methanol from pectin. Both *P. ruminicola* MAGs, *R. albus,* and *R. flavefaciens* made up the group of

MAGs expressing a broad spectrum of CAZymes (0.4-12.1% at G10, together comprising 24.0% of relative abundance) and produced all but one of the CAZymes found in these consortia.

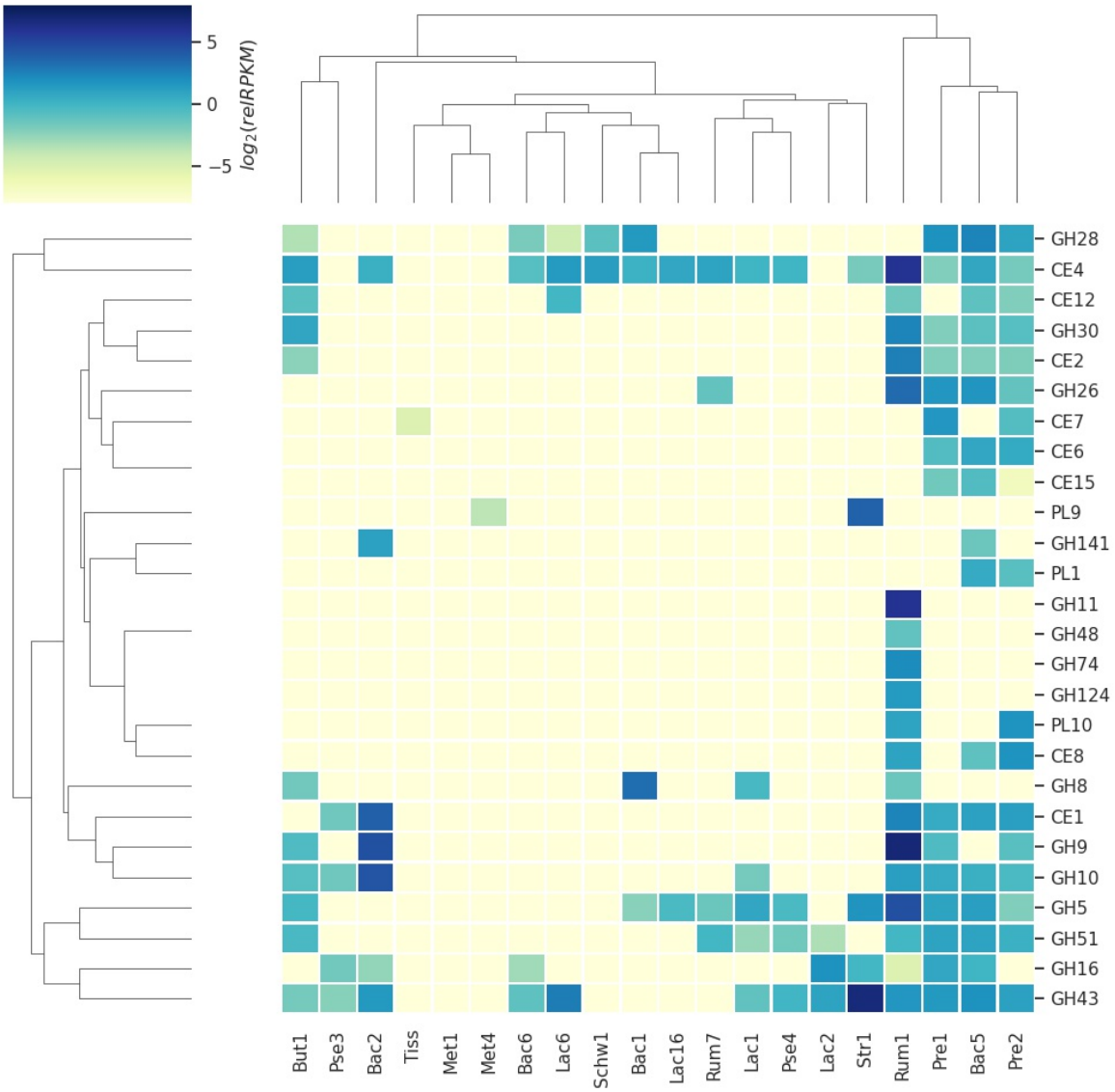


Figure 7. Clustermap of CAZyme expression in key MAGs in generation 9 on alfalfa. log_2(relRPKM) = log_2[(RPKM for designated ORF)/(median of all expressed RPKMs for that MAG in the sample)]. Cell values are means across all biological samples. Refer to Appendix 2, Supplemental Figure 1A for presence/absence of unexpressed enzymes.

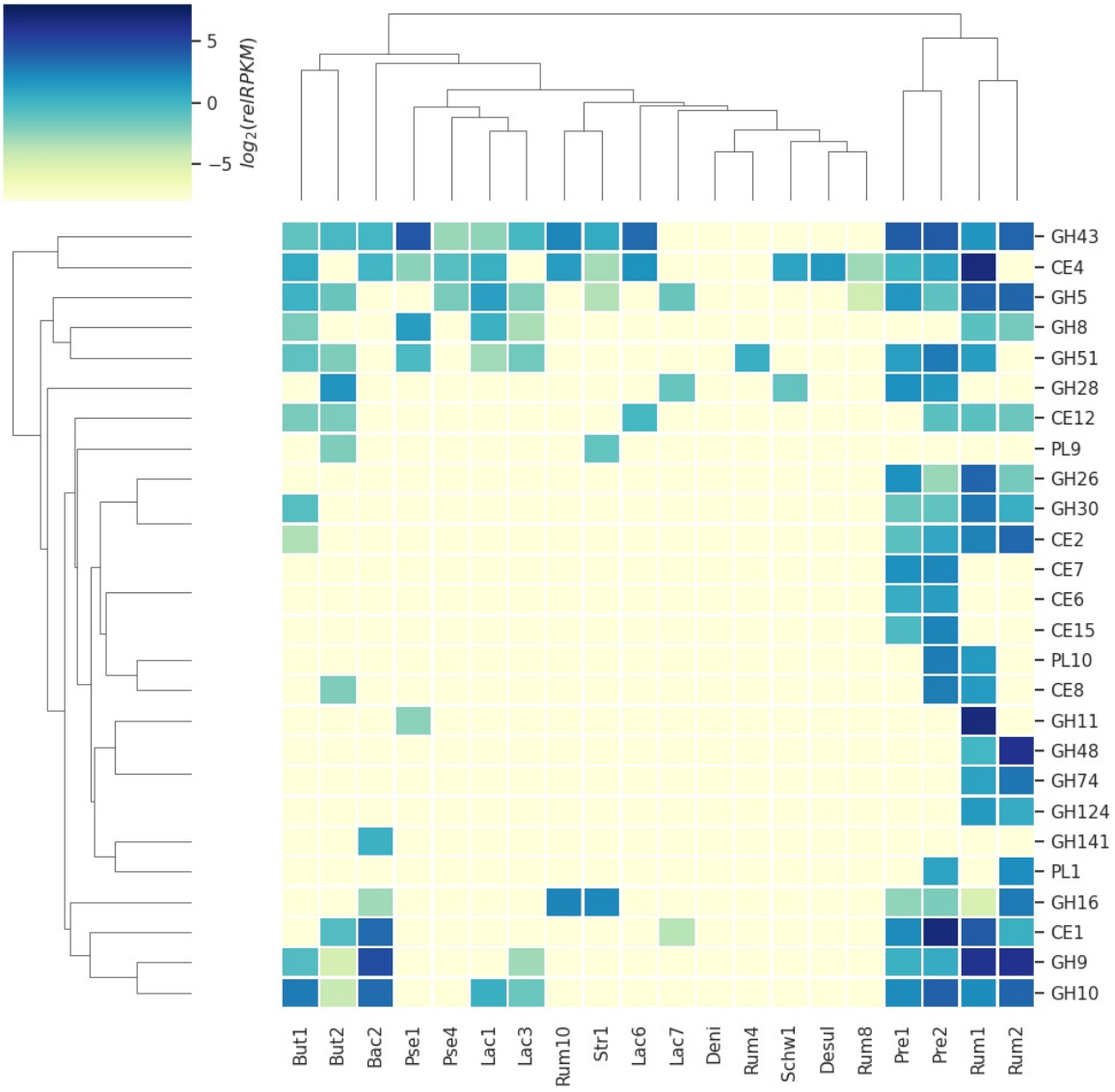


Figure 8. Clustermap of CAZyme expression in key MAGs in generation 9 on bagasse. log_2(relRPKM) = log_2[(RPKM for designated ORF)/(median of all expressed RPKMs for that MAG in the sample)]. Cell values are means across all biological samples. Refer to Appendix 2, Supplemental Figure 1A for presence/absence of unexpressed enzymes.

The consortia enriched on reed canary grass secreted a repertoire of CAZymes reflective of their low-pectin, relatively chemically homogeneous monocot substrate (**Figure 9**). The predominant enzymes participating in overall biomass breakdown were the families GH11 and CE4 from Lachnospiraceae 1XD8-76, the family CE1 from *Prevotella ruminicola*, the families GH5, GH9, GH11, CE1, and CE4 from *R. albus*, and the families GH9 and GH48

from *R. flavefaciens*. *Streptococcus equinus*, Lachnospiraceae 1XD8-76, and *Acetatifactor* sp., the most active MAGs in this consortium overall, expressed high levels of GH11, CE1, PL9, GH5, GH9, CE1, and CE4. However, *R. albus* and *R. flavefaciens* dominated CAZyme transcription both in magnitude and diversity. Both *P. ruminicola* MAGs, an unidentified *Prevotella* sp., *R. albus*, and *R. flavefaciens* formed a core of now relatively low-abundance (1.3-4.5% at G10, together comprising 9.8% of relative abundance) but active producers of a diverse CAZyme repertoire; taken together, these four MAGs alone expressed 24 of the 25 families of CAZymes found to be expressed in these consortia.

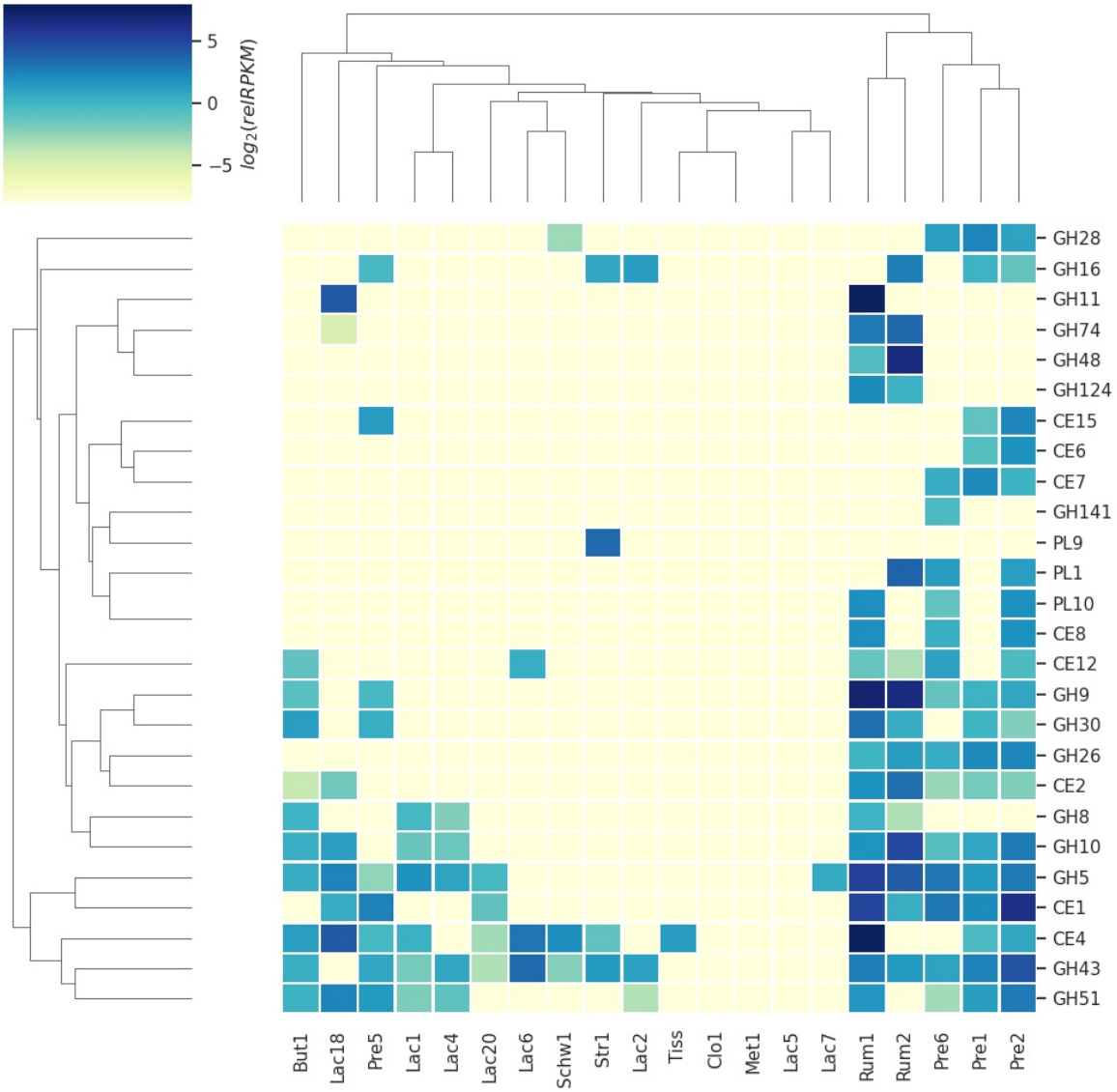


Figure 9. Clustermap of CAZyme expression in key MAGs in generation 9 on reed canary grass. log_2(relRPKM) = log_2[(RPKM for designated ORF)/(median of all expressed RPKMs for that MAG in the sample)]. Cell values are means across all biological samples. Refer to Appendix 2, Supplemental Figure 1A for presence/absence of unexpressed enzymes.

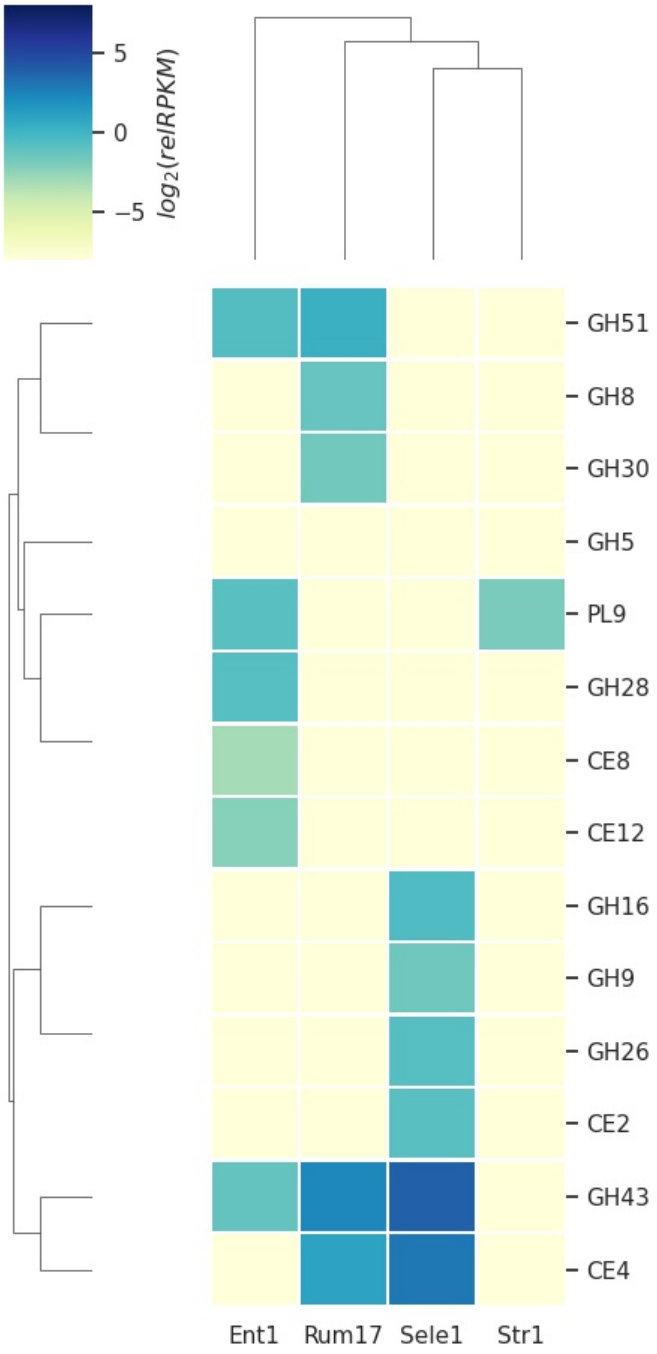


Figure 10. Clustermap of CAZyme expression in key MAGs in generation 9 on xylan. log_2(relRPKM) = log_2[(RPKM for designated ORF)/(median of all expressed RPKMs for that MAG in the sample)]. Cell values are means across all biological samples. Refer to Appendix 2, Supplemental Figure 1A for presence/absence of unexpressed enzymes.

CAZyme expression in xylan consortia key MAGs was dominated by the xylanase/xylosidase GH43 and acetyl xylan esterase CE4, predominantly from *Selenomonas ruminantium*, with additional degradative activities from GH16, its mannanase GH26, and its acetyl xylan esterase CE2 (**Supplemental Figure 1**). Of the four MAGs present, *Enterococcus lactis*, *Acutalibacteraceae UBA4871*, and *S. ruminantium* expressed the entire repertoire of CAZymes detected (**Supplemental Figure 1, Figure 10**). *Streptococcus equinus* expressed only PL9, and this overlapped with *E. lactis*'s expression of the same CAZyme. *E. lactis*, Acutalibacteraceae UBA4871, and *S. ruminantium* exhibited a low degree of redundancy in CAZyme expression; aside from GH43, GH51, CE4, and PL9, each CAZyme in the repertoire expressed by these MAGs was expressed only by one of the three MAGs. Most CAZymes expressed were associated with hemicellulase activity. In addition, GH28 (polygalacturonase), CE8 (pectin methylesterase) and PL9 (pectate lyase) were expressed by *E. lactis* and *St. equinus*.

2.3.3. Cellulolytic labor across substrates is divided into redundant and broad-spectrum strategies

The consortia enriched on alfalfa, bagasse, and reed canary grass shared a common core group of four or five MAGs that employed a broad-spectrum strategy for CAZyme expression in their respective consortia. This core was composed of two *Prevotella ruminicola* MAGs and one *Ruminococcus albus* MAG, with one or two additional MAGs dependent on substrate: either Bacteroidales in the alfalfa consortia, or *Ruminococcus flavefaciens* in the bagasse and reed canary grass consortia. Nearly all of the different CAZyme families expressed in a consortium were expressed by this core set of MAGs;

notably, none of the four core CAZyme producers on each substrate possessed a PL9 (pectate lyase) in their genome; this was typically produced mostly by *Streptococcus equinus*. In addition to the core group of CAZyme producers, each lignocellulosic enrichment consortium also had a group of auxiliary CAZyme producers, composed chiefly of members of the Lachnospiraceae (with two members of the Bacteroidales on alfalfa). The CAZymes expressed by this auxiliary group were generally functionally redundant with CAZymes heavily expressed by the core group.

The two *Prevotella ruminicola* MAGs and *Ruminococcus albus* exhibited similar patterns of complementary CAZyme expression across all three lignocellulosic substrates, with clusters of CAZyme expression corresponding to different enzymatic activities (**Figure 11**). *Ruminococcus albus* fielded CAZymes from families GH8, 11, 48, 74, and 124, corresponding mostly to endo-β-1,4-cellulase and -hemicellulase activity. Pre2 expressed PL1, a pectate lyase. *Ruminococcus albus* and Pre2 expressed PL10, CE8, and CE12, CAZyme families primarily involved in pectin degradation (specifically, families of pectate lyases, pectin methylesterases, and pectin acetylesterases, respectively). Notably, Pre1 did not express these pectinases. The two *P. ruminicola* MAGs expressed GH28, CE6, CE7, and CE15, CAZyme families; GH28 primarily attacks galacturonans (most often found in pectin), CE6 and CE7 attack xylan acetyl esters, and CE15 attacks lignin/xylan bonds. These groups clustered more similarly to each other on bagasse and reed canary grass than between either monocot substrate and alfalfa, probably reflecting similarity in substrate chemical composition. Pre1's absence of PL1, PL10, CE8, and CE12 expression but present GH28 expression, and Pre2's expression of all of these CAZymes, indicates Pre2 may preferentially attack pectin, whereas Pre1 may preferentially attack pectate. Two other

potential clusters emerged, one containing the xylan-degrading GH10 GH43, and GH51, and the other containing GH5, GH9, GH30, CE2, and CE4, a mixed group of cellulases, hemicellulases, and acetyl xylan esterases.

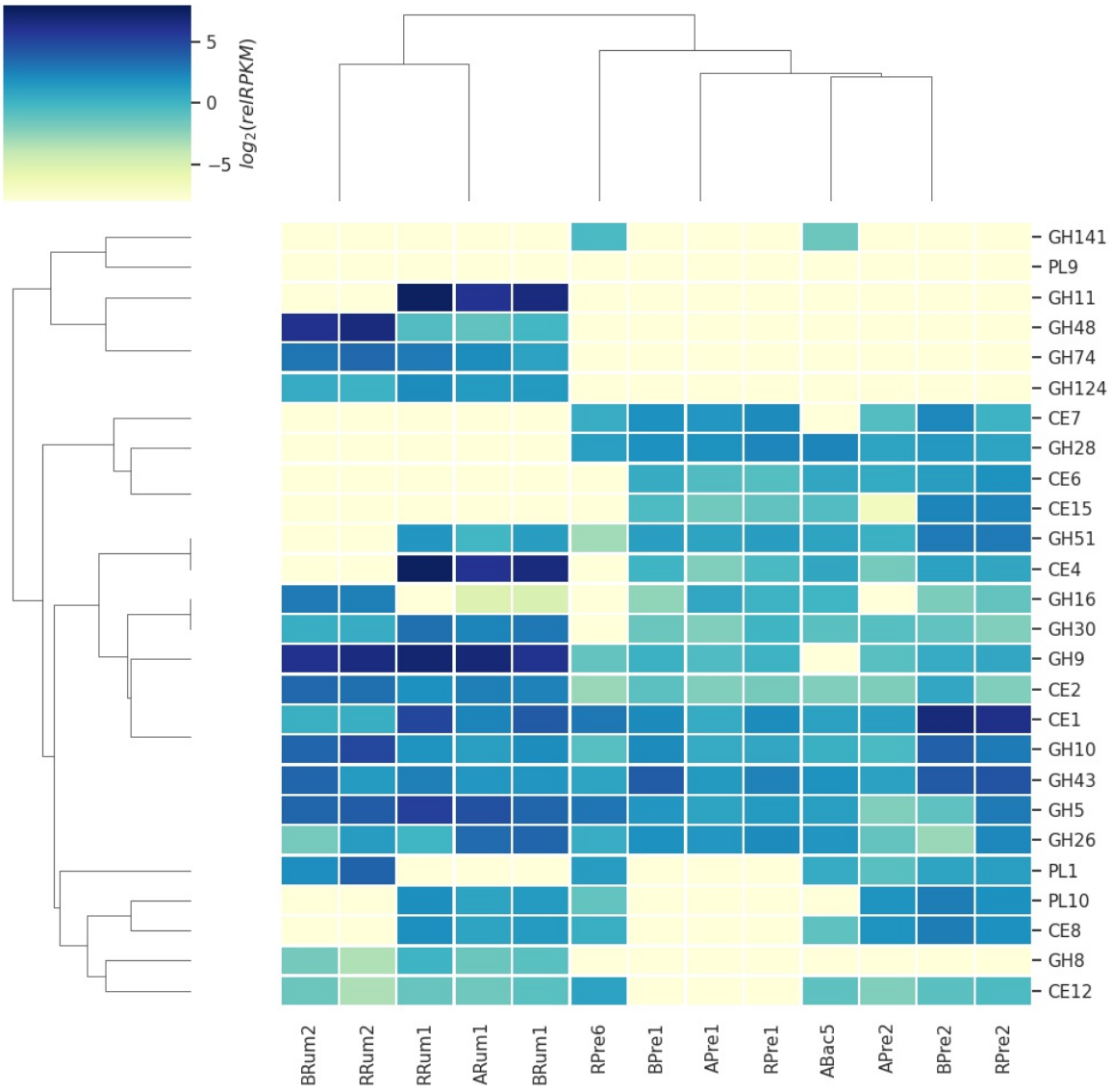


Figure 11. Clustermap of CAZyme expression in "broad-spectrum" transcriber key MAGs in generation 9 across alfalfa, bagasse, and reed canary grass. The first letter of the MAG name denotes the substrate (A=alfalfa, B=bagasse, R=reed canary grass). $\log_2$(relRPKM) = $\log_2$[(RPKM for designated ORF)/(median of all expressed RPKMs for that MAG in the sample)]. Cell values are means across all biological samples. Refer to Appendix 2, Supplemental Figure 1A for presence/absence of unexpressed enzymes.

Bacteroidales W3P20-009, *Ruminococcus flavefaciens*, and an unidentified *Prevotella* sp., the remaining members of the core group, augmented the expression of the other core members in a manner potentially reflecting substrate chemical composition. CAZyme families expressed by Bacteroidales W3P20-009 but not *Ruminococcus flavefaciens*, consisting of GH28, GH51, GH141, CE4, CE6, CE8, and CE15, reflecting alfalfa's pectin content. The CAZyme families expressed by *Ruminococcus flavefaciens* but not Bacteroidales W3P20-009, consisting of the endoglucanase families GH7, GH48, GH74, and GH124 and the mixed cellulase/hemicellulase family GH8, reflecting the increased cellulose content of bagasse and reed canary grass. Of the shared CAZyme families, the greatest differences in relRPKM were seen in two clusters of CAZymes. The primarily-cellulase family GH9, endoglucanase/endogalactanase family GH16, glucanase/xylosidase family GH30, and acetyl xylan esterase family CE2 exhibited higher relRPKMs in *Ruminococcus flavefaciens*. In contrast, the mannanase family GH26, the highly diverse esterase family CE1, and the acetylesterase family CE12 (attacking acetyl groups in pectin, rhamnogalacturonan, and xylan) exhibited higher relRPKMs in Bacteroidales W3P20-009.

2.3.4. Metabolic pathway utilization

To measure primary metabolite production, reducing sugar content in samples from G2 and G6 were measured using a dinitrosalicylic acid (DNS) assay, and short-chain fatty acids and reducing sugars were measured in generations 2, 6, and 9 using high-pressure liquid chromatography (HPLC). (**Supplemental Figure 2, Appendix 2**). Methane, H_2, and CO_2 gases were measured from generations 0, 5, and 10 using gas chromatography (GC) (Peng et al., 2021).

Lignocellulosic consortia utilized mostly arabinose, fructose, fucose, galactose, glucuronate, and glucose, with fewer MAGs fermenting rhamnose and xylose. Most key MAGs expressed genes for fermentation to acetate, and about half expressed genes for fermentation to formate or L-lactate. Most key MAGs expressed alcohol dehydrogenase, but only 4 or 5 key MAGs per substrate expressed acetaldehyde dehydrogenase, the preceding enzyme in the production of ethanol. Most of the key MAGs expressing butyrate kinase also expressed propionyl-CoA:succinate-CoA transferase. Most MAGs expressed a hydrogenase.

At generation 9, alfalfa consortia produced a slight excess of formate, butyrate, and valerate, and larger excesses of acetate and propionate (**Supplemental Figure 2**). At generation 10, alfalfa consortia produced an excess of methane and a slight excess of H_2 (Peng et al., 2021). Alfalfa consortia MAGs clustered into three groups with respect to metabolic activity (**Figures 12-14**). One cluster consisted of Tissierellales PP17-6a, Lachnospiraceae XBB1006, *Schwartzia succinivorans*, and both methanogens. These consumed only a limited set of sugars; Lachnospiraceae XBB1006 consumed fructose, galactose, and glucose, Tissierellales PP17-6a consumed rhamnose, and *S. succinivorans* consumed glucose. Tissierellales PP17-6a and Lachnospiraceae XBB1006 both produced acetate, propionate, butyrate, and hydrogen, and Tissierellales PP17-6a additionally produced ethanol. One cluster of *Prevotella ruminicola*, Bacteroidales (primarily Paludibacteraceae) MAGs, and Oscillospiraceae ER4 acted as primary fermenters not utiliPP17zing xylose or fructose and not generating L-lactate. Bacteroidales W3P20-009 and Paludibacteraceae RF16 sp. expressed xylose isomerase, but did not express xylulokinase, the next enzyme required in the xylose degradation pathway. Rhamnose was solely utilized by Bacteroidales W3P20-009 (making it a critical feeder on rhamnogalacturonan

degradation products) and glucose was solely utilized by the *P. ruminicola* MAGs. Arabinose, fucose, galactose, and glucuronate were degraded by most MAGs, though none of them were degraded by the exact same MAGs. All MAGs in this cluster produced acetate, and the *P. ruminicola* MAGs produced propionate. The remaining MAGs were primary fermenters degrading all assessed monosaccharides and not generating propionate. Of these, *Streptococcus equinus*, the most active overall MAG, primarily fermented arabinose, then galactose, then fructose, then glucose, and produced formate, ethanol, acetate, and L-lactate. All of these MAGs but *Pseudobutyrivibrio ruminis* generated formate, acetate, and H_2. Additionally, all of these MAGs but *Ruminococcus albus* produced L-lactate, and *Butyrivibrio* sp. produced butyrate. *R. albus*, tied for the second most active overall MAG, primarily fermented fucose, then galactose, glucuronic acid, xylulose, and fructose, and produced formate, alcohol, acetate, and hydrogen. Paludibacteraceae UBA1723, tied for the second most active overall MAG, fermented galactose into acetate.

At generation 9, bagasse consortia produced an excess of acetate. By generation 10, bagasse did not produce an excess of methane, but produced a slight excess of H_2 (**Supplemental Figure 2**). Bagasse consortia MAGs clustered into three groups of within-MAG relative metabolic enzyme transcriptional activity (**Figures 15-17**). One cluster consisted of *R. flavefaciens*, Butyricoccaceae sp., Paludibacteraceae UBA1723, and *Schwartzia succinivorans*, as well as the denitrifying bacterium *Denitrobacterium detoxificans* and the sulfate-reducing bacterium *Desulfovibrio desulfuricans*. *R. flavefaciens* degraded both fructose and galactose. Paludibacteraceae UBA1723 degraded only galactose. Butyricoccaceae sp. did not degrade any monosaccharides on this substrate. *D. detoxificans* degraded fructose and *D. desulfuricans* degraded glucose. *R. flavefaciens* generated formate

and L-lactate, *D. desulfuricans* generated ethanol, all but *S. succinivorans* generated acetate, and Butyricoccaceae sp. generated butyrate. In contrast to its demonstrated activity in the alfalfa consortia, *S. succinivorans* did not generate propionate on bagasse. All but Paludibacteraceae UBA1723 generated hydrogen. *S. equinus* clustered by itself; it fermented arabinose, fructose, galactose, and glucose into ethanol, acetate, and L-lactate. A third cluster contained the *Prevotella ruminicola* MAGs, *R. albus*, Lachnospiraceae 1XD8-76 and XBD2001, *Enterocloster clostridioforme*, *Oribacterium* sp., and *Butyrivibrio* spp.. Most MAGs in this cluster degraded most of the assessed monosaccharides; however, rhamnose was degraded by *Enterocloster clostridioforme*, *Oribacterium* sp., and one of the *Butyrivibrio* spp., and xylose was degraded by *Enterocloster clostridioforme* and *Pseudobutyrivibrio* spp.. Most MAGs in this cluster generated formate, lactate, and hydrogen, and all generated acetate. The *P.* ruminicola MAGs generated propionate, and the *P. ruminicola* and *Butyrivibrio* spp. MAGs generated butyrate. Finally, a fourth cluster contained Lachnospiraceae G11 and Oscillospiraceae CAG-110 and ER4. Lachnospiraceae G11 degraded all assessed monosaccharides but fructose, Oscillospiraceae CAG-110 degraded fucose, glucose, xylose, and xylulose, and Oscillospiraceae ER4 degraded arabinose, fructose, fucose, galactose, glucuronate, xylose, and xylulose. Lachnospiraceae G11 generated formate, D-lactate, and butyrate, and Oscillospiraceae CAG-110 generated D-lactate. All members of the cluster generated acetate and hydrogen. No L-lactate or propionate was generated by this cluster.

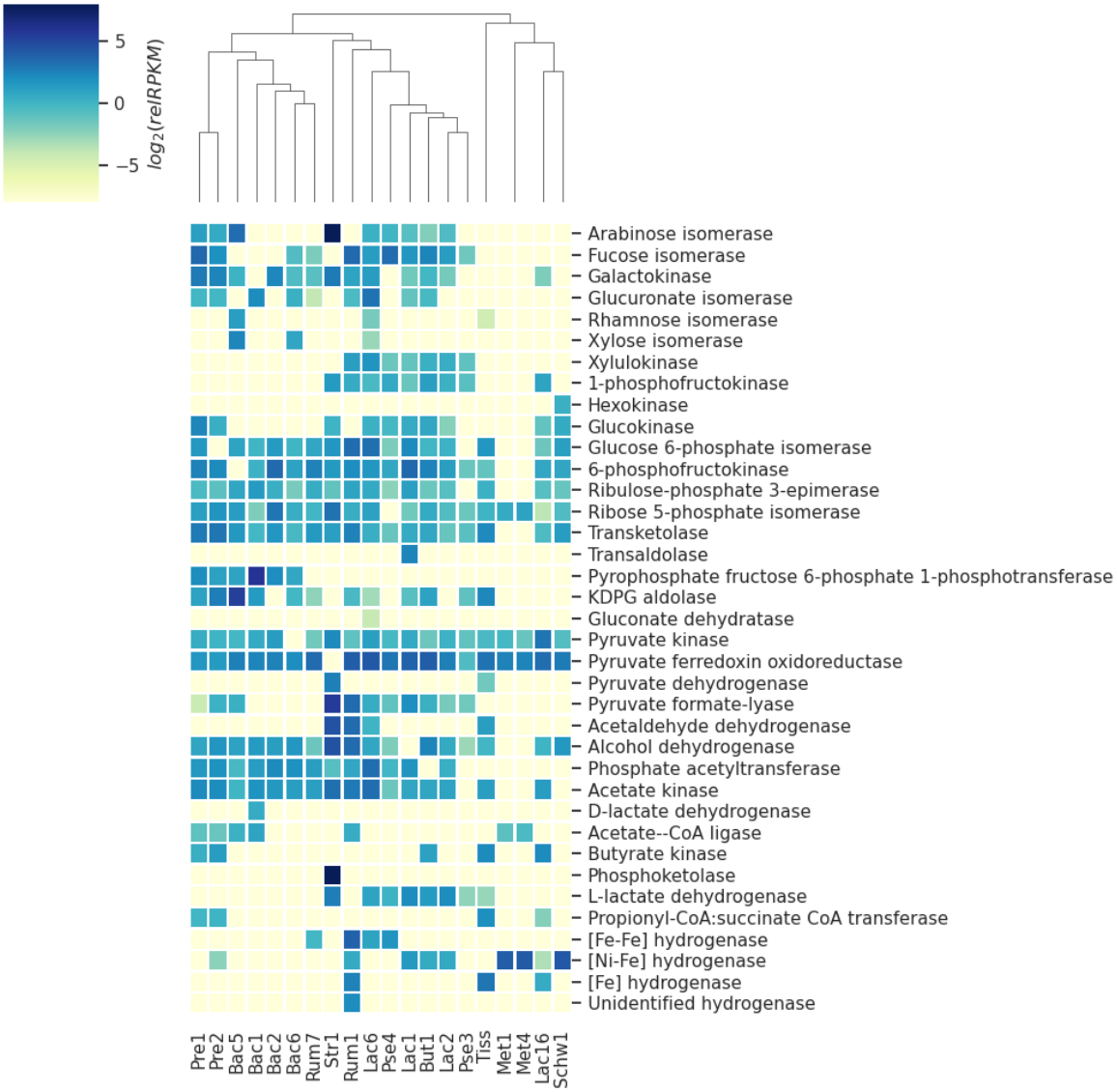


Figure 12. Clustermap of metabolic enzyme expression in key MAGs in generation 9 on alfalfa. $\log_2$(relRPKM) = $\log_2$[(RPKM for designated ORF)/(median of all expressed RPKMs for that MAG in the sample)]. Cell values are means across all biological samples. Refer to Supplemental Figure 1B for presence/absence of unexpressed enzymes.

Figure 13. Metabolic schematic of key MAGs in generation 9 on alfalfa. Made with BioRender.

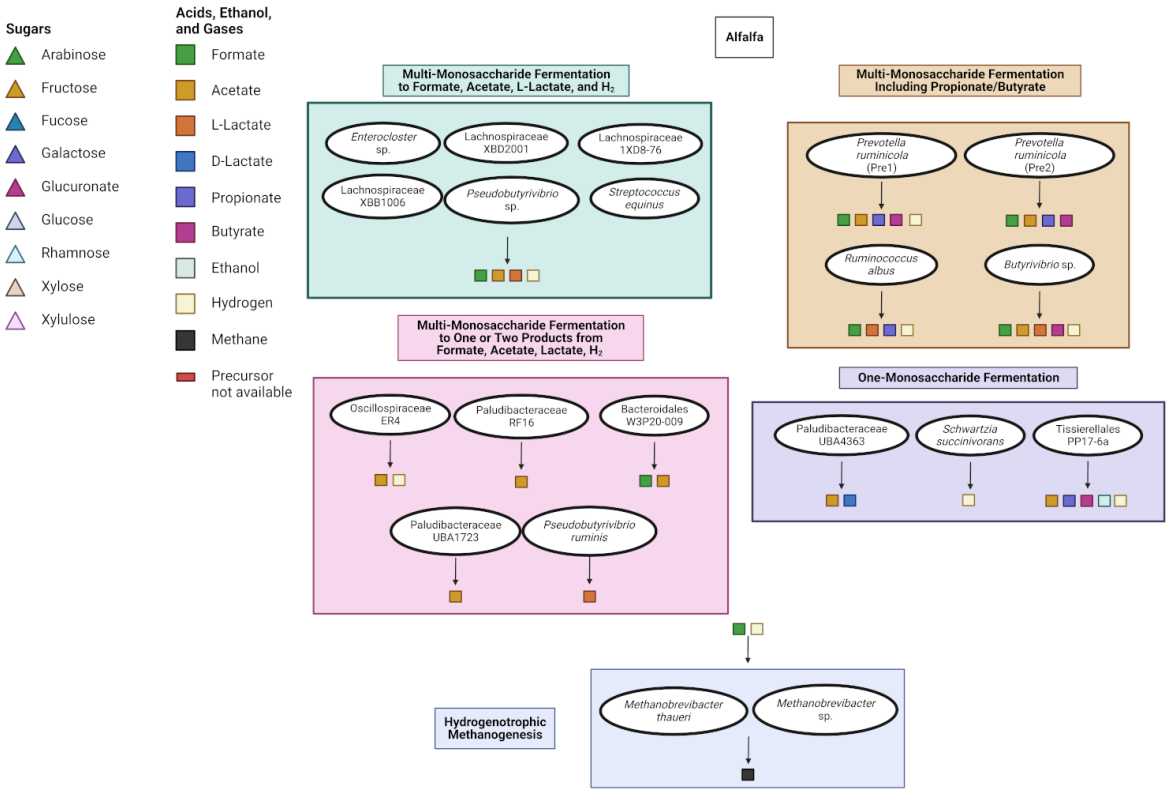


Figure 14. Interaction schematic of key MAGs in generation 9 on alfalfa. Made with BioRender.

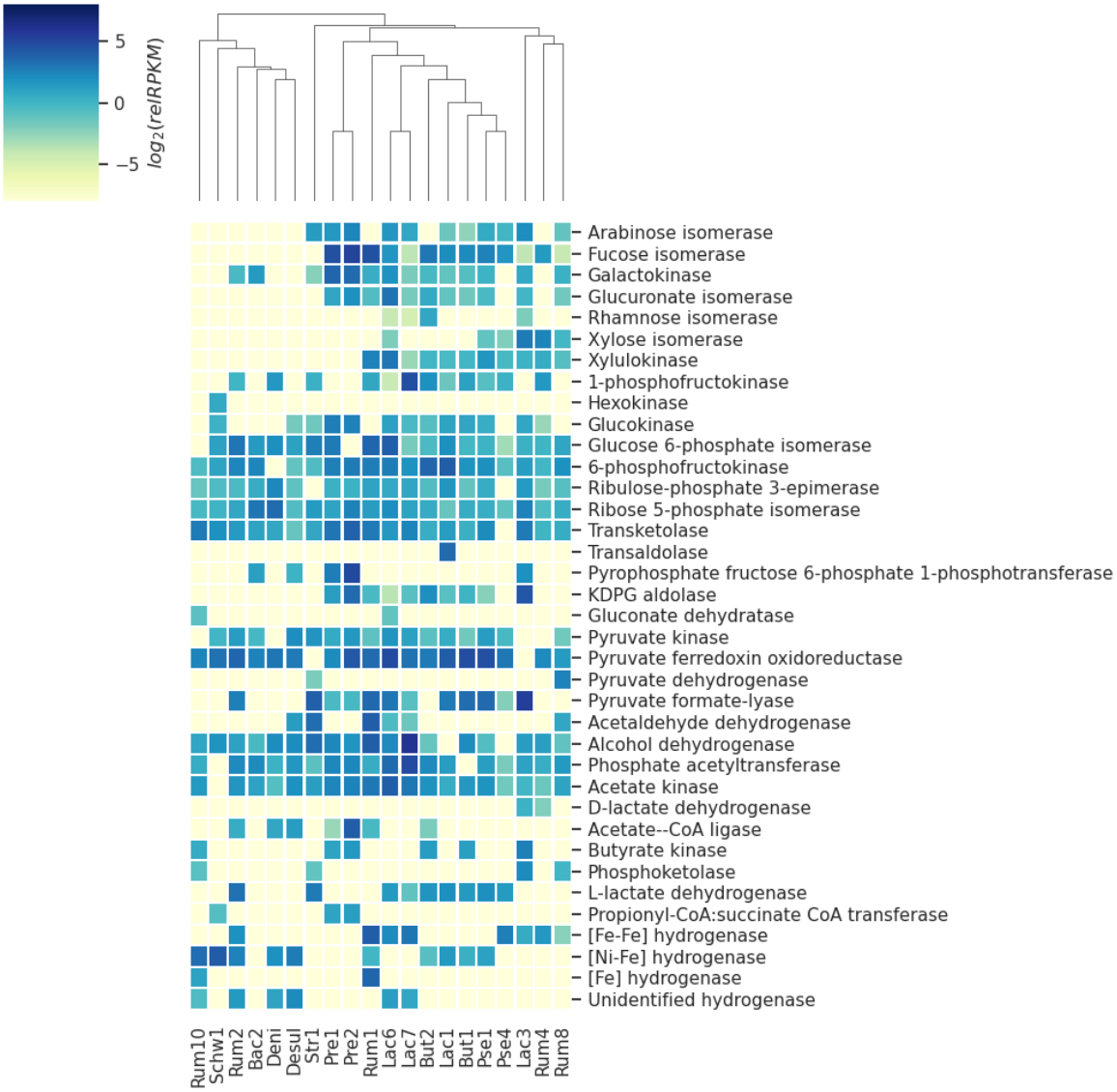


Figure 15. Clustermap of metabolic enzyme expression in key MAGs in generation 9 on bagasse. $\log_2$(relRPKM) = $\log_2$[(RPKM for designated ORF)/(median of all expressed RPKMs for that MAG in the sample)]. Cell values are means across all biological samples. Refer to Appendix 2, Supplemental Figure 1B for presence/absence of unexpressed enzymes.

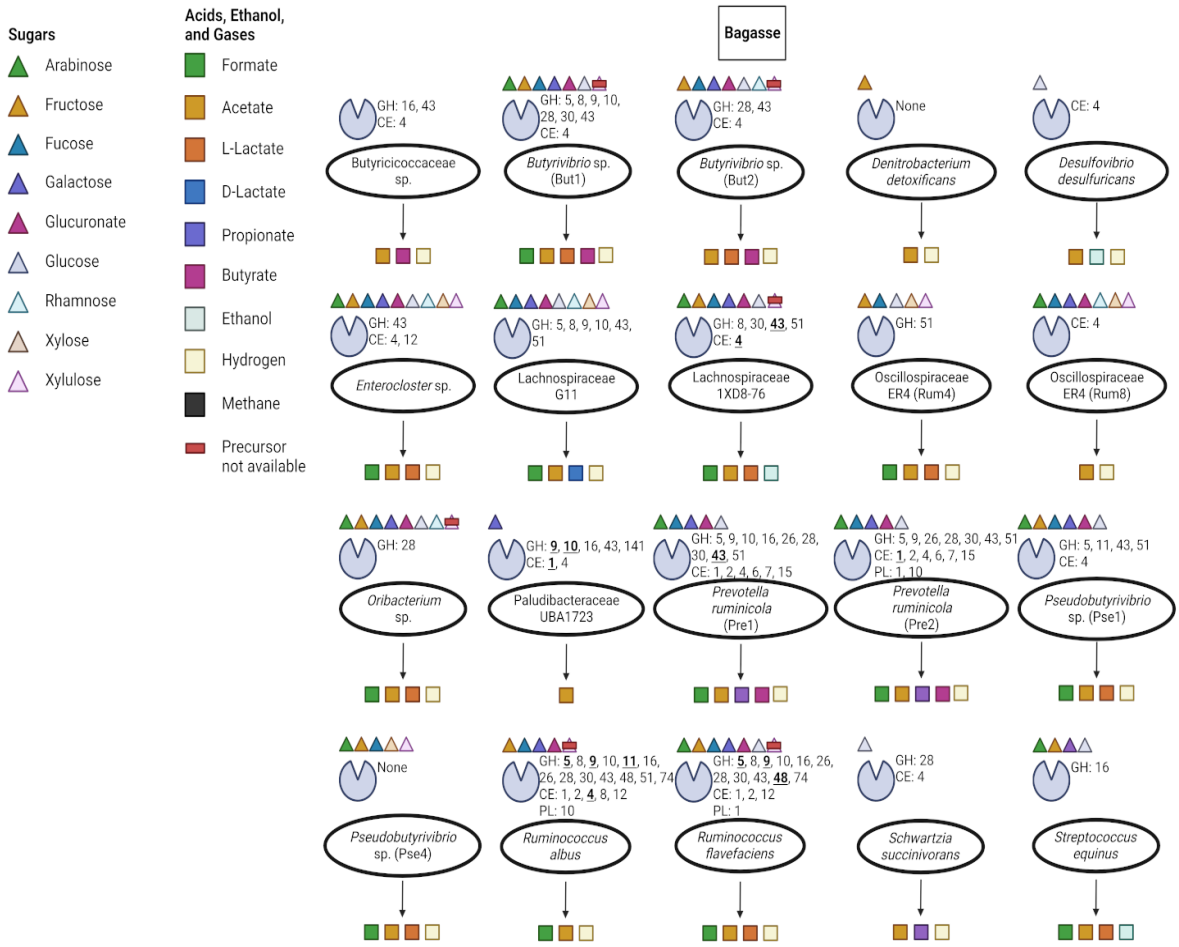


Figure 16. Metabolic schematic of key MAGs in generation 9 on bagasse. Made with BioRender.

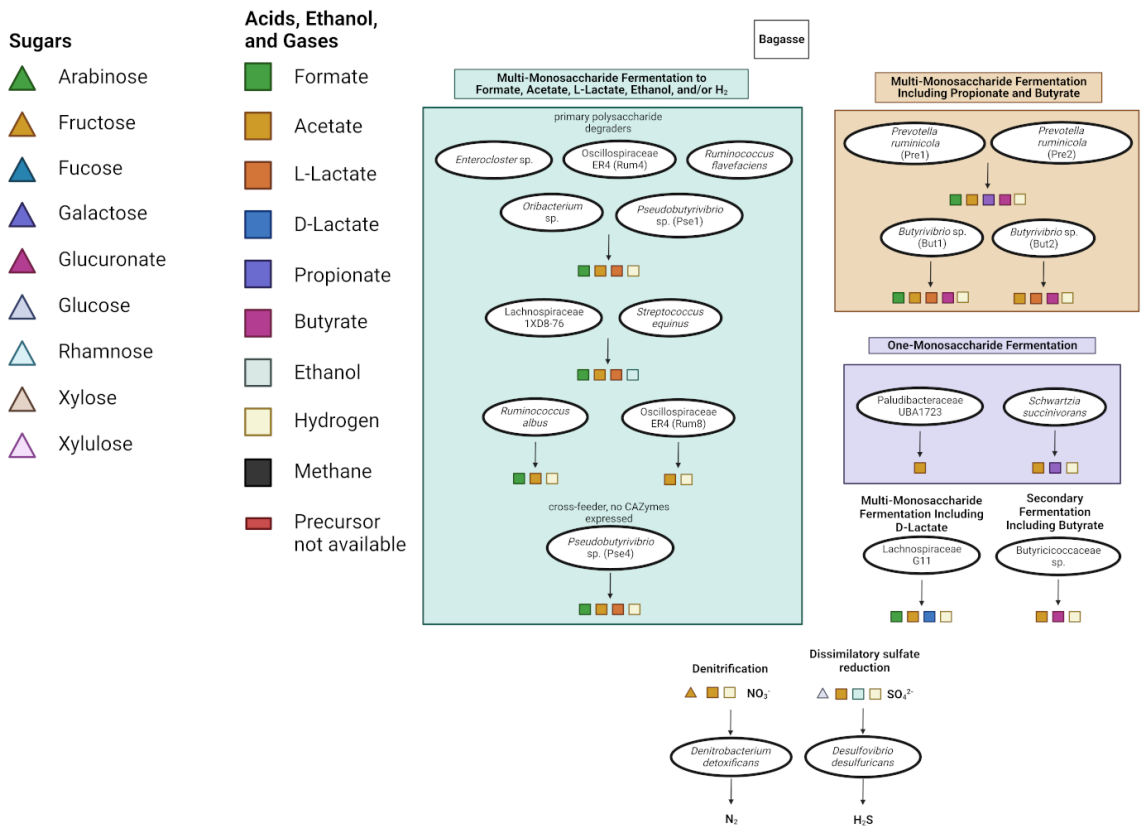


Figure 17. Interaction schematic of key MAGs in generation 9 on bagasse. Made with BioRender.

At generation 9, reed canary grass consortia produced an excess of acetate, propionate, and butyrate. By generation 10, reed canary grass produced an excess of methane and a slight excess of H_2 (**Supplemental Figure 2**). Reed canary grass consortium MAGs clustered into two groups of relative within-MAG metabolic activity (**Figures 18-20**). The first group, containing *Methanobrevibacter thaueri*, *Schwartzia succinivorans*, Anaerovoracaceae RUG099, and Tissierellales PP17-6a, did not degrade any of the assessed monosaccharides in this consortium (except for glucose by *S. succinivorans*), nor did they generate formate or D-lactate. All cluster members but *S. succinivorans* generated acetate, and all cluster members generated hydrogen. Anaerovoracaceae RUG099 generated butyrate, and Tissierellales PP17-6a generated L-lactate, propionate, and butyrate. The other group was composed of the remaining MAGs, including the four most active overall MAGs.

Streptococcus equinus, the most active overall MAG, fermented arabinose, fructose, and glucose into formate, ethanol, acetate, and lactate. *Ruminococcus albus,* the next most active overall MAG, fermented primarily fucose and xylulose, then galactose, glucuronic acid, and fructose, into formate, ethanol, acetate, and hydrogen. Lachnospiraceae 1XD8-76, the next most active overall MAG, fermented arabinose, fucose, galactose, glucuronic acid, xylulose, fructose, and glucose into formate, acetate, lactate, and hydrogen. *Acetatifactor* sp., the next most active overall MAG, fermented arabinose, fucose, galactose, glucuronic acid, xylose, xylulose, and glucose into formate, acetate, lactate, butyrate, and hydrogen. In general, this cluster followed the same overarching pattern seen in clusters in the other complex consortia where at least 50% of cluster members degraded most assessed monosaccharides, with the exception of rhamnose and xylose, which were degraded by a small set of one to four MAGs, and fermented into formate, acetate, and hydrogen by most cluster members, as well as into L-lactate by a significant portion. In this cluster, propionate was generated by *Prevotella* spp. (including both *P. ruminicola* MAGs) and butyrate was generated by *Prevotella* spp., Lachnospiraceae UBA2868 and CAG-194, and *Acetitomaculum* sp..

Of xylan consortium key MAGs, *Selenomonas ruminantium*'s metabolic repertoire was the most transcriptionally activated relative to the median RPKM of all of its ORFs (**Figures 21 and 22**). *S. ruminantium* expressed enzymes involved in the utilization of arabinose, fucose, xylose, and unidentified hexoses. Formate and lactate were manufactured by all key MAGs. *Enterococcus lactis* and Acutalibacteraceae UBA4871 manufactured acetate. *S. ruminantium* manufactured propionate via the succinate pathway. *E. lactis* and *A*cutalibacteraceae UBA4871 played secondary roles in sugar utilization, fermentation, and fatty acid metabolism. *E. lactis* and Acutalibacteraceae UBA4871 primarily deployed sugar

utilization enzymes oriented toward metabolizing a broader variety of carbohydrates, or carbohydrates not found in xylan. *Streptococcus equinus* utilized only fructose. *S. ruminantium* was the primary producer of hydrogen gas, with a small contribution from Acutalibacteraceae UBA4871. At generation 9, xylan consortia produced an excess of formate, acetate, lactate, propionate, butyrate, and reducing sugar. By generation 10, xylan did not produce an excess of methane, but produced a large excess of H_2.

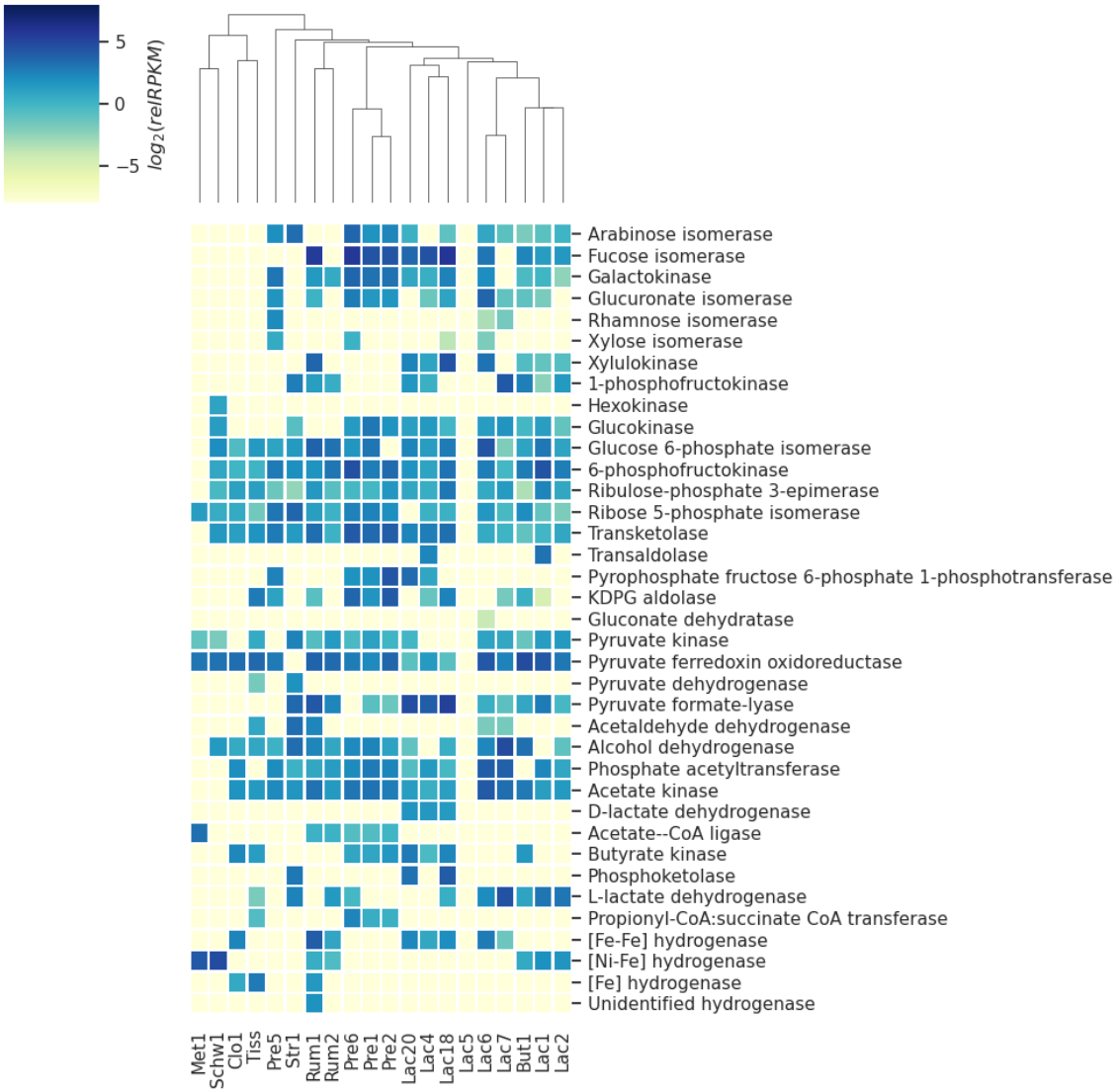


Figure 18. Clustermap of metabolic enzyme expression in key MAGs in generation 9 on reed canary grass. log_2(relRPKM) = log_2[(RPKM for designated ORF)/(median of all expressed RPKMs for that MAG in the sample)]. Cell values are means across all

biological samples. Refer to Appendix 2, Supplemental Figure 1B for presence/absence of unexpressed enzymes.

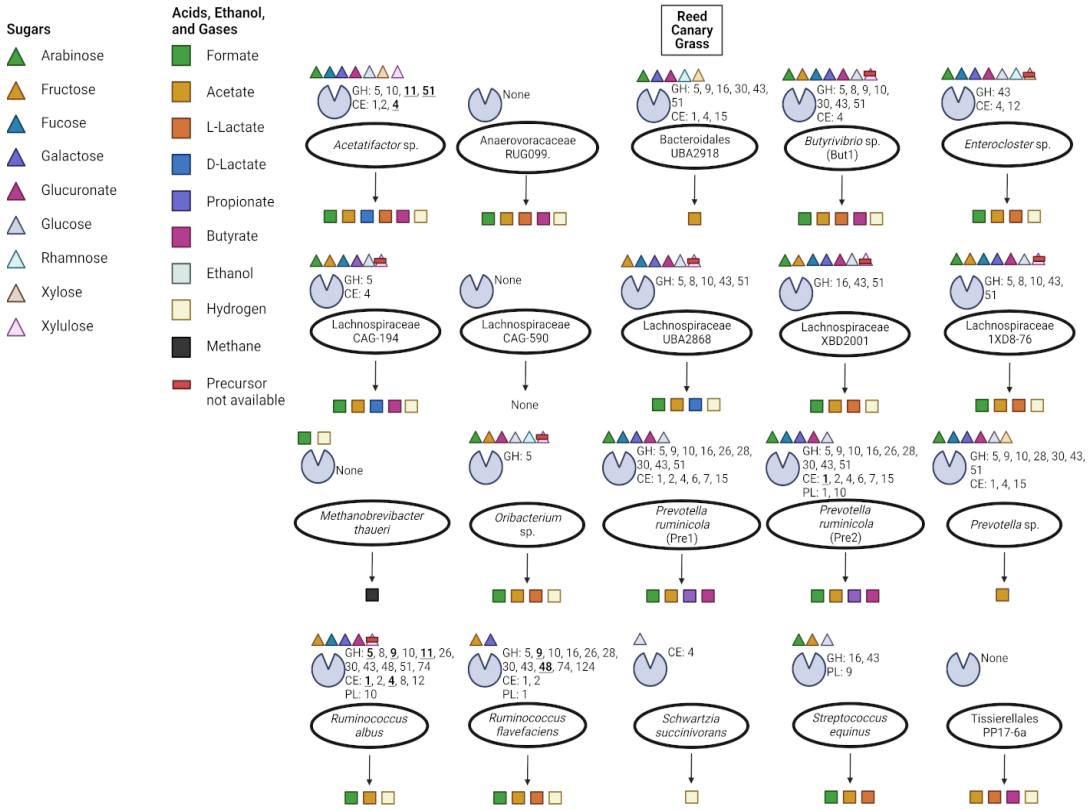


Figure 19. Metabolic schematic of key MAGs in generation 9 on reed canary grass. Made with BioRender.

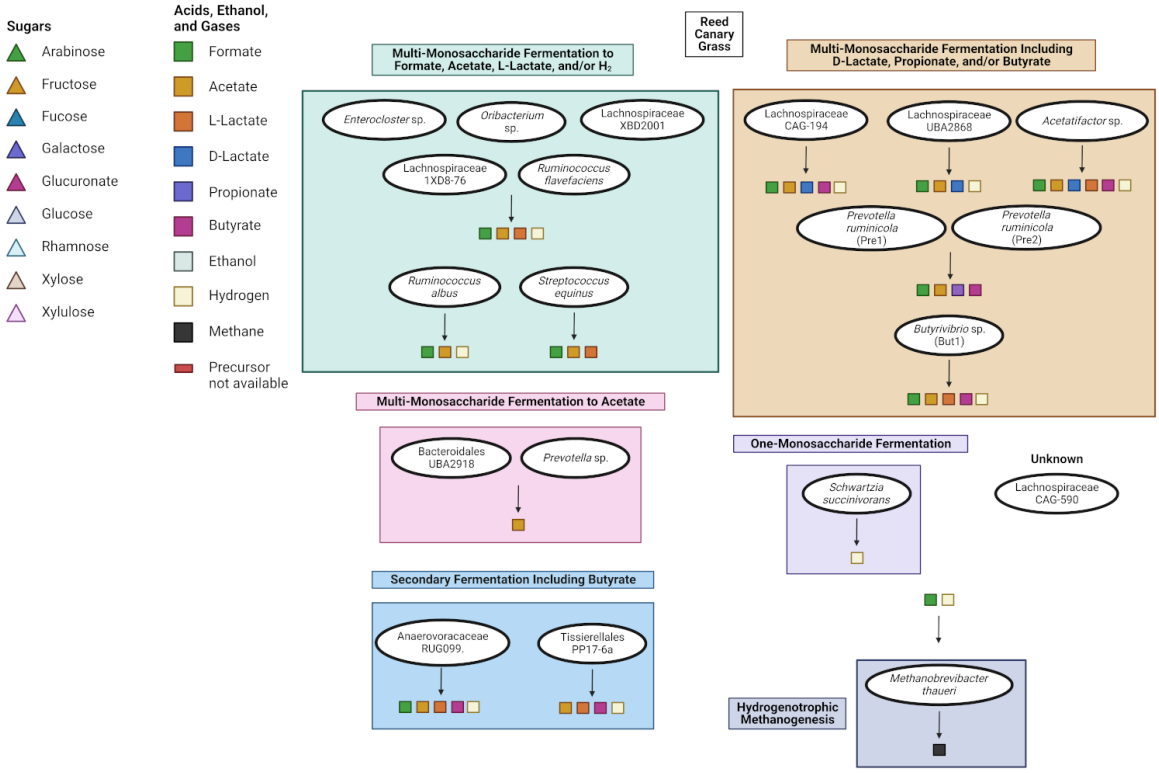


Figure 20. Interaction schematic of key MAGs in generation 9 on reed canary grass. Made with BioRender.

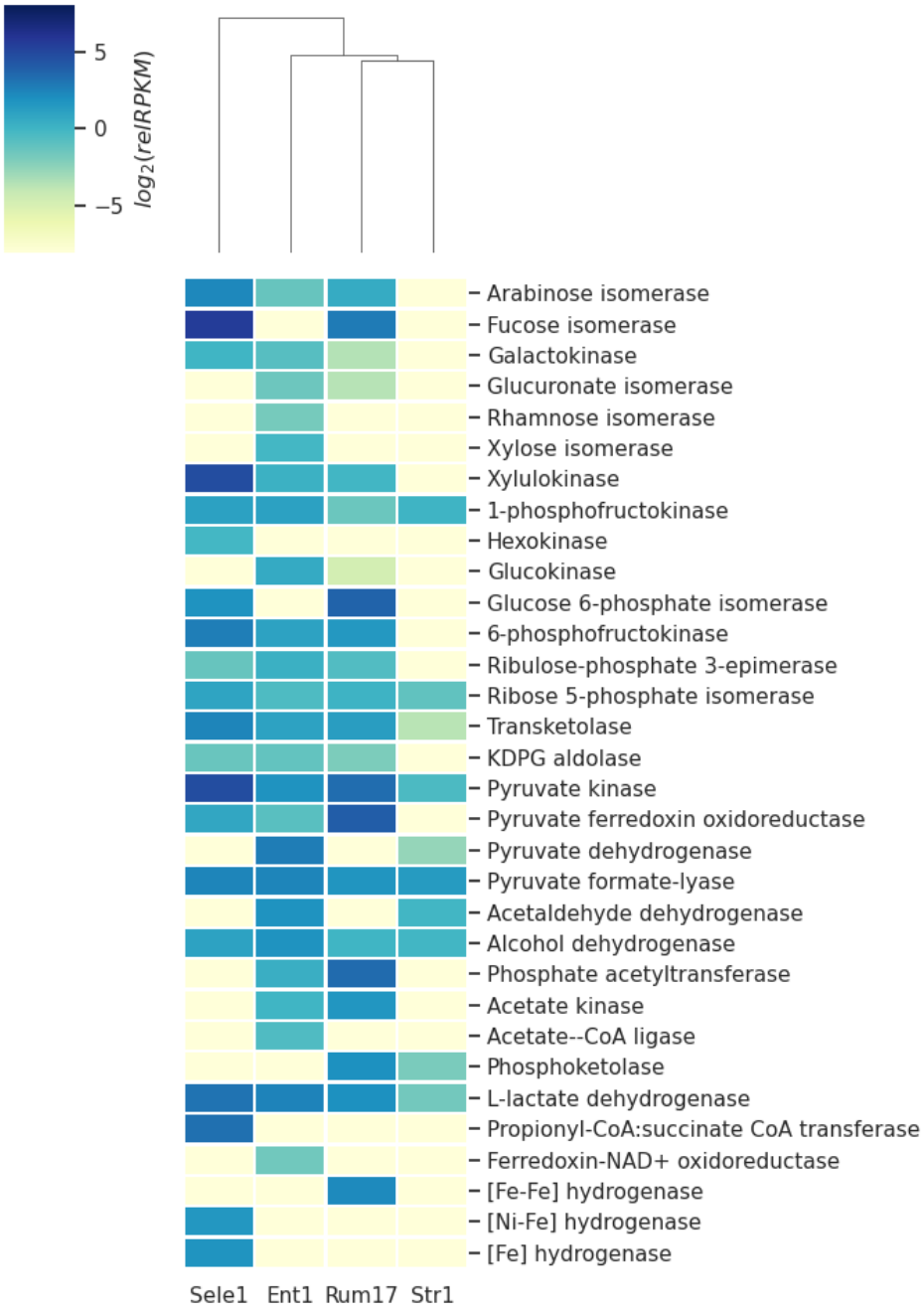


Figure 21. Clustermap of metabolic enzyme expression in key MAGs in generation 9 on xylan. log_2(relRPKM) = log_2[(RPKM for designated ORF)/(median of all expressed RPKMs for that MAG in the sample)]. Cell values are means across all biological samples. Refer to Appendix 2, Supplemental Figure 1B for presence/absence of unexpressed enzymes.

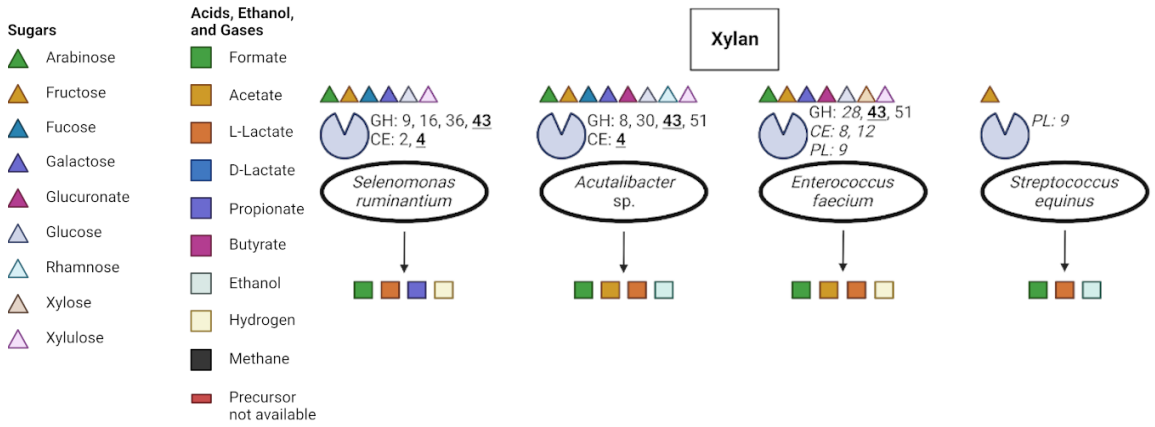


Figure 22. Metabolic schematic of key MAGs in generation 9 on xylan. Made with BioRender.

2.3.5. Methanogenesis pathways

Alfalfa consortium methanogen activity was dominated by the hydrogenotroph *Methanobrevibacter thaueri*, but another *Methanobrevibacter sp.*, as well as three obligate methylotrophs, *Methanosphaera stadtmanae* and two undescribed members of the Thermoplasmata, also contributed (**Figure 20**). Alfalfa culture methanogenesis occurred entirely through hydrogenotrophy, since no expression of the *mcr* gene was observed in methylotrophs. Methane in reed canary grass consortia was generated by *M. thaueri* with no contributions from the other *Methanobrevibacter* sp. or the two novel Thermoplasmata present.

In bagasse consortia, the methylotrophic gene expression of *Methanosphaera stadtmanae* was elevated relative to its activity in alfalfa consortia. *M. stadtmanae* was the only methanogen MAG remaining after the early demise of hydrogenotrophic methanogen MAGs, which took place before generation 5. Notably, the sulfate-reducing bacterium *Desulfovibrio desulfuricans* was enriched in these consortia and not in alfalfa or reed canary grass consortia, and exhibited one of the highest relative contributions to bagasse consortium

transcriptional activity. *D. desulfuricans* and other sulfate-reducing bacteria outcompete hydrogenotrophic methanogens for hydrogen, which it uses for dissimilatory sulfate reduction, a process more thermodynamically favorable than hydrogenotrophic methanogenesis under anoxic conditions (Thauer et al., 2007). *D. desulfuricans* does not compete with methylotrophs for substrate. Bagasse typically contains more lignin, about 25% of dry mass, than either alfalfa (6-9% of dry mass) or reed canary grass (Su et al., 2015;). The addition of lignin enhances the growth of *Desulfovibrio spp.* in laboratory culture, and some *Desulfovibrio spp.* have exhibited the ability to degrade some of the aromatic compounds liberated by lignin degradation (Ziomek & Williams, 1989). It is likely that methanogenesis inhibition in bagasse consortia may be driven by the combined effects of enhanced growth of *D. desulfuricans* on the relatively lignin-rich substrate and preferential consumption of hydrogen by sulfate reduction over methanogenesis.

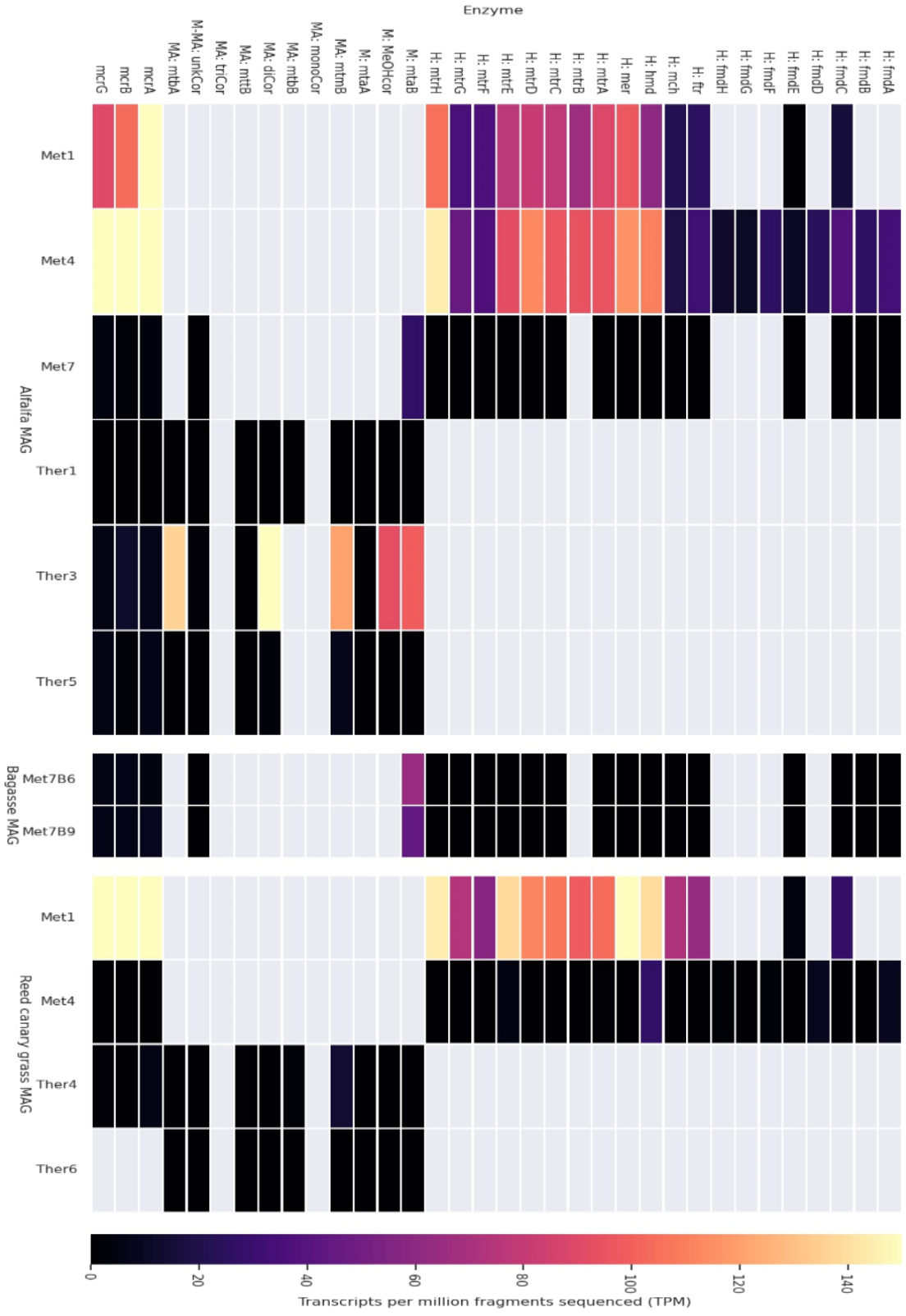


Figure 23. Methanogenesis pathway enzyme expression in key MAGs across the carbohydrate substrates alfalfa, bagasse, and reed canary grass. Cell values are means across all biological samples. Grey = enzyme absent. "H:" represents enzymes in the

hydrogenotrophic pathway. "M:" represents enzymes in the methylotrophic pathway from methanol. "MA:" represents enzymes in the methylotrophic pathway from methylamines. mcrA-G are common to all pathways. Enzyme abbreviations: fmdA-H: formyl-methanofuran dehydrogenase subunits A-H; ftr: formyl-tetrahydromethanopterin formyltransferase; mch: methenyl-tetrahydromethanopterin cyclohydrolase; hmd: H_2-forming methylene-tetrahydromethanopterin dehydrogenase; mer: methylene-tetrahydromethanopterin reductase; mtrA-H: methyl-tetrahydromethanopterin subunits A-H; mtaAB: methanol-cobalamin methyltransferase, subunits A and B; MeOHcor: methanol corrinoid protein; mtmB: monomethylamine methyltransferase subunit B; monoCor: monomethylamine corrinoid protein; mtbB: dimethylamine methyltransferase; diCor: dimethylamine corrinoid protein; mttAB: trimethylamine methyltransferase; triCor: trimethylamine corrinoid protein; unkCor: unknown corrinoid protein; mcrABG: methyl-coenzyme M reductase.

2.4. Discussion

This study builds on a previous investigation of parallel enrichments from the goat gut microbiome that sought to separate this microbiome into specific taxa and analyze the impact of substrate on the composition of communities dominated by bacteria and fungi, and in which we reconstructed the metabolic networks of over 700 high-quality MAGs constituting most of the microbiome. In this follow-up analysis, we identified the most transcriptionally active members of antibiotic-free consortia enriched on four different carbon substrates, then analyzed CAZyme and metabolic enzyme expression during the second-to-last passage to understand the extent of the impact of substrate on consortium metabolic activity.

2.4.1. CAZyme expression strategies are inherent to a goat fecal source microbiome across substrates

This study is, to our knowledge, the first study examining the division of polysaccharide-degrading labor across a set of complex carbohydrate substrates of biotechnological

importance. Most experimental studies of labor division in polysaccharide degradation have employed simple substrates in synthetic consortia engineered in a bottom-up fashion, via the combination of isolates and/or engineered strains, and division of labor in lignocellulosic consortia, as with other polysaccharide-degrading consortia, may occur through four different mechanisms: complementary differences in gene content, alternate regulation of polysaccharide degradation genes, differences in CAZyme function, and specialization in transport and consumption of liberated carbohydrates (Lindemann, 2020).

UPGMA clustering of CAZyme expression data revealed two functional clusters of key MAGs involved in CAZyme expression in lignocellulosic consortia. One was composed of four or five MAGs which expressed nearly all of the diversity of CAZymes expressed in the consortium, with the exception of the pectate lyase PL9 in all consortia and the hemicellulase GH141 in bagasse consortia. This cluster had three members in common across all lignocellulosic consortia: two *Prevotella* ruminicola MAGs and *Ruminococcus albus.* One or two other MAGs depending on substrate (Bacteroidales W3P20-009 in alfalfa, *Ruminococcus flavefaciens* in bagasse, *R. flavefaciens* and *Prevotella* sp. in reed canary grass) completed this set. The other cluster was composed of the remaining key MAGs and expressed a smaller set of CAZymes including the most common CAZyme families in the prokaryotic portion of the source microbiome. Across all lignocellulosic substrates, this set included the following families: the cellulase family GH9, the hemicellulase families GH10, GH16, and GH43, the combined cellulase/hemicellulase family GH5, the pectinase family GH28, and the esterase family CE4. These families clustered together across all substrates with the cellulase/hemicellulase family GH51 and the esterase family CE1. The way in which they clustered differed by substrate. Substrate type was not consistently associated

with the degree of expression of a given CAZyme family relative to the median RPKM of expressed ORFs in each MAG, as exemplified by differences in expression of CAZyme families across substrates in the broad-spectrum strategy MAGs; this was expected, likely due to the different degradative demands associated with each substrate and the different gene regulatory architecture in each MAG. Taken together, this evidence indicates that while the degree to which a given set of CAZymes is expressed may be affected by substrate, underlying subdivisions of labor in lignocellulolytic source microbiota are likely to be preserved regardless of substrate, and the members of these groups are likely to be relatively consistent. In addition, the observed two-part strategy of both a small group of broad-spectrum CAZyme producers and a much larger group of more focused, highly functionally redundant producers is likely inherent to this source microbiome. It is unclear how generalizable this strategy is to other microbiota, but this strategy is an appropriate one for attacking a set of disaccharide bonds that are likely to be present across a wide variety of lignocellulosic substrates while retaining the capacity to attack rarer ones that may vary from substrate to substrate: the large population of functionally redundant producers mobilize their CAZymes against the former, whereas the smaller subset of producers with a particularly broad variety of expressed CAZymes act against the latter. Viewed through the lens of Lindemann's 2020 review, these consortia accomplish division of labor through the first and third of the four methods put forward, and a follow-up study of these consortia may reveal they employ the other two as well. Goats are natural browsers and consume highly variable diets, so this strategy is highly appropriate for a goat gut microbiome (National Research Council, 2006).

2.4.2. Cross-substrate metabolic commonalities suggest limits to the capacity of substrate to shape the metabolism of a source microbiome under uniform enrichment conditions

The results we report here generally recapitulate three governing principles of microbial consortium enrichment and function on a given substrate: the substrate provided to a microbiome shapes its composition and activity, that a substrate selects for microbes with a metabolic repertoire suited to survive on it, and that more chemically complex substrates enrich more phylogenetically diverse and metabolically complex communities of microbes. Substrate choice shaped the metabolic output of these consortia. According to our HPLC analysis, the small community of specialists enriched on xylan, dominated by *Selenomonas ruminantium*, fermented xylan constituents into primarily acetate, lactate, propionate, and hydrogen, with some excess formate and reducing sugar. Bagasse was largely unproductive, except for a slight excess of acetate and an excess of hydrogen comparable to that of alfalfa without an accompanying excess of methane, likely influenced by the collapse of the hydrogenotrophic methanogen population. Alfalfa produced mostly acetate, with some formate and propionate. Reed canary grass produced acetate, propionate, and butyrate. Substrate chemical complexity does not appear to have an impact on the degree of diversity of metabolites generated in prokaryotic consortia, particularly short-chain fatty acids. Neither does it seem to necessarily restrict by much, in a global sense, the transcriptional array that is mobilized to produce them – a few secreted CAZymes in the lignocellulosic consortia were not present in the xylan consortia. Despite the fact that this substrate selected for a small specialist community many times smaller than those in the lignocellulosic consortia, most key enzymes of all evaluated sugar utilization and fermentation pathways

that were expressed in the lignocellulosic consortia were expressed by all xylan key MAGs as a group.

Taken together, especially in light of the HPLC results, this evidence fits a hypothesis that the collective transcriptional networks of a lignocellulolytic consortium, particularly mechanisms of catabolite repression, are the central factor determining that consortium's response to substrate in conjunction with their network of metabolic pathways, directly linking these pathways to the external environment. Most experiments of transcriptional regulation in substrate uptake have taken place in small synthetic consortia with defined substrates; in these consortia, interactions are generally engineered into them at the transcriptional network level (Deter & Lu, 2022; see F. Li et al., 2019 as an example).

Genome-scale metabolic models (GMMs) have been successfully used to predict the metabolic responses of and drive the engineering of prokaryotes and eukaryotes, including those that play important roles in biomass degradation (Blazeck & Alper, 2010; Simeonidis & Price, 2015; Aung et al., 2013; Wilken et al., 2021). Whole consortia may be modeled with multiple GMMs to study microbial interactions, and this method is readily scalable to study the exchange of multiple metabolites between different species (Basile et al., 2020; Garza et al., 2018). Integration of other networks, such as transcriptional regulatory networks (TRNs) and signal transduction networks (STNs), with GMMs into a unified multiscale model has been accomplished for individual organisms (e.g. Filho & Paula, 2021), but not for consortia of organisms, to our knowledge. These tools may be useful in the future to successfully predict the impact of substrate on consortium enrichment.

Lignocellulolytic microbiota enriched from herbivore digestive tracts are excellent sources of consortia, and consortium members, to use to degrade lignocellulose into

valuable products such as methane, hydrogen, and volatile fatty acids. Our results suggest that the metabolic activity of a prokaryotic consortium enriched on a variety of lignocellulolytic substrates will be shaped by the substrate, but that broad patterns of lignocellulolytic labor distribution among consortium members are preserved across substrates. This analysis of the effects of substrate on the carbohydrate-degrading activity of gut-derived consortia provides insight into how lignocellulolytic consortia divide labor during the degradation of complex substrates in the environment of the herbivore gut, and presents for the first time one strategy by which carbohydrate-degrading labor is distributed in an herbivore gut consortium to address the problem of degrading multiple substrates. In doing so, it expands our knowledge about the role of carbon substrates in the design of lignocellulolytic synthetic consortia for efficiently degrading waste biomass and/or converting it into platform chemicals and other valuable metabolites, as well as their role in designing interventions to modulate the activity of existing consortia in the gut and in the industrial bioreactor, and reveals a new avenue along which to effectively divide degradative labor in these engineered consortia.

2.5. Additional supplemental data availability

Additional supplemental data for this manuscript is available on the eScholarship platform entry for this dissertation.

Supplemental Data 1. Alignment statistics for all replicates in antibiotic-free alfalfa, bagasse, reed canary grass, and xylan consortia at generation 9.

Supplemental Data 2. Functional annotations and gene expression values associated with all samples. This supplement provides functional annotations of all ORFs for the 719+G1

MAG dataset analyzed in this study. Gene expression results are included as reads per kilobase million (RPKM) values.

Supplemental Data 3. Functional and pathway tables for all key MAGs. This supplement includes results from annotation of key MAGs with the CAZy database.

Supplemental Data 4. CAZyme expression results. This supplement includes results from annotation of key MAGs with the CAZy database.

Supplemental Data 5. Metabolic enzyme expression results.

3. Enrichment of fecal samples of captive *Colobus guereza* and *Gorilla gorilla gorilla* reveals a diverse community of lignocellulolytic prokaryotes and fungi with potential key roles in herbivory and health

3.1. Introduction

The mammalian gut microbiome – a complex community of bacteria, archaea, fungi, protists, and viruses – plays a central role in the life of its host, with systemic effects extending far beyond the digestive system (McFall-Ngai et al., 2013). For almost two decades, the human gut microbiome has been and continues to be a subject of extensive and well-funded research, from which we have gained a basic understanding of its role in nutrition, and in overall health and disease (Clavel et al., 2022). The gut microbiomes of non-human primates (NHP) relatives are less well-studied, but studies of gut microbiota in our closest relatives are critical to understanding the emergence of the diverse dietary strategies and gut physiologies of primates, including our own, and the evolutionary forces that shape them (Amato, 2019). The study of gut microbiota is also of critical importance to NHP conservation; detectable differences in microbiome composition and activity have been associated with food quality and abundance, social organization, season, age, sex, human incursion on habitat, and life in captivity (Amato et al., 2013; Barelli et al., 2020; Clayton et al., 2016; Greene et al., 2018; Perofsky et al., 2017; B. Sun et al., 2018).

Due to the absence of cellulolytic and hemicellulolytic enzymes in vertebrate genomes (in fact, in all but a few phyla within the Metazoa), herbivorous NHPs are completely dependent on their microbiota to convert ingested plant biomass into absorbable energy compounds, which occurs via fermentation by bacteria and fungi expressing unique

repertoires of carbohydrate-active enzymes (CAZymes) (Watanabe & Tokuda, 2001). The two species highlighted in this study, Western lowland gorillas (*Gorilla gorilla gorilla*) and members of the subfamily Colobinae ("colobines"), represent two highly contrasting evolutionary approaches to the degradation of lignocellulose-heavy diets among NHPs. Uniquely among the primates, colobines carry out fermentation in a three- or four-chambered foregut, and consume leafy diets that may contain as much as 52% acid detergent fiber in the wild, sometimes subsisting exclusively on leaves during parts of the year (Gruninger et al., 2014; Matsuda et al., 2019; Lambert, 2011; Matsuda et al., 2017; Nijboer & Clauss, 2006). Gorillas, by contrast, are hindgut fermenters like all primates outside the subfamily Colobinae (Wolfensohn, 2004). They ferment plant biomass from a variety of sources, including grasses, figs, bark, and pith, in their colon, with as much as 57.3% of energy from short-chain fatty acids (SCFAs) deriving from fermentation of fiber (Doran-Sheehy et al., 2009; Popovich et al., 1997).

The gut mycobiome of NHPs is poorly explored but key to host digestion and health, and serves as an especially critical component of the microbiome in herbivores both within and outside the primates. Though the population of fungi in the gut by total number of organisms is orders of magnitude smaller than that of prokaryotes (*e.g.* in humans: Qin et al., 2010 and Underhill & Iliev, 2014; in ruminants: Druzhinina & Kubicek, 2012, Orpin & Ho, 1991, and Trinci et al., 1994), studies in humans and mice have associated the composition of this portion of the gut microbiome with the regulation of a variety of health states such as host immune responses and chronic gastrointestinal disease etiology(Limon et al., 2019; van Tilburg Bernardes et al., 2020; Wheeler et al., 2016; Xinyun Qiu et al., 2017). A variety of studies in NHPs that explore the mycobiome describe a tendency for fungi of the

Ascomycota and Basidiomycota to be the most abundant of all phyla, with significant representation from yeasts (Barelli et al., 2020; Borruso et al., 2021; James et al., 2022; Mann et al., 2019; Schulz et al., 2018; B. Sun et al., 2021). Obligate anaerobic fungi of the early-diverging fungal phylum Neocallimastigomycota, in particular, have been identified as active members of the gut mycobiome of the strongly herbivorous gorilla, and are likely to be crucial members of the microbiota of other herbivorous NHPs with high-fiber diets, including colobines. Anaerobic fungi are recognized as key participants in plant biomass degradation in the digestive tracts of many herbivores, and live in the gut of a phylogenetically diverse set of herbivorous mammals (Gruninger et al., 2014; Liggenstoffer et al., 2010). 12 separate clones of ITS1 sequences from anaerobic fungi were identified in 2018 from fecal samples of *G. gorilla gorilla* (Schulz et al., 2018). A 2022 preprint explored the gut metatranscriptome of zoo-housed gorillas, and found that anaerobic fungi were not only present, but disproportionately transcriptionally active, compared to other fungi in the gorilla gut microbiome (Houtkamp et al., 2022). Given that anaerobic fungi appear to be present in all clades of mammalian foregut-fermenting herbivores where previous investigations have sought them out (Gruninger et al., 2014), anaerobic fungi are likely to be present in the foregut of members of the Colobinae. Amato, Clayton, and Hale demonstrated, in a study of 16S sequences from fecal samples of 19 species of wild colobines, that the taxonomic composition of prokaryotic communities of colobine fecal microbiota converges strongly with that of ruminants, whose gut mycobiota are mostly or exclusively composed of anaerobic fungi and who make up the majority of known hosts (2020). If present, anaerobic fungi are a minority among fungi in the colobine gut, since previous studies examining colobine mycobiota indicate that fungal populations in the

colobine gut are dominated by the phyla Ascomycota and Basidiomycota (Xu et al., 2015; Mann et al., 2019; Barelli et al., 2020).

In this study, we sought to characterize the prokaryotic and fungal microbiota of two captive representatives of strongly herbivorous foregut-fermenting and hindgut-fermenting NHPs. We attempted to enrich obligate anaerobic fungi of the Neocallimastigomycota from fecal samples of zoo-housed Eastern black and white colobus monkeys (*Colobus guereza*) and Western lowland gorillas (*G. gorilla gorilla*) using serial dilution methods pioneered by Theodorou *et al*. (2005). In addition, we employed a high-throughput amplicon sequencing approach targeting the 16S and ITS2 regions of prokaryotic and fungal ribosomal DNA, respectively, using a primer pool optimized to capture a larger alpha diversity of fungi than that seen in previous studies of the mycobiota of these two primates, in order to reveal the composition and diversity of both the prokaryotic and fungal communities in the digestive tracts of these two herbivorous non-human primate species, and to compare these two different digestive strategies. We used a subset of this data to perform a case study of differences in microbiome composition between two captive *G. gorilla gorilla* brothers, one of whom was reported to suffer from gastric dysfunction. Using these findings, we hope to illuminate more of the full diversity of the gut microbiota of these two physiologically divergent herbivores, and to understand how their unique microbiota contribute to their dietary strategies, with the aim to further understand the evolution of herbivory as a dietary strategy among the primates and to aid in the conservation of these vulnerable primate taxa.

3.2. Methods

3.2.1. Sampling and fungal enrichment procedures

Fecal samples were obtained from six *C. guereza* individuals of mixed sex and age housed at the Cincinnati Zoo (Cincinnati, OH, USA) and two adult male *G. gorilla gorilla* individuals, brothers "Bangori" and "Nzinga", housed in the same enclosure at the Santa Barbara Zoo (Santa Barbara, CA, USA). Identifying data for each individual, as well as dietary information, may be found in **Supplemental Data 6** (uploaded with dissertation). Fresh colobine fecal samples were pooled from all six individuals into one 50 mL sterile Falcon tube per individual immediately after defecation, in the case of *C. guereza*, or as soon as safely possible thereafter, in the case of *G. gorilla gorilla*. *C. guereza* feces was stored on ice on-site in Cincinnati and shipped overnight to Santa Barbara, CA, USA. Fresh gorilla fecal samples were collected in four 50 mL sterile Falcon tubes, two for each individual, delivered to the laboratory on the day of defecation and stored overnight in a Styrofoam™ box at room temperature prior to the isolation procedure. Dietary and health information were obtained for all captive primates. Immediately upon receipt, samples were transferred to the laboratory and the isolation process was initiated.

Enrichment of anaerobic fungi was performed in an anaerobic glove bag (Sigma-Aldrich, catalog number Z555525-1EA), fitted with two in-line Whatman HEPA-CAP™ 36 HEPA filters (Whatman, catalog number 6702-3600) for incoming and outgoing gas, under 100% CO_2, with an outgoing vacuum. For all inoculations, less than 1 g of feces was used to inoculate each of 3 (gorilla) or 5 (colobine) Hungate tubes containing 10 mL of Medium C, a complex medium containing ovine rumen fluid (Theodorou, Brookman, and Trinci, 2005), with 100 mg of insoluble carbon substrate (1% weight/volume), and with a final

concentration of 0.01 mg/ml chloramphenicol to reduce contamination by prokaryotes. Stock chloramphenicol (100x) was prepared in 40% ethanol, resulting in 0.4% ethanol in the enrichment tubes. Alfalfa stems were used as carbon substrate to match the predominantly dicotyledonous leaf diet of *C. guereza*, and reed canary grass was used to match the more lignified vegetative diet of *G. gorilla gorilla.* Both alfalfa and reed canary grass were provided by the US Dairy Forage Research Center (Agricultural Research Service, US Department of Agriculture) and were milled in a Model 4 Wiley Mill (Thomas Scientific) using a 4-mm screen size. For gorilla enrichment cultures, initial inoculation was performed with fecal samples from only one brother, Bangori, due to gastric health problems present in the other brother, Nzinga, as reported by the gorillas' caretakers. Hungate tubes were incubated in a shared incubator at 39°C. The remainder of the feces was preserved in 100% ethanol for DNA extraction and stored at -80°C. For colobine samples, five samples were immediately inoculated, then allowed to incubate at 39°C for four days, then diluted tenfold into five new tubes containing Medium C, substrate, and chloramphenicol, which were themselves diluted tenfold into five other new tubes containing Medium C, substrate, and chloramphenicol. For gorilla samples, three inoculated tubes were harvested in triplicate for immediate DNA sequencing, and three inoculated tubes were allowed to incubate at 39°C for four days, then diluted tenfold using a sterile needle into new tubes containing Medium C, substrate, and chloramphenicol.

3.2.2. DNA extraction, quality measurements, and sequencing

Metagenomic DNA was extracted from ethanol-preserved fecal samples and solids in enrichment tubes using the QIAamp PowerFecal kit (QIAGEN, catalog number 12830-50). The resulting concentration and quality of DNA was measured using a NanoDrop (Thermo

Scientific). The V3/V4 region of the prokaryotic 16S ribosomal DNA and the second internal transcribed spacer (ITS2) region of the fungal ribosomal DNA operon were amplified from extracted genomic DNA using the primers listed in **Appendix 3, Supplementary Table 3**. The 16S V3/V4 primers were derived from Illumina's 16S amplicon sequencing protocol (2013), which selected these primers from a study published by Klindworth et al. (2013). The ITS2 primers were derived from Canarini et al.'s (2021) and Tedersoo and Lindahl's (2016) studies. 16S sequences were amplified using Illumina's 16S protocol: initial denaturation at 95°C for 3 minutes, then 25 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 30 seconds, and a final elongation of 72°C for 5 minutes. ITS2 sequences were amplified using the following PCR conditions in Op De Beeck et al. (2014): initial denaturation at 95°C for 2 minutes, followed by 40 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 60 seconds, and a final elongation at 72°C for 10 minutes. Detailed library preparation and sequencing methods are presented in **Supplementary Data 7**. Libraries were multiplexed and sequenced on an Illumina MiSeq with MiSeq reagent kit v3, with a 600-bp indexed run.

3.2.3. Amplicon sequence variant analysis

A total of 18,047,870 paired-end 16S reads and 26,671,582 paired-end ITS2 reads from colobine and gorilla samples were obtained from sequencing; total read counts and bp for each sample are listed in **Supplemental Data 7** (uploaded with dissertation). Demultiplexed raw reads were loaded as artifacts into QIIME2 (version 2023.5; Bolyen et al., 2019). Read quality was checked at each processing stage with the q2-demux plugin's "summarize" function. For 16S reads, adapters were trimmed using the q2-cutadapt plugin's "trim-pairs" function, using default settings. For ITS2 reads, unmerged reads were trimmed to the ITS2

regions using the “trim-pair-output-unmerged” function of q2-itsxpress, which implements BBmap, BBmerge, and ITSx, with the parameter “p-cluster-id” set to 1.0 and default settings otherwise. Quality filtering was performed with the q2-quality-filter plugin’s “q-score” function. Trimmed, filtered, unmerged reads were then denoised using the “denoise-paired” function in q2-dada2 (the QIIME2 implementation of DADA2). For *C. guereza* 16S samples, errors learned from 476,711,800 total bases in 2,072,660 reads from 12 samples were used for DADA2 sample inference, and for ITS2 samples, this was done with errors learned from 425,074,056 total bases in 2,088,816 reads from 11 samples. For *G. gorilla gorilla* 16S samples (both Bangori and Nzinga combined), errors learned from 632,938,380 total bases in 2,751,906 reads from 15 samples were used for DADA2 sample inference, and for ITS2 samples, this was done with errors learned from 521,105,956 total bases in 2,679,208 reads from 10 samples. Additional run metadata may be accessed in **Supplemental Data 8** (uploaded with dissertation). Taxonomy was assigned using the SILVA v138.1 database for 16S reads, with the pre-trained “full-length animal corpus” classifier for the *C. guereza* samples and the pre-trained “full-length animal distal gut” classifier for the gorilla samples (Glöckner et al., 2017; Kaehler et al., 2019). For fungi, taxonomy was assigned using the UNITE QIIME2 classifier 9.0 (Nilsson et al., 2019). Samples were rarefied only prior to alpha and beta diversity calculation and calculation of differential abundance of more-abundant (>5% of relative abundance after rarefaction) taxa, as rarefaction may remove reads from rare taxa from the sample (McMurdie & Holmes, 2014), and the detection of rare taxa is a major focus of this work. 16S and ITS2 alpha diversity in each sample was represented by Shannon entropy, calculated using the “core-metrics-phylogenetic” function of the QIIME plugin “diversity” (**Appendix 3,**

Supplementary Table 4). Beta diversity was assessed with non-metric multidimensional scaling using the Bray-Curtis dissimilarity, unweighted UniFrac, and weighted UniFrac distance metrics, calculated using the "beta-group-significance" function of the QIIME2 plugin "beta-group-significance". Permutational multivariate analysis of variance (PERMANOVA) was performed with 999 permutations to test the hypothesis that the 16S and ITS2 composition of the fecal samples of the gorillas Bangori and Nzinga were the same, using the "beta-group-significance" function of the QIIME2 plugin "beta-group-significance". Individual differences in taxonomic groups between fecal samples and enrichments were compared by rarefying samples to the lowest sample depth above 500 (samples with depths lower than this were considered too small to be used in these analyses; sampling depths are available in **Supplemental Data 8**) using the "ancombc" function in QIIME2's "composition" plugin (tables of differential abundance used with full labels are available in **Supplemental Data 9**, uploaded with dissertation). All phylogenetic trees were visualized with the Interactive Tree of Life website.

3.2.4. Data availability

16S and ITS2 sequencing reads will be uploaded to NCBI's Genbank. Tables of ASVs associated with all samples may be found in **Supplemental Data 10** (uploaded with dissertation).

3.3. Results

3.3.1. The Colobus guereza *fecal microbiome*

We enriched fecal samples from *Colobus guereza* on Medium C containing chloramphenicol and alfalfa, then extracted DNA from the enrichment cultures and the fecal

samples we used to inoculate them and sequenced 16S and ITS2 sequences from them. After bioinformatic processing, a total of 4,277,471 non-chimeric merged reads were recovered from 16S sequencing runs and a total of 2,673,174 merged, non-chimeric reads were recovered from ITS2 sequencing runs, from fecal samples and enrichments combined. A total of 2382 prokaryotic 16S ASVs and 1,087 different fungal ITS2 ASVs were recovered from *C. guereza* fecal samples and enrichments on Medium C containing chloramphenicol and alfalfa. Bacterial ASVs spanned 16 phyla (**Figure 24**). Archaeal ASVs spanned three phyla, comprising methanogens from Euryarchaeota and Thermoplasmatota from genera known to be capable of producing methane from hydrogen, methanol, and methylamines, as well as an unidentified Woesearchaeales archaeon from the Nanoarchaeota. Most of the diversity (1727/2372 ASVs, 72.8%) of 16S features found in fecal samples and enrichments belonged to the phylum Firmicutes. An additional 245 (10.3%) belonged to the Bacteroidota. Of the total of 2372 16S ASVs acquired from the *C. guereza* gut microbiota, only 1,450 (61.1%) were detected in fecal samples.

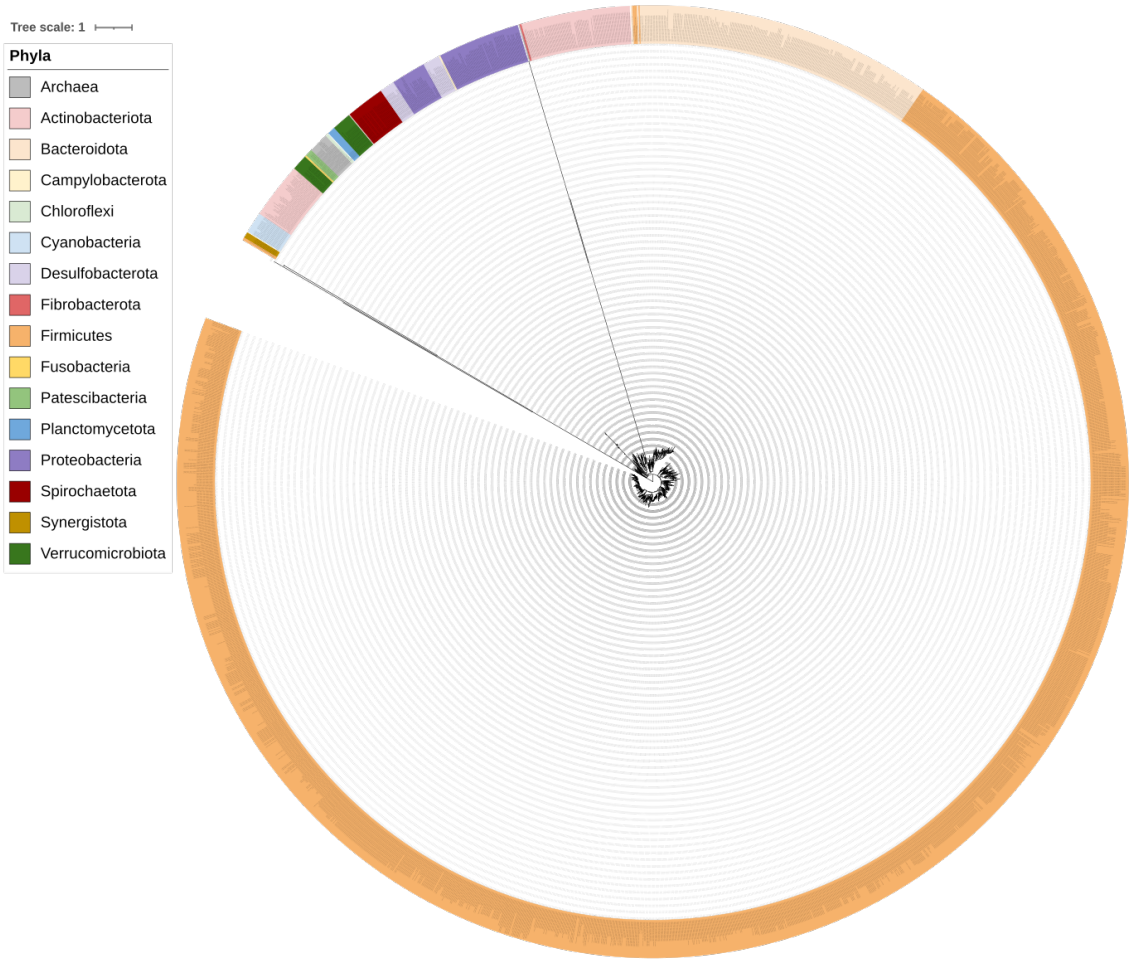


Figure 24. Phylogenetic tree of 16S ASVs from *Colobus guereza* fecal samples and enrichments.

Most (834/1087, 76.7%) fungal ASVs clearing quality filtration were unidentified at the phylum level (**Figure 25**). All but four that were identifiable at the phylum level or lower belonged to the Ascomycota (183/1087, 16.8%) or Basidiomycota (66/1087, 6.07%). Of the remaining four, all were early-diverging fungi. Two were identified as *Mucor circinelloides*, of the phylum Mucoromycota. One was identified as belonging to the Chytridiomycota with no further classification. The last was identified as *Caecomyces churrovis*, a member of the class Neocallimastigomycota (subkingdom Chytridiomyceta), a

clade of fungi widespread in the digestive tracts of large herbivores that boasts the largest arsenal of carbohydrate-active enzymes (CAZymes) of all organisms on Earth (Seppälä et al., 2017). Of the 1087 ITS2 ASVs acquired from *C. guereza* gut microbiota, 401 (36.8%) were detected in fecal samples.

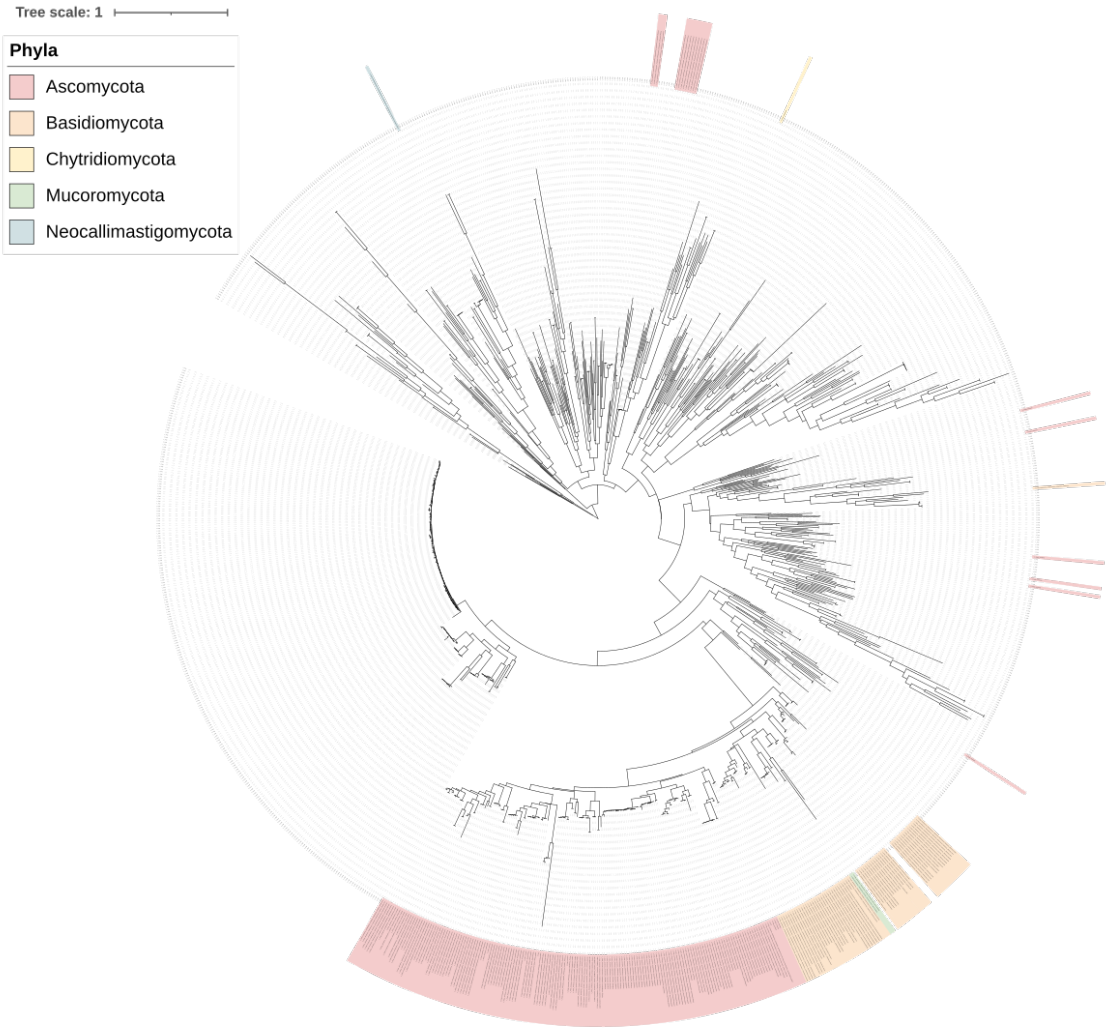


Figure 25. Phylogenetic tree of ITS2 ASVs from *Colobus guereza* fecal samples and enrichments.

More than 80% of 16S sequences found in captive *C. guereza* feces, averaged across technical replicates from all three biological samples, came from just 10 prokaryotic orders

(**Figure 26**): Bacteroidales, Oscillospirales, Lachnospirales, Christensenellales, Spirochaetales, Methanobacteriales, Peptostreptococcales-Tissierellales, Bacilli RF39, and Erysipelotrichales. Bacteroidales, Oscillospirales, and Lachnospirales contain many cellulolytic bacteria (L. Liu et al., 2021). Bacilli RF39 is a Tenericutes lineage of Bacilli widespread in the gut of humans and domestic animals, and is likely a clade of acetate and hydrogen producers (Yong Wang et al., 2020), potentially indicating dysbiosis of the source microbiome due to its association with diarrhea in captive Sichuan golden snub-nosed monkeys (*Rhinopithecus roxellana*) (Zhu et al., 2018). The Christensenellales were most highly represented by the Christensenellaceae R-7 group, which is mainly involved in amino acid, peptide, and lipid metabolism and is widespread among the gut bacteria of primates (Waters & Ley, 2019). The most abundant archaea belonged to the Methanobacteriales, encompassing approximately 1-4.7% of prokaryotic abundance. At the phylum level, Firmicutes and Bacteroidota together accounted for more than 80% of the prokaryotic population, and together with Verrucomicrobiota, Spirochaetota, and Euryarchaeota they accounted for more than 90%. No more than 0.44% of prokaryotes in any fecal sample were unidentifiable at at least the phylum level.

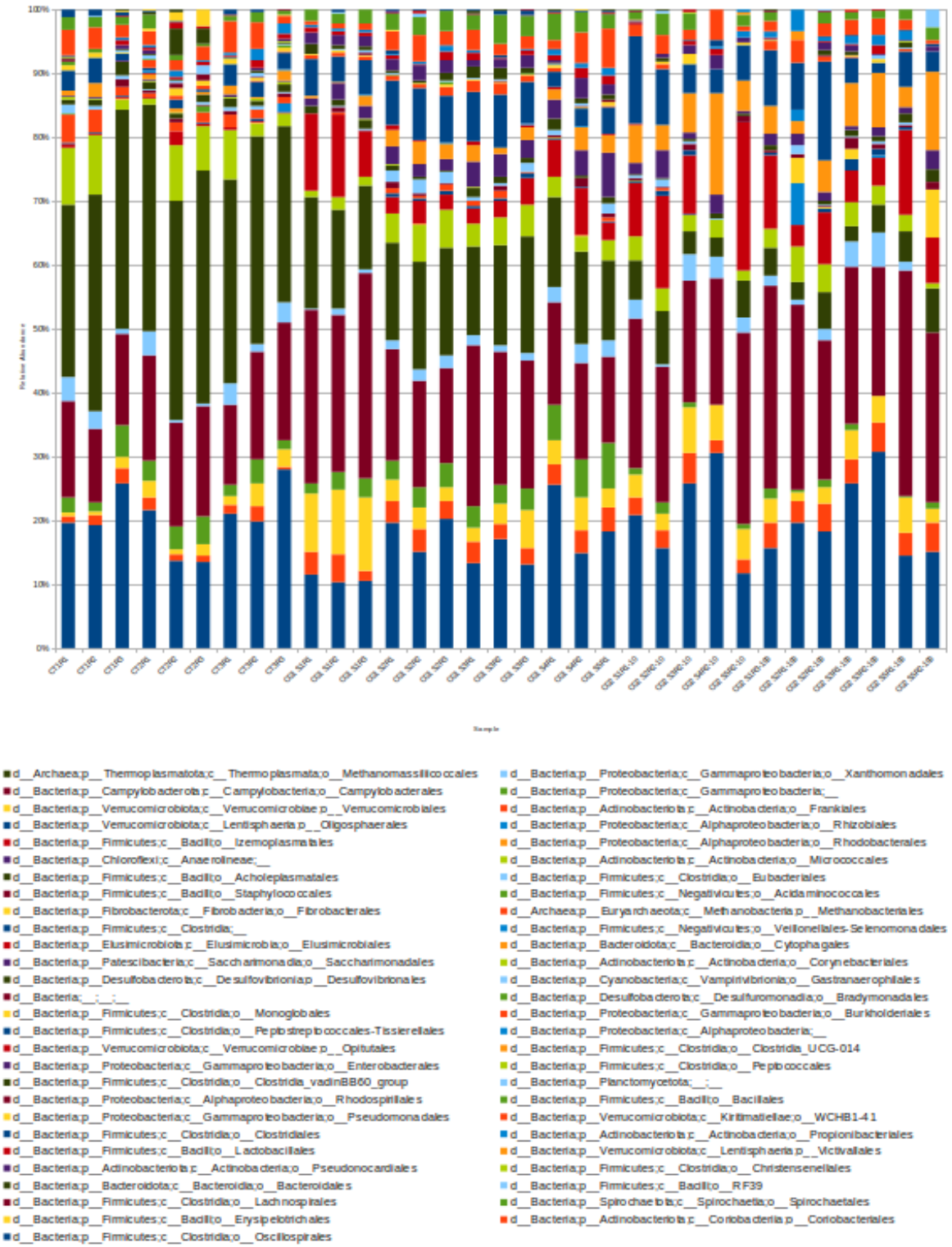


Figure 26. Barplot of relative abundance of prokaryotic taxa from fecalsamples and enrichments from captive *Colobus guereza*, as measured by 16S sequence abundance, at the order level. "CT#R#": fecal samples. "CG#S#R#": enrichments. CG1: enrichment generation 1. CG2: enrichment generation 2. R#: replicate. #-10: 10-fold

dilution from the previous generation. #-100: 100-fold dilution from the previous generation.

Most ITS2 sequences recovered from *C. guereza* fecal samples and enrichments, by relative abundance, were unidentified at the phylum level (**Figure 27**). Of taxa identified at the phylum level, most fecal samples contained mostly yeasts of the Saccharomycetales (Ascomycota), of which the most prevalent member was *Cyniclomyces* sp. One sample contained mostly fungi of the Pezizales (Ascomycota). Smaller relative abundances of sac fungi of the Eurotiales (Ascomycota), Microstromatales (Basidiomycota), and Taphrinales (Ascomycota) were also found. 63.4% to 100% of ASVs by relative abundance were unidentified at the genus or species level.

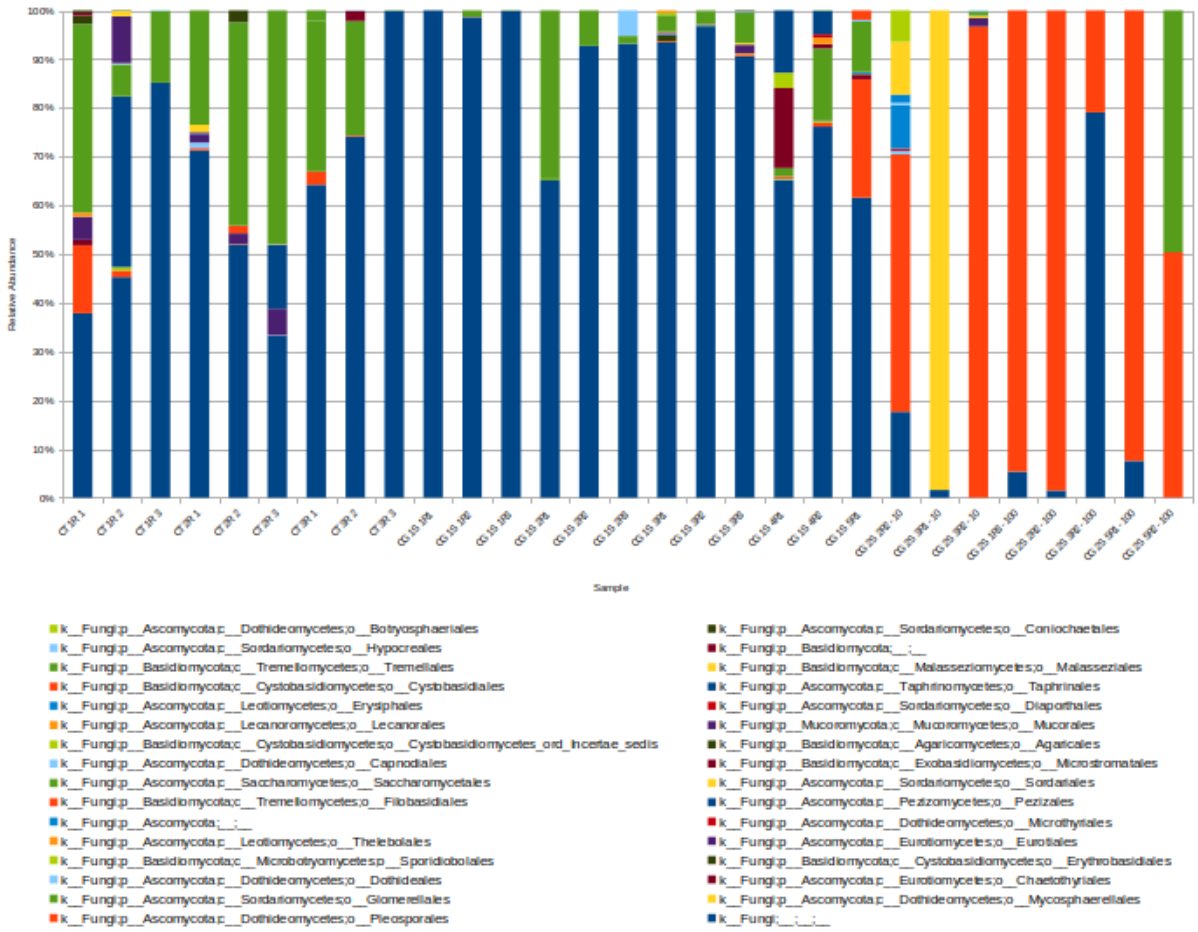


Figure 27. Barplot of relative abundance of fungal taxa from fecal samples and enrichments from captive *Colobus guereza*, as measured by ITS2 sequence abundance, at the order level. "CT#R#": fecal samples. "CG#S#R#": enrichments. CG1: enrichment generation 1. CG2: enrichment generation 2. R#: replicate. #-10: 10-fold dilution from the previous generation. #-100: 100-fold dilution from the previous generation.

3.3.2. The gorilla fecal microbiome

We enriched fecal samples from a healthy male individual of *Gorilla gorilla gorilla* on Medium C containing chloramphenicol and reed canary grass, then extracted DNA from the enrichment cultures and the fecal samples we used to inoculate them and sequenced 16S and ITS2 sequences from them. After bioinformatic processing, a total of 1,292,546 merged,

non-chimeric reads were recovered from 16S sequencing runs and 2,300,626 merged, non-chimeric reads were recovered from ITS2 sequencing runs, from fecal samples and enrichments combined. A total of 1367 prokaryotic 16S ASVs and 1,822 fungal ITS2 ASVs were recovered from healthy *G. gorilla gorilla* fecal samples and enrichments on Medium C containing chloramphenicol and reed canary grass. Prokaryote ASVs spanned two phyla of archaea (methanogens from the Euryarchaeota and Thermoplasmatota) and 13 phyla of bacteria (**Figure 28**). Most of the diversity of 16S features (1037/1367 features, 75.85%) came from the Firmicutes, with an additional 165 (12.07%) from the Bacteroidota. Of the 1,367 total 16S ASVs acquired from the healthy gorilla gut microbiome, 865 (63.2%) were detected in fecal samples.

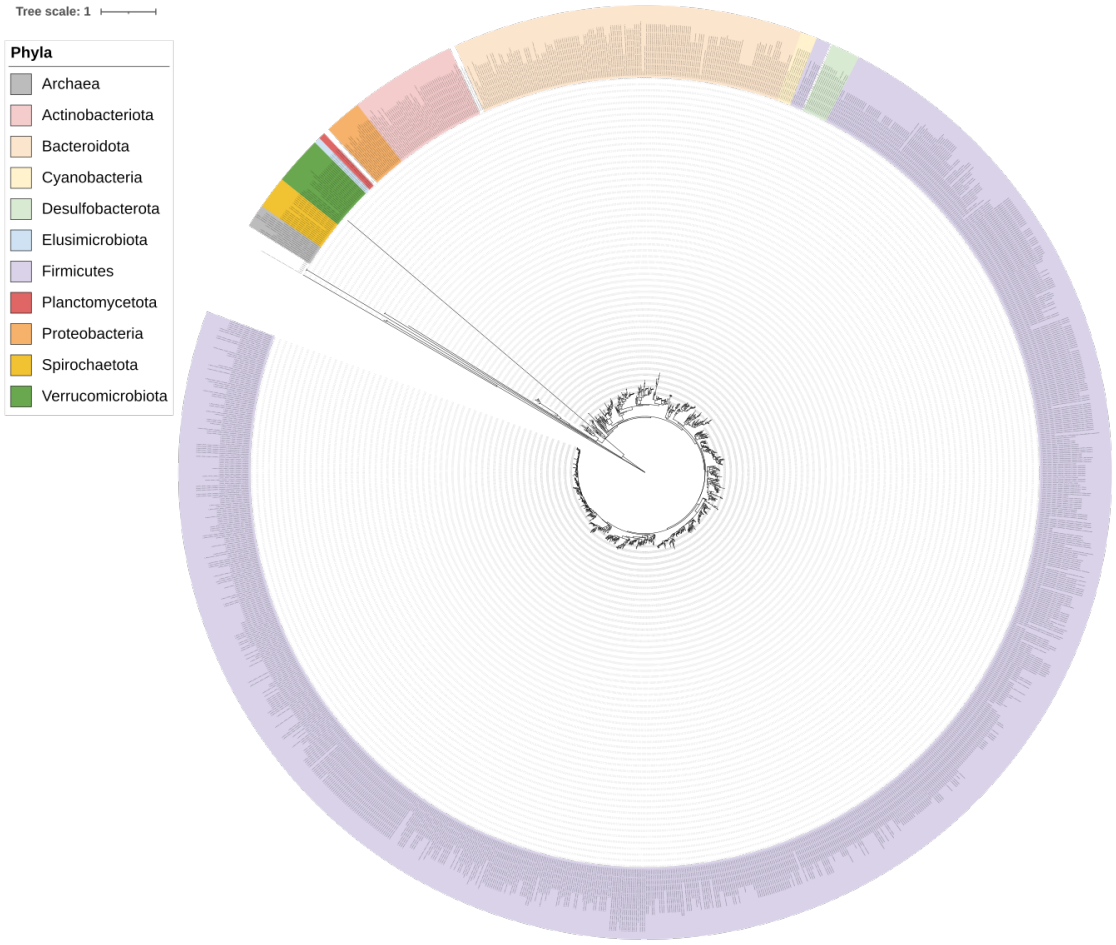


Figure 28. Phylogenetic tree of 16S ASVs from gorilla fecal samples and enrichments.

As with *C. guereza* samples, most (1247/1822, 68.4%) fungal ASVs recovered from healthy gorilla samples were unidentified at the phylum level (**Figure 29**). 401 (22%) ASVs were classified in the Ascomycota, 152 (8.3%) were classified in the Basidiomycota, and 6 (0.32%), 1 (0.054%), and 15 (0.82%) ASVs were classified in the Chytridiomycota, Mortierellomycota, and Mucoromycota, respectively. Of the 1,822 total ITS2 ASVs acquired from the healthy gorilla gut microbiome, 264 (14.4%) were detected in fecal samples.

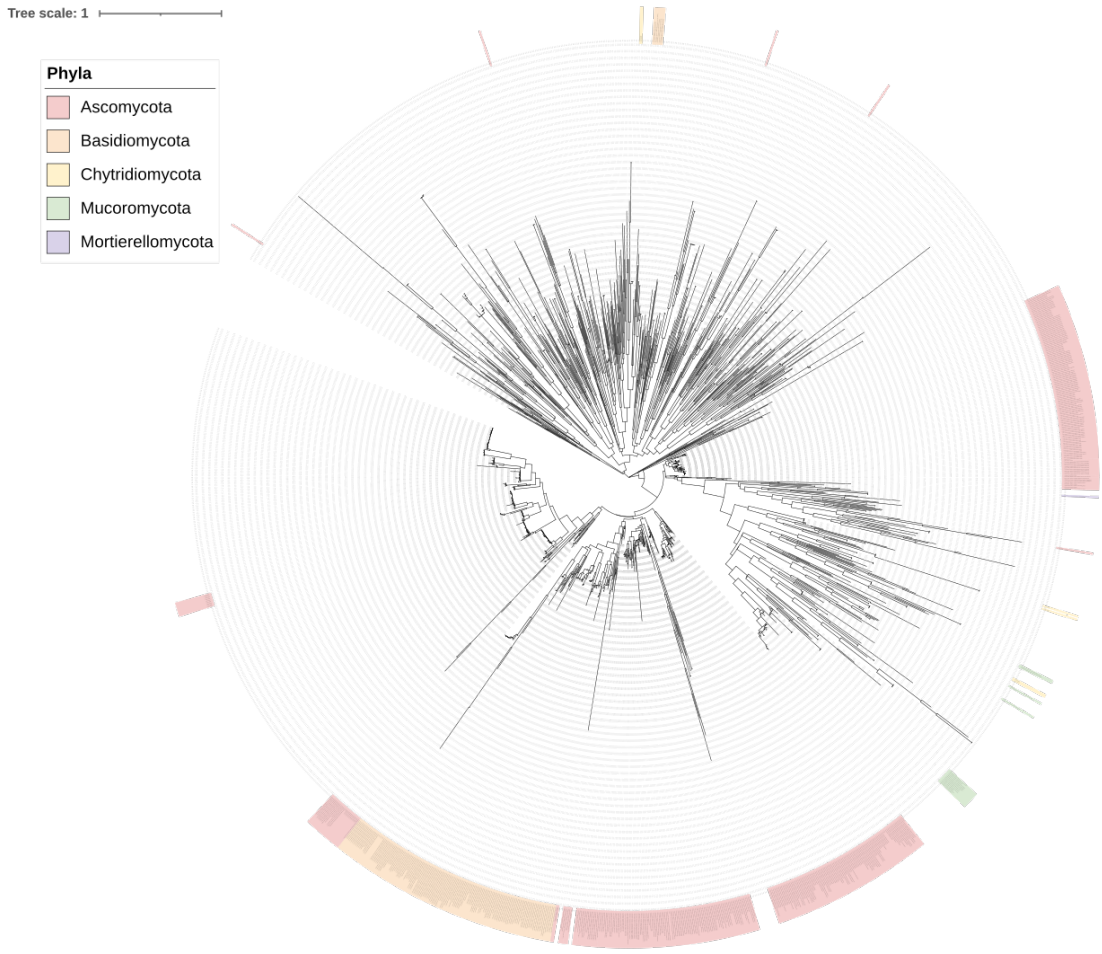


Figure 29. Phylogenetic tree of ITS2 ASVs from gorilla fecal samples and enrichments.

More than 80% of 16S sequence variants found in healthy captive *G. gorilla gorilla* feces came from just 10 prokaryotic orders (**Figure 30**): Oscillospirales, Bacteroidales, Lachnospirales, Christensenellales, Spirochaetales, Erysipelotrichales, Methanobacteriales, Clostridia UCG-014, Coriobacteriales, and Izemoplasmatales. Coriobacteriales is made up of cellulolytic species (Gupta et al., 2017). Archaea made up 1.5-6% of relative abundance, with the majority belonging to the Euryarchaeota (1.2-5.2%) and a smaller population belonging to the Thermoplasmatota (0.4-1.2%). At the phylum level, Firmicutes and Bacteroidota together accounted for more than 70% of the prokaryotic population, and

together with Verrucomicrobiota, Spirochaetota, and Euryarchaeota they accounted for more than 90%. More than 90% of strains were unidentified at the species level.

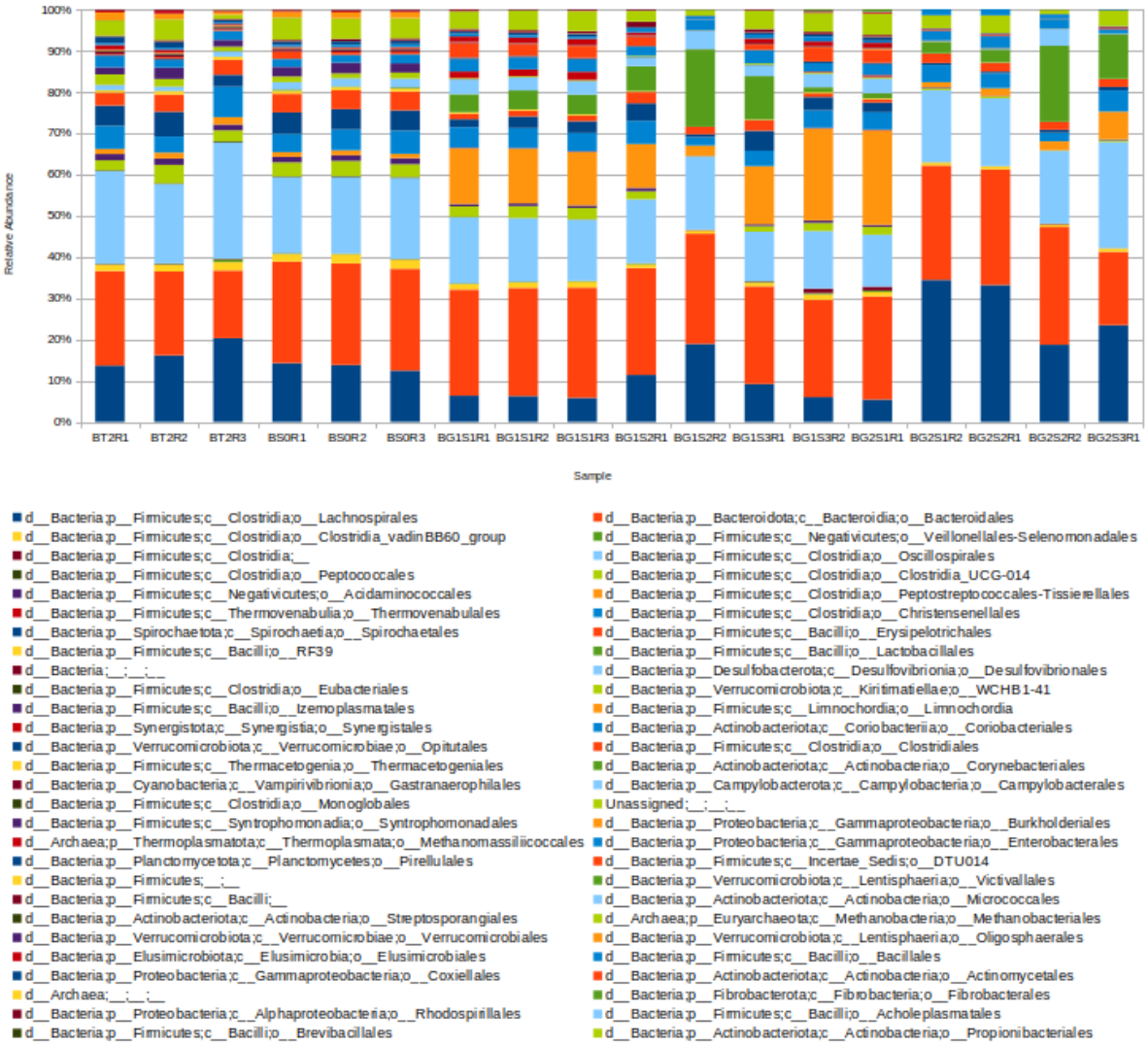


Figure 30. Barplot of relative abundance of prokaryotic taxa in *Gorilla gorilla gorilla* fecal samples and enrichments, as measured by 16S sequence abundance, at the order level. "BT#R#": fecal samples. "BS#R#"/"BG#S#R#": enrichments. S#: BG1: enrichment generation 1. BG2: enrichment generation 2. R#: replicate.

Most fungal ITS2 sequences recovered after denoising were unidentified at the phylum level (**Figure 31**). Of identified ITS2 sequences, most from fecal samples belonged to the Eurotiales (Ascomycota). Two samples also contained at least 0.87% of relative abundance of the Pleosporales (Ascomycota) and Tremellales (Basidiomycota), and one

contained approximately 5.5% of relative abundance of Agaricales (Basidiomycota). 1.8% to 10.4% of relative abundance was composed of *Aspergillus* sp. in all three fecal samples, including 1.87% and 10.4% *Aspergillus ruber* in two samples.

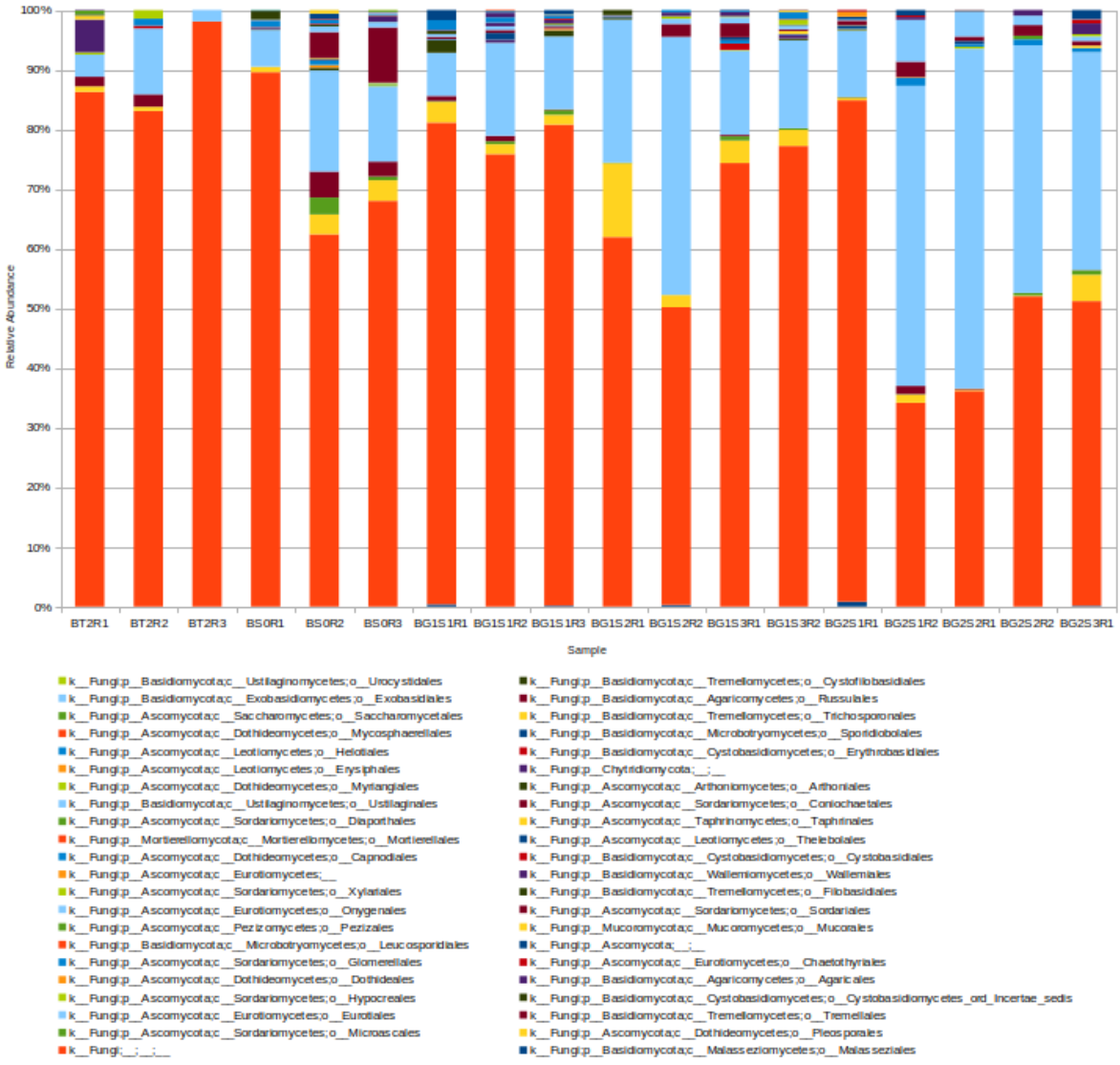


Figure 31. Barplot of relative abundance of fungal taxa in Gorilla gorilla gorilla fecal samples and enrichments, as measured by ITS2 sequence abundance.

3.3.3. Unsuccessful anaerobic fungal enrichment attempts nonetheless enrich diverse consortia of taxa undetected in primate fecal samples

After receiving fecal samples from *C. guereza* and *G. gorilla gorilla*, we immediately carried out enrichment under CO_2 in Medium C containing chloramphenicol, to select for fungi, as well as containing alfalfa (*C. guereza*) and reed canary grass (*G. gorilla gorilla*), to match each primate's diet and aid in enriching fungi of the Neocallimastigomycota. We did not successfully enrich anaerobic fungi from either *C. guereza* or *G. gorilla gorilla* feces. However, we enriched a variety of bacterial and fungal taxa that we did not detect in DNA from fecal samples alone and that, to our knowledge, are previously unrecorded members of *C. guereza* and *G. gorilla gorilla* gut microbiota.

Of the 2372 total 16S ASVs detected from *C. guereza* gut microbiota, 1943 (81.9%) were detected in enrichments, a 1.34-fold increase over the 1450 (61.1%) detected in fecal samples (**Figure 26**). Among the archaea, one member of the Woesearchaeales was detected in enrichment consortia and undetected in fecal samples. Additional members of the Actinobacteria were detected in fecal samples; one member of the Bifidobacteriales (*Pseudoscardovia suis*), several additional members of the Corynebacteriales, one member of the Kineosporiales (*Kineococcus* sp.), several additional members of the Micrococcales and Propionibacteriales, and one member of the Thermoleophilia (*Solirubrobacter* sp.). Among the Bacteroidota, additional members of the Tannerellaceae, Chitinophagales, Cytophagales, and Flavobacteriales were detected. Three members of the Campylobacteria (*Helicobacter* sp.) and Chloroflexi (Anaerolineae, including order SBR1031) were detected. Among the Firmicutes, two members of *Anoxybacillus* sp. (including *A. voinovskiensis*) were detected, as well as Erysipelotrichaceae genus UCG-009, *Aerococcus sanguinicola*,

several Lactobacillaceae (including *Lactobacillus intestinalis*, *Latilactobacillus* sp., *Ligilactobacillus* sp., *Secundilactobacillus* (*Lactobacillus*) *kimchicus*, *Weissella confusa*, and *W. oryzae*), several Streptococcaceae, one Thermoactinomycetales (*Thermoactinomyces* sp.), *Clostridium botulinum*, two Lachnospiraceae (*Coprococcus comes* and *Frisingicoccus* sp.), two Oscillospirales (*Anaerofilum* sp. and *Paludicola* sp.), several members of Family XI of the Peptostreptococcales-Tissierellales, and two members of the Veillonellales-Selenomonadales (*Selenomonas* sp. and *Megasphaera* spp.). One member of the Fusobacteria (*Leptotrichia* sp.) and one member of the Gracillibacteria of the Patescibacteria (Abscondibacteriales SR1) were detected. Among the Proteobacteria, members of the Rhizobiales, Rhodobacterales, and Sphingomonadales of the Alphaproteobacteria were detected, as were members of the Acidithiobacillales and Xanthomonadales of the Gammaproteobacteria. *Cloacibacillus everyensis* of the Synergistota was detected. Among bacteria whose ASVs were not filtered out during rarefaction to calculate differential abundance at the order level, Lactobacillales, Enterobacterales, Clostridia order UCG-014, Peptostreptococcales-Tissierellales, Saccharimonadales, Erysipelotrichales, and Coriobacteriales were differentially enriched in the enrichments relative to the fecal samples, and Peptococcales, Bacteroidales, and Kiritimatiellae WCHB1-41 were differentially depleted (**Supplemental Data 9**).

Of the 1087 ITS2 ASVs acquired from *C. guereza* gut microbiota, 775 (71.2%) were detected in enrichment consortia, a 1.93-fold increase over the 401 (36.8%) detected in fecal samples (**Figure 27**). *Caecomyces churrovis*, an anaerobic fungus detected in *C. guereza* fecal samples, was notably absent from enrichment consortia, as were two *Mucor circinelloides* (Mucoromycota) ASVs. From the Ascomycota, ASVs were recovered from

the Botryosphaeriales (*Botryosphaeria scharfii*), the Microthyriales (*Neoscolecobasidium agapanthi*), the Mycosphaerellales *Lapidomyces aloidendricola* and *Parateratosphaeria* sp.), the Lecanoromycetes (*Scoliciosporum fabisporum*), and the Taphrinomycetes (*Taphrina inositophila*). From the Basidiomycota, ASVs were recovered from the Ustilaginomycetes and the Wallemiomycetes (*Wallemia canadensis*). One ASV from the Chytridiomycota was recovered. Among fungi whose ASVs were not filtered out during rarefaction to calculate differential abundance at the order level, no order was significantly enriched, but the Saccharomycetales were depleted. At the family to species level, fungi of the Didymellaceae (Ascomycota: Dothideomycetes) were enriched, whereas *Cyniclomyces* sp. (Ascomycota: Saccharomycetes) were depleted. In the last stages of enrichment, consortia were entirely, or almost entirely, overtaken by fungi of the Pleosporales.

Of the 1,366 total 16S ASVs acquired from the healthy gorilla gut microbiome, 1,233 (90.2%) were detected in enrichment consortia, a 1.43-fold increase over the 862 (63.1%) detected in fecal samples (**Figure 30**). No change in composition was detected in the archaeal community. Enrichment produced an expansion of detectable taxa in almost all phyla of bacteria. Among the Actinobacteriota, few additional taxa were observed in enrichments, with a loss of the Corynebacteriales. Among the Bacteroidota, observable species diversity increased in the genus *Bacteroides*; *B. ovatus* and *B. uniformis* were detectable in fecal samples, whereas *B. fragilis* and *B. thetaiotaomicron* became detectable in enrichments with no loss of the former two species. Among the Firmicutes, several taxa from class to species level were only detectable in enrichments. *Enterococcus durans* was only detectable in enrichments, as well as 4 additional members of the Streptococcaceae (*Streptococcus infantis* was detectable in fecal samples, whereas *Lactococcus garvieae*,

Leuconostoc pseudomesenteroides, *Streptococcus alactolyticus*, and an unidentified *Streptococcus* sp. were detectable only in enrichments). Among the Lachnospirales, cellulolytic bacteria highly abundant in both fecal samples and enrichments, *Cellulosilyticum* sp., *Eisenbergiella* sp., and *Hungatella* sp. were only detectable in enrichments. Among the Oscillospirales, also highly abundant in both fecal samples and enrichments, *Intestinimonas butyriciproducens*, *Anaerotruncus* sp., and *Caproiciproducens* spp. were only detectable in enrichments. Among the Peptococcales, *Peptococcus* sp. was only detectable in enrichments. Among the Peptostreptococcales-Tissierellales, several members of Family XI were enriched: *Anaerococcus* spp. (including *A. vaginalis*), *Anaerosalibacter* spp., *Sporanaerobacter* spp., *Gallicola* sp., *Peptoniphilus* spp. (including *P. urinimassiliensis*), and *Tepidimicrobium* spp. One member of the Limnochordia (*Hydrogenispora* sp.) was only detectable in enrichments. Members of the Thermacetogenia and Thermovenabula were only detectable in enrichments. Finally, members of the phylum Synergistota (*Pyramidobacter* spp., including *P. piscolens*) were only detectable in enrichments. Among bacteria whose ASVs were not filtered out during rarefaction to calculate differential abundance at the order level, Lactobacillales, Thermovenabulales, unidentified Firmicutes, and Peptostreptococcales-Tissierellales were differentially enriched in the enrichments relative to the fecal samples, and Clostridia vadinBB60, Erysipelotrichales, Spirochaetales, Victivallales, Clostridia UCG-014, Gastranaerophilales, Pirellulales, Acidaminococcales, Rhodospirillales, Elusimicrobiales, Bacilli RF39, Campylobacterales, Acholeplasmatales, Kiritimatiellae WHCB1-41, Burkholderiales, Peptococcales, Monoglobales, Oligosphaerales, and Izemoplasmatales were differentially depleted (**Supplemental Data 9**).

Enrichment on reed canary grass grew a rich community of fungi not detected in fecal samples, particularly from the Mucoromycota; three *Mucor* sp. ASVs were recovered from fecal samples, whereas from enrichments, 15 ASVs from five families of the Mucoromycota (Cunninghamellaceae, Lichtheimiaceae, Mucoraceae, Pilobolaceae, Rhizopodaceae) were recovered. Of the 1,822 total ITS2 ASVs acquired from the healthy gorilla gut microbiome, 1,706 (93%) were detected in enrichment consortia, a 6.46-fold increase over the 264 (14.4%) detected in fecal samples (**Figure 31**). From the Ascomycota, ASVs were recovered from the Arthoniomycetes (*Pentagenella langei*), Dothideales, Mycosphaerellales (*Vermiconia antarctica*, Mycosphaerellacaceae), representatives of 14 families of the Dothideomycetes (in contrast to unidentified Didymellaceae and *Massarina papulosa*), Chaetothyriales, 10 species of *Aspergillus* and 22 different *Aspergillus* sp. ASVs (in contrast to *Aspergillus ruber* and *A. xerophilus* and 3 *Aspergillus* spp.), 10 different *Penicillium* spp., *Talaromyces* spp., *Thermomyces crustaceus*, *T. dupontii*, *T. ibadanensis*, Onygenales, Leotiomycetes, Pezizomycetes, 5 additional orders of Sordariomycetes, *Taphrina inositophila*, and an unidentified Protomycetaceae bacterium. From the Basidiomycota, ASVs were recovered from 6 families of the Agaricomycetes, one member of the Agaricostilbomycetes, 4 families of the Cystobasidiomycetes (and one unidentified member), *Graphiola phoenicis* (Exobasidiomycetes), three different members of the Malasseziomycetes, two *Leucosporidium* spp. (Microbotryomycetes), members of 4 orders of the Tremellomycetes, members of 2 orders of the Ustilaginomycetes, and four species of the Wallemiomycetes (*Wallemia canadensis, W. muriae, W. sebi*, and one unidentified *Wallemia* sp.). One species, *Mortierella polygonia*, was recovered from the Mortierellomycota. Among fungi whose ASVs were not filtered out during rarefaction to

calculate differential abundance at the order level, Sordariales, Onygenales, Microascales, Wallemiales, and Glomerellales were differentially enriched in the enrichments relative to the fecal samples, and no orders were differentially depleted (**Supplemental Data 9**). At the family to species level, *Aspergillus fumigatus*, *Thermomyces ibadanensis*, *T. dupontii*, *Chrysosporium lobatum*, *A. protuberus*, *Chaetomium thermophilum*, and *Malbranchea cinnamomea* were enriched, whereas unidentified fungi were depleted.

3.3.4. A tale of two brothers: a case study of gastric dysfunction and dysbiosis in two captive gorillas

Gastrointestinal issues are common among captive primates, including gorillas and colobines (Amato et al., 2016; Frankel et al., 2019; Lukas, 1999; McKenzie et al., 2017; Strong et al., 2016; Zhu et al., 2018). The two gorillas from which fecal samples were collected, Bangori and Nzinga, were brothers housed in the same enclosure at the Santa Barbara Zoo and fed similar diets. Both were administered chlorhexidine mouthwash (ChloraZinc™) once a day and a probiotic (Culturelle™ Probiotics Ultimate Strength) twice a day. Nzinga's keepers reported current chronic symptoms of gastric irritation, loose stools, and diarrhea. We took the opportunity to compare the microbiota of two genetically closely-related adult male gorillas housed in the same enclosure and fed similar diets, in an effort to illuminate the potential source of Nzinga's symptoms and improve the welfare of other captive gorillas with similar symptoms. After bioinformatic processing, Nzinga's fecal samples generated a total of 230,228 merged, non-chimeric reads from 16S sequencing runs and 321,178 merged, non-chimeric reads from ITS2 sequencing runs. A total of 808 different 16S ASVs and 299 different ITS2 ASVs were recovered from Nzinga's fecal samples.

All but one of the 807 16S ASVs recovered from Nzinga's fecal samples were assigned to either the Bacteria or Archaea. Of these 807 prokaryotic ASVs, 13 were assigned to the Archaea (3 to the Euryarchaeota, 10 to the Thermoplasmatota), and 800 to the Bacteria. The majority (589/807, 72.9%) of 16S ASVs were assigned to the Firmicutes, and 93 (11.5%) were assigned to the Bacteroidota (**Figure 32**).

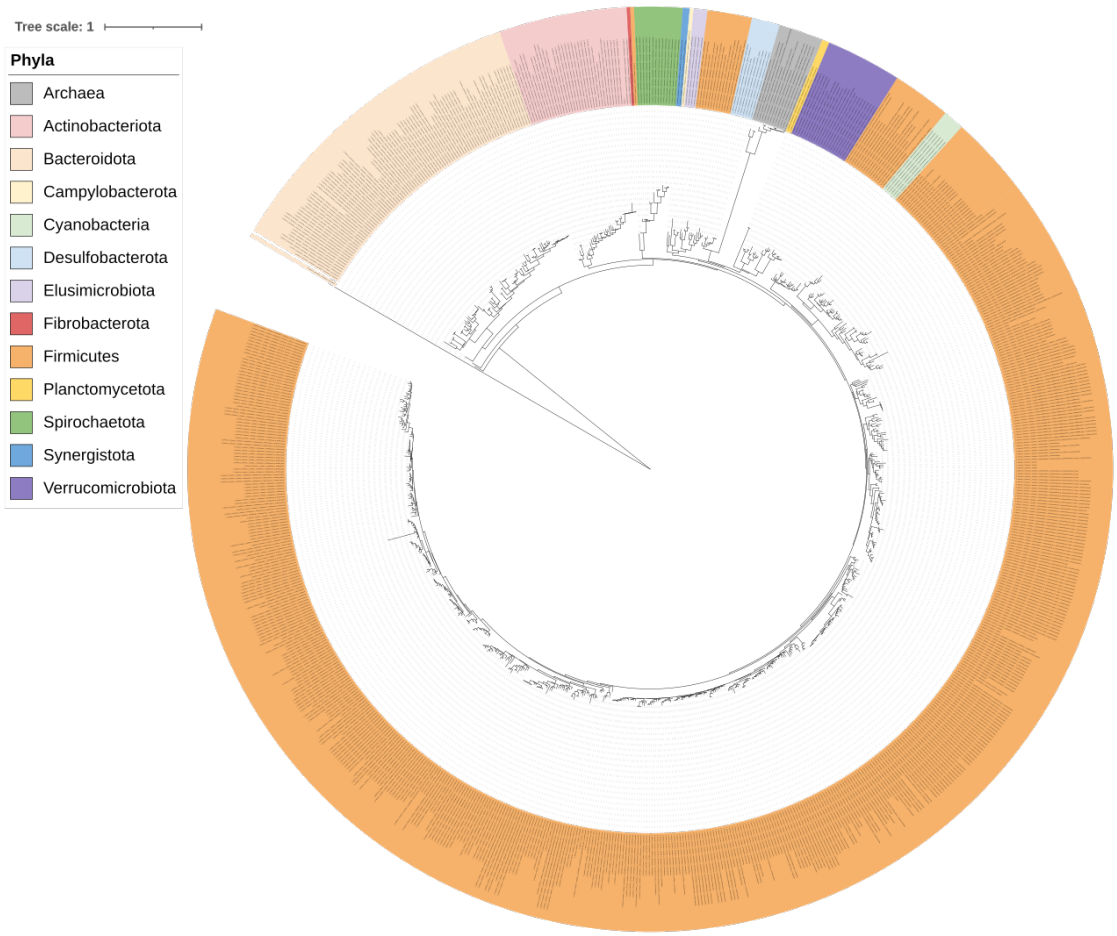


Figure 32. Phylogenetic tree of 16S ASVs in fecal samples from an individual of *Gorilla gorilla gorilla* with gastric dysfunction.

234 of 299 (78.2%) of fungal ASVs identified in Nzinga's fecal samples were unidentified at the phylum level (**Figure 33**). Of those identified at the phylum level, 43 (14.3%) belonged to the Ascomycota, predominantly Dothideomycetes (Capnodiales,

Dothideales, and Pleosporales) and Eurotiomycetes (split almost evenly between *Aspergillus* sp., including *Aspergillus ruber*, and *Penicillium* sp., including *Penicillium tropicum*), but also including the Lecanoromycetes, Leotiomycetes, Saccharomycetes, and Sordariomycetes. 20 (6.6%) belonged to the Basidiomycota (5 from the Cystobasidiomycetes, 1 from the Microbotryomycetes, 12 from the Tremellomycetes, and 2 from the Wallemiomycetes), and 2 (0.66%) belonged to the Mucoromycota (both were *Mucor* sp.).

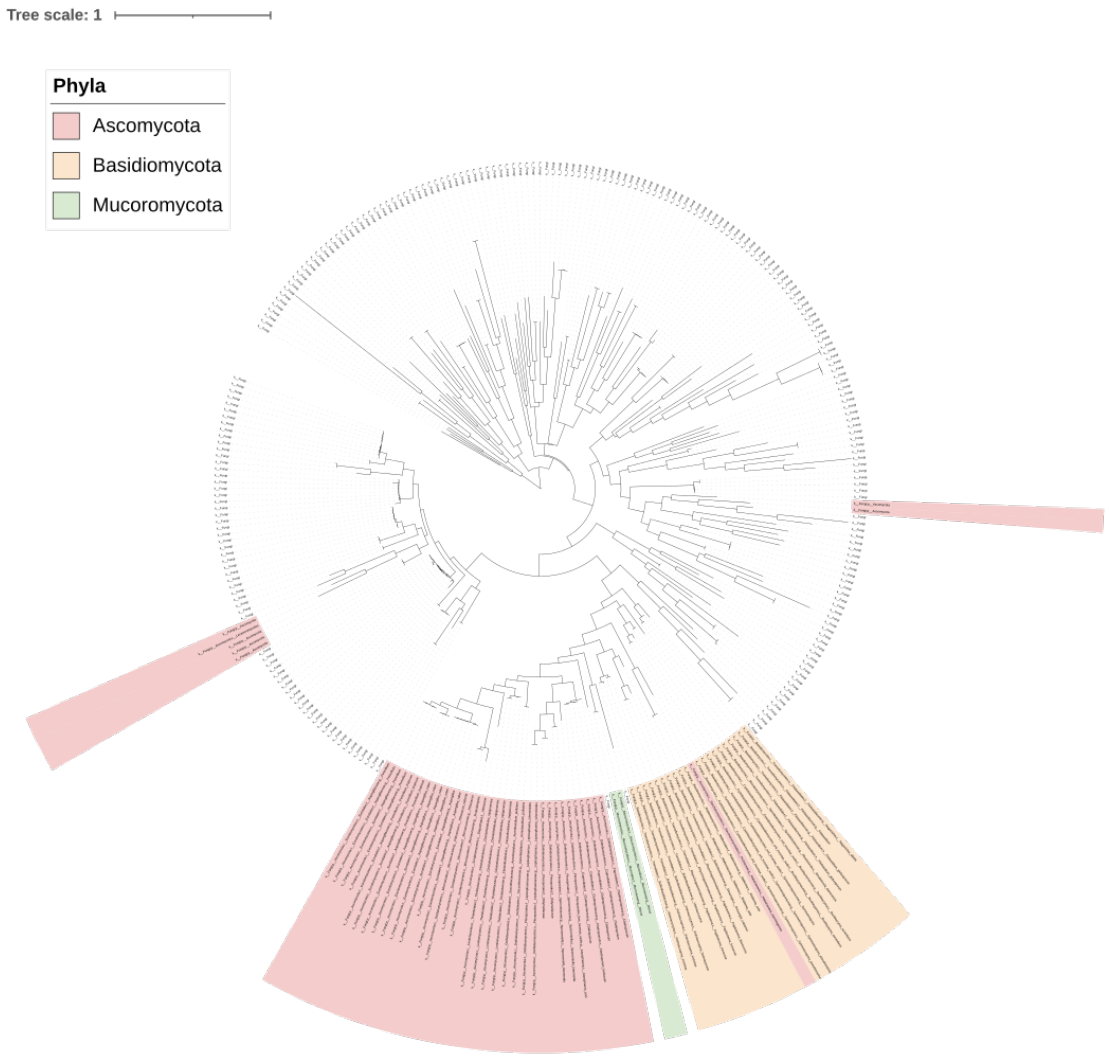


Figure 33. Phylogenetic tree of ITS2 ASVs in fecal samples from an individual of *Gorilla gorilla gorilla* with gastric dysfunction.

Bangori's and Nzinga's prokaryotic gut communities exhibited numerous differences in the relative abundance of prokaryotic taxa (**Figure 34**). The prokaryotic community in Bangori's feces was significantly more diverse than the prokaryotic community in Nzinga's feces, as measured by a Kruskal-Wallis test comparing Shannon entropies in their fecal samples (n = 3 replicates for each gorilla, H = 3.857142857142854, p = 0.049534613435626915) (**Figure 35**). When evaluating beta diversity, Bangori's and Nzinga's prokaryotic communities were not significantly different from each other in

composition by simple PERMANOVA (test statistic = 5.163007, p = 0.1), or unweighted (test statistic = 5.163007, p = 0.098) and weighted (test statistic = 6.42439, p = 0.094) Unifrac distances (**Figure 36**). Among bacteria whose ASVs were not filtered out during rarefaction to calculate differential abundance at the order level, Lactobacillales, Bacillales, Synergistales, and Coxiellales (Legionellales) were more abundant in Nzinga's fecal samples relative to Bangori's fecal samples (in particular, lactic acid bacteria of the Lactobacillales were present in Nzinga's feces and entirely undetected in Bangori's feces), and Bacilli RF39, Burkholderiales, Clostridiales, Elusimicrobiales, Eubacteriales, Gastranaerophilales, Izemoplasmatales, unidentified members of class Bacilli, and Acholeplasmatales were less abundant (**Supplemental Data 9**).

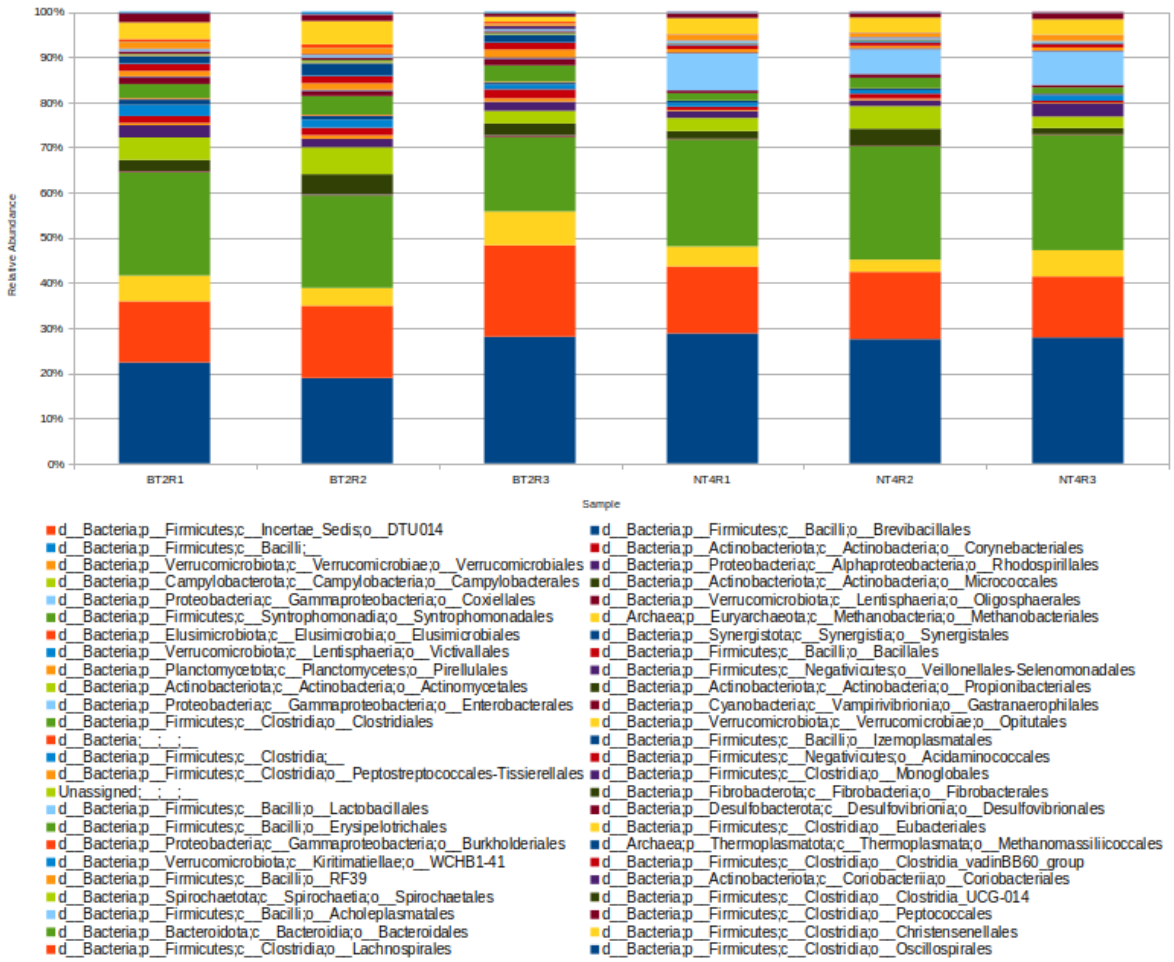


Figure 34. Barplot of relative abundance of 16S ASVs in fecal samples from one healthy *Gorilla gorilla gorilla* individual and one with gastric dysfunction. "BT#R#": samples from Bangori, a healthy male gorilla. "NT#R#": samples from Nzinga, a male gorilla with gastric dysfunction.

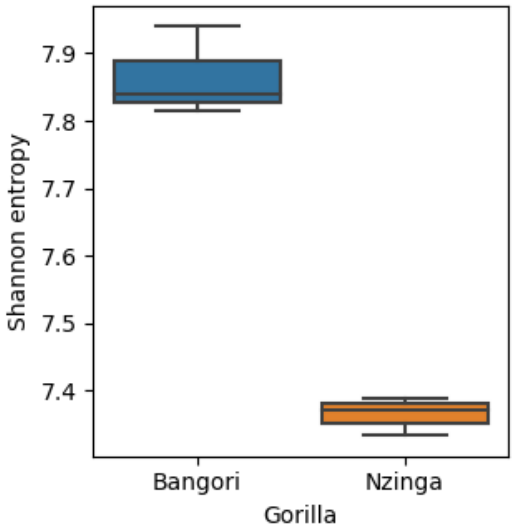


Figure 35. Plot of Shannon entropies of 16S ASVs in Bangori's and Nzinga's fecal samples.

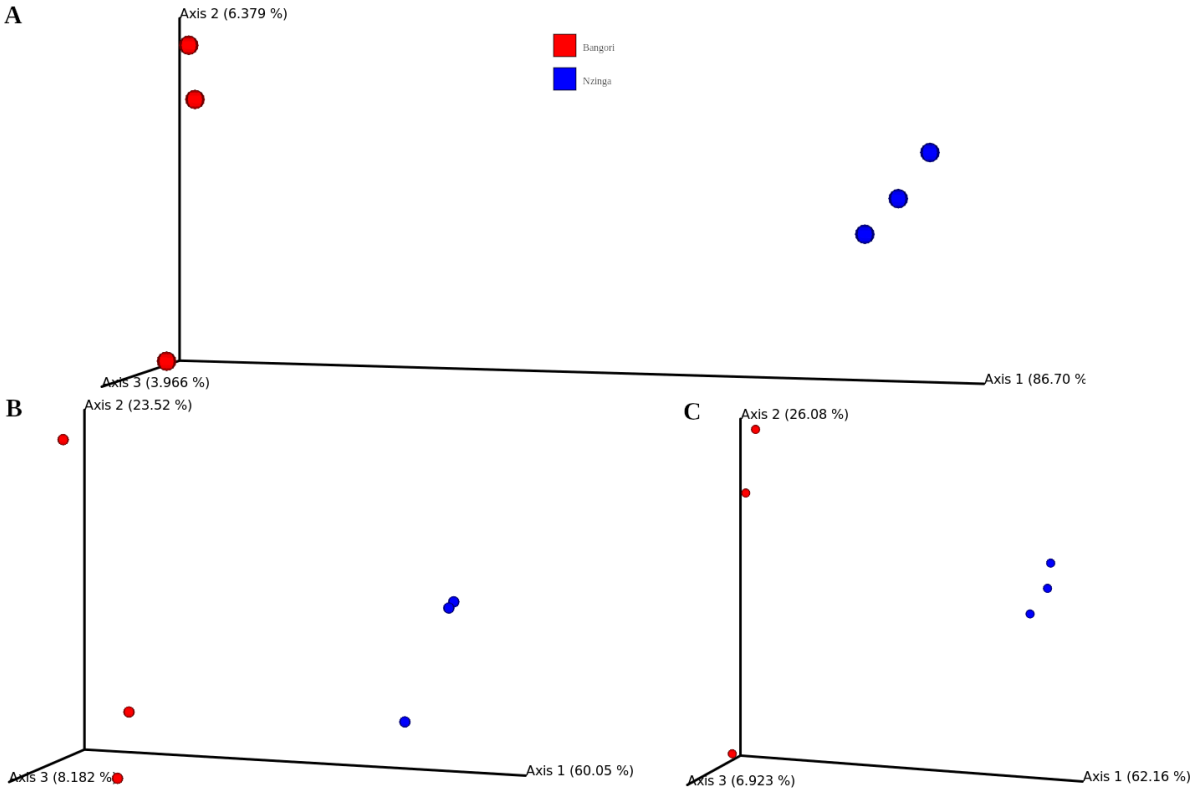


Figure 36. Beta diversity in 16S ASV communities between Bangori's and Nzinga's fecal samples. A: Bray-Curtis dissimilarity. B: Unweighted Unifrac distance. C: Weighted Unifrac distance.

Fewer differences were detectable between the fecal mycobiota of Bangori and Nzinga (**Figure 37**). Their mycobiota did not differ significantly in alpha diversity and were not significantly different from each other according to a simple PERMANOVA (test statistic = 1.210927, p = 0.104), or according to unweighted (test statistic = 1.210927, p = 0.089) or weighted (test statistic = 1.452759, p = 0.077) Unifrac distances (**Figures 38 and 39**). Among fungi whose ASVs were not filtered out during rarefaction to calculate differential abundance, no fungi were more abundant in Nzinga's fecal samples relative to Bangori's fecal samples, whereas unidentified Saccharomycetales and *Aspergillus xerophilus* were depleted.

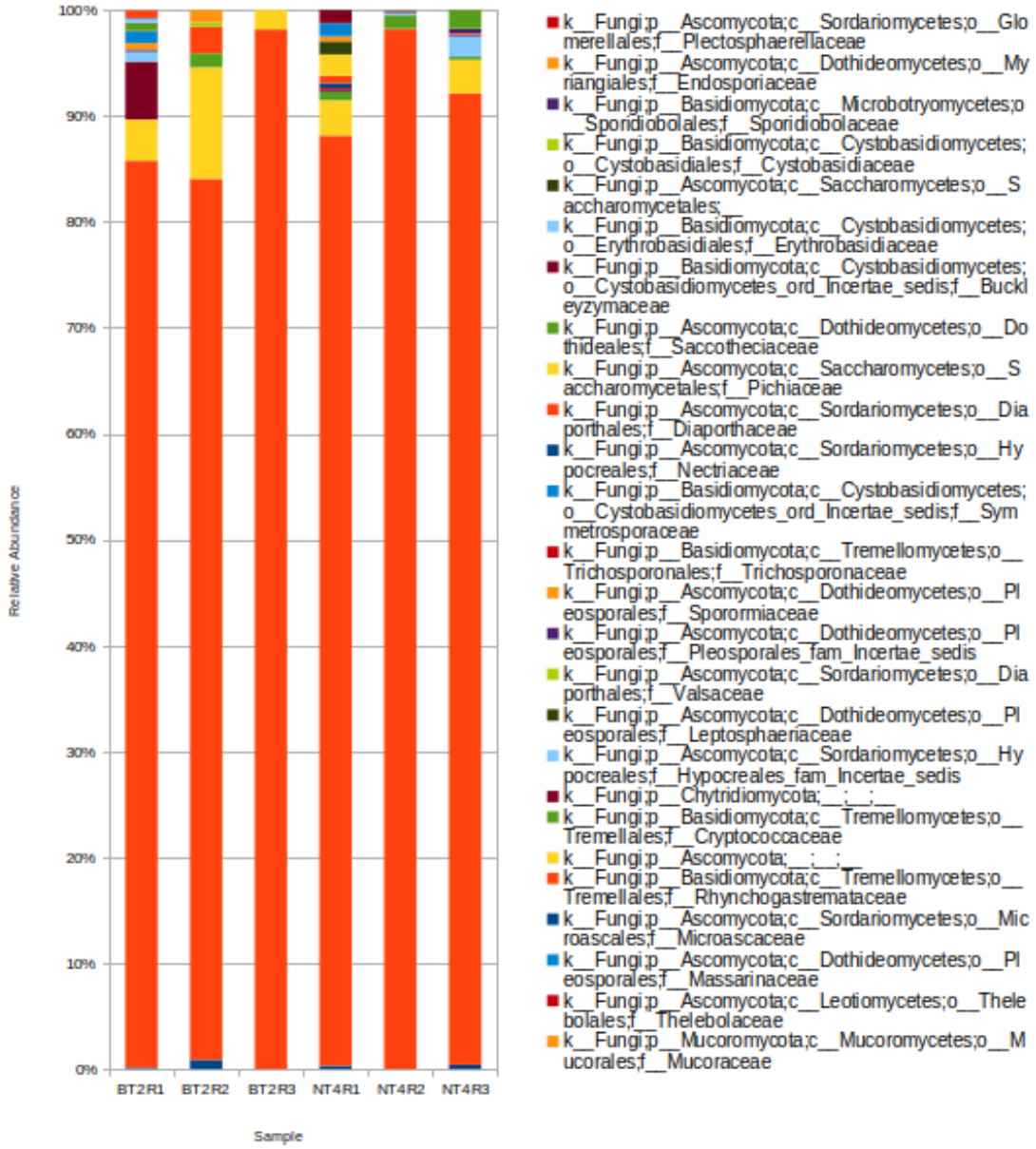


Figure 37. Barplot of relative abundance of ITS2 ASVs in fecal samples from one healthy *Gorilla gorilla gorilla* individual and one with gastric dysfunction. "BT#R#": samples from Bangori, a healthy male gorilla. "NT#R#": samples from Nzinga, a male gorilla with gastric dysfunction.

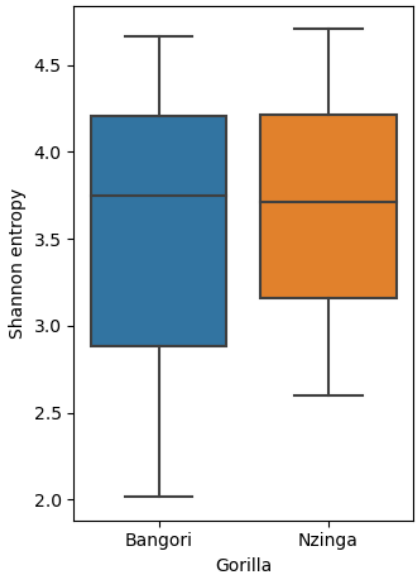


Figure 38. Plot of Shannon entropies of ITS2 ASVs in Bangori's and Nzinga's fecal samples.

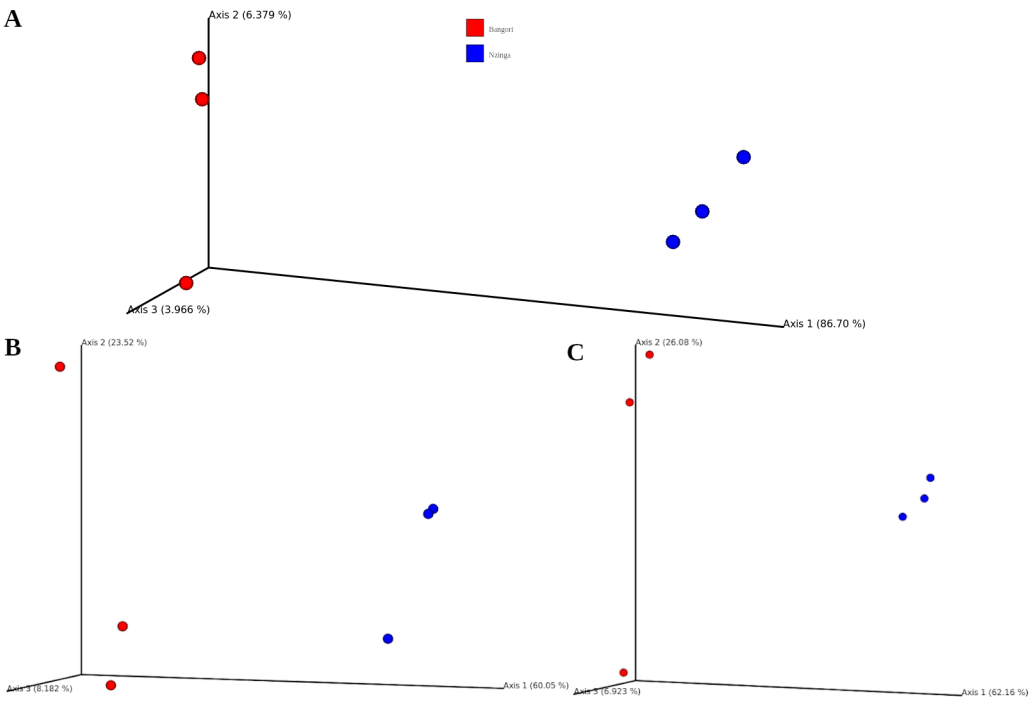


Figure 39. Beta diversity in ITS2 ASV communities between Bangori's and Nzinga's fecal samples. A: Bray-Curtis dissimilarity. B: Unweighted Unifrac distance. C: Weighted Unifrac distance.

3.4. Discussion

3.4.1. Increasing resolution of captive C. guereza *and gorilla microbiota through anaerobic enrichment culture*

In this study, we sought to characterize the fecal prokaryotic and fungal populations of *Colobus guereza* and *Gorilla gorilla gorilla* to understand how the composition of the bacterial and fungal communities in these primates, particularly the presence of rare taxa, contributes to lignocellulolysis and anaerobic digestion in their digestive tracts. Additionally, we sought to enrich and isolate anaerobic fungi of the Neocallimastigomycota, who carry out degradation of highly recalcitrant biomass in the gut of a variety of terrestrial herbivores (Liggenstoffer et al., 2010). This is the first study, to our knowledge, to utilize parallel enrichment on complex substrate to enrich rarer members of the gut microbiome to enable the resolution of rare taxa in non-human primate microbiota. Enrichment culture on alfalfa and reed canary grass revealed a rich community of lignocellulolytic prokaryotes and fungi not detected in fecal samples. Even though captivity humanizes and decreases the diversity of primate gut communities (Clayton et al., 2016; Frankel et al., 2019; McKenzie et al., 2017), we may underestimate the diversity they still possess.

We enriched fecal samples anaerobically on alfalfa and reed canary grass – complex substrates chemically matched as closely as possible to the leafy and more pectinous, or more fibrous and lower in average pectin content, wild diets of the animals from which we obtained fecal samples (Doran-Sheehy et al., 2009; Matsuda & Hummel, 2022). Across both species and sets of ASVs, a larger proportion of ASVs were found in enrichment consortia than in fecal samples. In *C. guereza*, this enabled the detection of 1.34 times as many prokaryotic ASVs and 1.93 times as many fungal ASVs. In *G. gorilla gorilla*, this enabled

the detection of 1.43 times as many prokaryotic ASVs and 6.46 times as many fungal ASVs. Many of the newly-detected enriched ASVs (such as the newly-detectable members of the Lactobacillales, Lachnospirales and Oscillospirales in both *C. guereza* and *G. gorilla gorilla* samples) belonged to members of taxa known to be resident in the gut of other herbivorous primates, as well as other large herbivores (Biddle et al., 2013; Ford et al., 1958), and many of these were obligate anaerobes with no spore-forming stage or other ability to survive for long periods of time in an oxygenated environment, suggesting that these strains were less likely to be transients and more likely to be rarer residents of the microbiota. Conversely, many of the newly detected fungal ASVs were likely to be transients, such as *Botryosphaeria scharfii*, a member of a genus of plant pathogens found in *C. guereza* samples (Dai et al., 2022), or *Mortierella polygonia*, a filamentous fungus from *G. gorilla gorilla* feces found in plant tissues, soil, and the rhizosphere (Vandepol et al., 2020).

3.4.2. Captive C. guereza *and gorilla gut microbiota are rich in lignocellulolytic taxa with complementary degradative abilities*

At a larger scale, the prokaryotic communities detected from both fecal samples and enrichments from *C. guereza* and *G. gorilla gorilla* reflect many previously-published studies on their gut prokaryotic communities (for example, Amato et al., 2020; Frankel et al., 2019; Hale et al., 2018; Houtkamp et al., 2022; Mccord et al., 2014; Yildirim et al., 2010; Zeng et al., 2022; Zhu et al., 2018), describing a community dominated chiefly by Firmicutes and Bacteroidota, with abundant cellulolytic and fermentative bacteria dominated by the Oscillospirales, Lachnospirales, and Bacteroidales and containing a variety of clades particularly prevalent among primate microbiota, such as the Christensenellales. Both gut archaeal communities were similar in both taxonomic composition and relative abundance;

the Nanoarchaeota ASV observed in *C. guereza* enrichment consortia was likely transient. The enrichment process enriched a variety of less-abundant prokaryotic taxa in both species' microbiota. In *C.* guereza, there was an expansion of ASVs in members of fermentative clades such as the Propionibacteriales, Lactobacillaceae, Lachnospiraceae, and Oscillospirales, and several additional phyla, such as the Campylobacteria, Chloroflexi, and Synergistota were also detected. In *G. gorilla gorilla*, a similar expansion was observed, in cellulolytic taxa such as *Cellulosilyticum* sp., *Caproiciproducens* sp., and *Hungatella* sp., and particularly in mostly-fermentative taxa such as *Intestimonas butyriciproducens*, members of the genus *Bacteroides*, and *Streptococcus* spp.; in addition, additional clades such as Family XI of the Peptostreptococcales-Tissierellales and the Synergistota were only detectable in enrichments. Increased observed diversity in the bacterial community in the gut suggests a likely underestimation of the lignocellulolytic and fermentative repertoire of these microbiota; some of these rare taxa may disproportionately contribute to lignocellulolysis.

The fungal communities in both *C. guereza* and *G. gorilla gorilla* were both dominated by unknown fungal clades, and otherwise primarily composed of fungi of the Ascomycota and Basidiomycota, which likewise conforms to previous detailed studies of the mycobiota of these primates and their relatives (Barelli et al., 2020; Houtkamp et al., 2022; Mann et al., 2019; Sharma et al., 2022; Xu et al., 2015). In *C.* guereza, enrichment eventually skewed recovered fungal ASVs toward the Pleosporales and depleted the *Cyniclomyces* sp. population, but facilitated the detection of what were largely clades of plant pathogens or lichen-forming fungi most likely introduced from the diet – the Lecanoromycetes, Microthyriales, Mycosphaerellales, and the Taphrinomycetes. Botryosphaeriales have been previously detected in colobine gut mycobiota (Barelli et al., 2020). By contrast, in *G.*

gorilla gorilla, the dramatic 6.46-fold increase in the number of ITS2 ASVs recovered from enrichment consortia revealed increases in diversity found in the early-diverging Mucoromycota, as well as *Aspergillus* spp. and the Sordariomycetes, from the species to order levels. As with cellulolytic bacteria, this observed diversity, especially in clades already established as gut fungi within a species, may imply an underestimation of gut degradative capacity.

To our knowledge, this is the most detailed picture yet presented of the composition of captive *C. guereza* and *G. gorilla gorilla* fecal mycobiota. However, we chose not to remove ASVs present in only one sample in order to capture rare taxa in the microbiome. This increases the probability that some amount of the dataset being analyzed will constitute taxonomic "noise" from either artifacts of data processing or from fungi that are more likely to be transient through the gut and not permanent residents. Firm establishment of residency by gut fungi is difficult, but the likelihood that singletons represent taxonomic "noise" may be mitigated with a careful examination of the enriched taxa and their environment: in short, if the fungi in question belong to taxa known not to belong in a digestive tract in the first place, whether via their environmental source or metabolic capacity, they are transient (Lavrinienko et al., 2021). Of the fungi identified in this study from either *C. guereza* or *G. gorilla gorilla* digestive tracts, most are likely to be transients due to the inability to carry out anaerobic metabolism, as the environment of the vertebrate gut is anoxic (Ley et al., 2008). Aside from the obligately anaerobic Neocallimastigomycota, all fungi capable of growth in anaerobic environments are facultative anaerobes, and this ability is distributed sporadically across fungal taxa, with the vast majority of fungi being obligate aerobes (Alexopoulos et al., 1996). To date, no fungi of the Chytridiomycota are known to be either

obligately or facultatively anaerobic, though some species of the Chytridiomycota may survive, but not grow, under anaerobic conditions (Gleason et al., 2007). *Mortierella polygonia*, a filamentous fungus identified from *G. gorilla gorilla* feces, is, like other Mortierellomycota, an obligately aerobic fungus found in plant tissues, soil, and the rhizosphere (Li, C.Y. et al., 2021; Vandepol et al., 2020). Of the Ascomycota and Basidiomycota, several species regularly found in the human gut microbiome may also be permanent residents of *C. guereza* and *G. gorilla gorilla*: *Rhodotorula* spp. (Basidiomycota: Microbotryomycetes) and *Malassezia* spp. (Basidiomycota: Malasseziomycetes), present in both colobus and gorilla samples, are frequently isolated from human gut samples and widely regarded as residents of human gut microbiota (Miceli et al., 2011; Raimondi et al., 2019). Few yeasts are facultative anaerobes, though several taxa present in both primates, including *Candida* spp., *Cyniclomyces* spp. and *Saccharomyces* spp. (Ascomycota: Saccharomycetes), and the anaerobic yeast form of *Mucor* spp. (Mucoromycota: Mucoromycetes), are capable of carrying out fermentation anaerobically, and are regularly isolated from other primate samples or from those of other vertebrate herbivores (Forsythe & Parker, 1985; Miceli et al., 2011; Raimondi et al., 2019; Sypherd et al., 1979). An unidentified fungus in the genus Fusarium was identified; at least one member of this genus (*F. oxysporum*) exhibits anaerobic metabolism, but the rest of the genus is anaerobic (Shoun et al., 1991).

The fungal “dark matter” of a NHP gut sample – the proportion of fungi unidentified at the phylum level, constituting the majority of both *C. guereza* and *G. gorilla gorilla* mycobiota – may contain important lignocellulolytic fungi, both facultative and obligate anaerobes. The UNITE database, used to classify all ITS2 ASVs in this study, contains a

substantial number of unidentified clades at the phylum, class, and order levels (Nilsson et al., 2016). The vast majority of fungi in fecal samples were unidentified, in both species. Tedersoo et al.'s analyses of the full rRNA operon from global soil samples placed most unidentified sequences in known clades, but found tens of monophyletic groups and single sequences representing novel clades, some likely to be novel at the phylum level (2020). Employing Tedersoo et al.'s strategy of sequencing of the full rRNA operon is likely the best option to identify those fungi that were unidentified in this study, but successful identification of all fungi is contingent on description and classification of these fungi by other investigators. Early-diverging fungi are likely to be overrepresented among unidentified fungal ITS sequences (Lazarus & James, 2015; Picard, 2017; Tedersoo et al., 2017, 2020). The ITS region itself poses problems as a taxonomic barcode for fungi due to its variation in length and intra-strain variability, and ITS surveys may underestimate the true diversity of fungi present in an environment; the 28S D1/D2 has been suggested as an alternate due to its lower degree of variation in length and less intra-strain variability, and its use has been validated in the Dikarya and Neocallimastigomycota (Callaghan et al., 2015; Edwards et al., 2019; Fell et al., 2000; Gade et al., 2017; Hanafy et al., 2020; Koetschan et al., 2014; Kwiatkowski et al., 2012; Scorzetti et al., 2002).

A small number of ASVs of the early-diverging phyla Chytridiomycota, Mortierellomycota, and Mucoromycota were identified in fecal samples of both *C. guereza* and *G. gorilla gorilla*. In addition, two sequences of one fungus, *Caecomyces churrovis* were identified in fecal samples of *C. guereza*. Of the early-diverging fungal clades, the anaerobic fungi of the Neocallimastigomycota are prevalent in the gut of herbivores, and are responsible for breaking down the most recalcitrant fiber (Haitjema et al., 2017).

Assignment of fungi from the Neocallimastigomycota to the Chytridiomycota is not unheard of in fungal rDNA amplicon surveys (for example, Reynolds et al., 2022); the Neocallimastigomycota were formerly classified with the Chytridiomycota, and are included with this phylum and the phyla Monoblepharidomycota and Caulochytriomycota in the subkingdom Chytridiomyceta (Tedersoo et al., 2018). However, they are the only class of the Chytridiomyceta already known, through an extensive body of research, to live as permanent residents of the herbivore gut, as the others carry out their life cycles in freshwater and soil (Gleason et al., 2008; Hess et al., 2020). Zoospores from this clade are capable of surviving desiccation outside the animal, and it is possible this ASV was derived from transient zoospores, but Davies et al. showed that the population of survivors from zoospores derived from cattle – whose mycobiome is mostly composed of Neocallimastigomycota – halved after 18 weeks (Davies et al., 1993). This increases the likelihood that members of the Neocallimastigomycota are resident in at least *C. guereza*, and potentially that of both species. However, if anaerobic fungi are not permanent residents of the colobine gut, these primates represent an exception to the general rule that foregut fermenters will harbor Neocallimastigomycota (Gruninger et al., 2014). Anaerobic fungal genomes contain an arsenal of carbohydrate-active enzymes exceeding that of other fungi by approximately an order of magnitude (Seppälä et al., 2017b). In this case, one critical future research direction involves understanding how fungi with drastically smaller repertoires of degradative enzymes work together to carry out sufficient plant degradation for their colobine host. As previously mentioned, among the Mucoromycota, several members of the genus *Mucor* were enriched from both *C. guereza* and gorilla samples and may be gut residents. *Mucor* sp. are known for their dimorphic growth, dependent on oxygen

concentration; under aerobic conditions, they grow as a filamentous fungus, but under anaerobic conditions, they grow as a yeast (Sypherd et al., 1979).

We failed to enrich fungi of the Neocallimastigomycota in enrichment consortia, most likely due to the near-absence of anaerobic fungi in *C. guereza* fecal samples, except for *Caecomyces churrovis*, and their ostensible complete absence in *G. gorilla gorilla* fecal samples. One factor potentially limiting the enrichment of *C. churrovis* is its non-rhizoidal morphology and substrate preference. Previous studies of this fungus indicate that even though it is capable of growth on complex plant biomass, which is infiltrated by the rhizoids of rhizoidal Neocallimastigomycota during degradation, it grows more quickly on simple sugars and hemicellulose and is easily outcompeted by other fungi on plant biomass (Brown et al., 2021; Henske et al., 2017; Peng et al., 2021). The *C. guereza* diet is primarily composed of leaves, and they prefer leaves that are lower in cellulose, higher in hemicellulose and lignin, rich in overall neutral detergent fiber, higher in proteins, and lower in secondary metabolites (Matsuda et al., 2022). Future anaerobic fungal enrichment attempts from colobine feces should involve multiple substrates, including simple substrates such as glucose and xylan. ITS sequences and shotgun metagenomic DNA from these fungi have been obtained previously from both wild and captive *G. gorilla gorilla* feces. 78 fungal clones representing 12 unique Neocallimastigomycota ASVs, closely related to the hindgut fermenter clade AL3, were isolated from wild gorilla feces (Schulz et al., 2018). They were found to be disproportionately metabolically active in the gut of one captive gorilla, and transcripts were assigned to the genera *Piromyces*, *Pecoramyces*, *Neocallimastix*, and *Anaeromyces* (Houtkamp et al., 2022). This is the first detection of Neocallimastigomycota from captive *C. guereza* feces. All terrestrial foregut-fermenting herbivores whose gut

mycobiota have been surveyed contain Neocallimastigomycota in their gut (Robert J. Gruninger et al., 2014). However, the dependence of anaerobic fungal occupancy on gut transit time, the dependence of transit time on size, and the small size of colobines compared to other established hosts may present a limiting factor to anaerobic fungal residence in the colobine gut (Demment & Soest, 1985; Fonty & Grenet, 1994). Fungi of the Neocallimastigomycota, if actually resident, deploy carbohydrate-active enzymes (CAZymes) in a manner complementary to that of cellulolytic bacteria (Peng et al., 2021). Alternatively, if these fungi are absent from the colobine gut, it is possible that groups of fungi with smaller arsenals of CAZymes that are complementary to the other fungi in the mycobiome also fill a complementary niche in relation to cellulolytic bacteria.

3.4.3. Gastric dysbiosis in captive gorillas

As part of this study, we conducted a small case study comparing the prokaryotic and fungal gut communities of two male gorilla brothers, Bangori and Nzinga, housed in the same enclosure at the Santa Barbara Zoo. Their keepers reported that Nzinga suffered from gastrointestinal dysfunction, which is prevalent among captive herbivorous primates (Amato et al., 2016; Frankel et al., 2019; Lukas, 1999; McKenzie et al., 2017; Strong et al., 2016; Zhu et al., 2018). This is the first recorded comparison of captive gorilla microbiota within a single zoo population to determine The predictive power of this study was limited by small sample size (one gorilla in each treatment group), but the shared housing of the two gorillas, as well as the close genetic relatedness of the two gorillas, allowed us to control for the impact of the external environment on the gorillas' fecal microbiota.

Compared to Bangori's fecal samples, bacteria of the Bacillales, Synergistales, and Coxiellales (Legionellales) were enriched; bacteria of the Bacilli (especially lineage RF39),

the Burkholderiales, Clostridiales, Elusimicrobiales, Eubacteriales, Gastranaerophilales, Izemoplasmatales, and Acholeplasmatales, as well as fungi of the Saccharomycetales and *Aspergillus xerophilus*, were depleted. Lactobacillales and Synergistales were present in Nzinga's feces, but completely absent from Bangori's, and *Lactobacillus gorillae*, the representative of the Lactobacillales present in Nzinga's gut, constituted 5.4-8.0% of prokaryotic relative abundance, whereas Synergistales and Coxiellales constituted less than 1% of relative abundance. *L. gorillae*, as a member of the lactic acid bacteria, produces both D-lactate and L-lactate, and strains from wild gorillas are known to ferment D-xylose, arbutine, cellobiose, and trehalose into lactate more efficiently than those from captive gorillas, and has been proposed as a probiotic for captive gorillas (Tsuchida et al., 2018). Excess lactate does not usually accumulate in the healthy gut due to microbes able to convert it to other acids such as butyrate (D. J. Morrison & Preston, 2016). Among the depleted bacterial taxa, *Clostridium* sensu stricto 1 (Clostridiales) and *Anaerofustis* spp. (Eubacteriales) are the most likely depleted candidates to ferment lactate into other short-chain fatty acids (Rivière et al., 2016; S. Zhang et al., 2022). None of the depleted orders of bacteria were represented by more than a few ASVs: the most representatives belonged to the Gastranaerophilales, with 7 in Bangori's samples. Alpha and beta diversity measures of the fungal communities of Bangori and Nzinga suggest no significant overall difference in their mycobiota, which suggests that it likely does not have a role in the etiology of Nzinga's gastric dysfunction. The significant depletion of an unidentified Saccharomycetales fungus and *A. xerophilus,* considered in the context of an overall lack of depletion of fungi of the Saccharomycetales and of the Aspergillaceae, is unlikely to play a significant role in Nzinga's dysbiosis.

This study represents the first use of anaerobic enrichment on complex substrate to investigate prokaryotic and fungal diversity in non-human primate consortia, and the first attempt to enrich anaerobic fungi of the Neocallimastigomycota from *C. guereza* feces. We have shown that prokaryotic diversity and (especially) fungal diversity are underestimated by rDNA amplicon surveys, that enrichment enhances detection of these rare taxa, and that the rare taxa uncovered by enrichment are likely to contribute critical metabolic redundance and diversity to these lignocellulolytic consortia. As the choice of substrate biases the microbes enriched, future studies should integrate parallel enrichment on a larger panel of substrates, and metatranscriptomes and metametabolomes on each of these substrates may be taken to interrogate the metabolic relationships of these microbes through their gene expression and fermentative output; as described in Chapter 2 and Peng et al. (2021), enrichment is an effective strategy for uncovering microbial partnerships in the herbivore gut ecosystem. Enrichment on these primates' preferred foods will also enable dissection of mechanisms of community assembly. Using methods of community composition analysis such as LSU D1/D2 amplification and full-length rRNA operon sequences, and validating them for use with additional fungal clades, in addition to improving methods of fungal metagenome-assembled genome (MAG) assembly to facilitate shotgun sequencing of the mycobiome and employing lyticase resolution (Pierre et al., 2021) for further improvement of taxonomic resolution, will aid in resolving interactions between bacteria, fungi, and other microbes in the primate gut and understanding the metabolic outcomes and consequences of these interactions for their herbivore host.

4. Conclusions

4.1. Perspectives and future directions

4.1.1. The study of consortia growing on complex substrates enables the study of biomass-degrading strategies emergent from concerted activity by microbial partners in complex consortia

Carbon substrates shape the composition and activity of microbial consortia. Much of our current understanding of how this occurs arises from studies of simple substrates and small, defined consortia. Natural microbiota, however, are larger and more metabolically diverse, and for most of evolutionary history, have fed on complex biomass substrates. The consortia presented in chapter 2 were enriched from inocula taken from the feces of goats, herbivores who, like other ruminants, have consumed chemically complex lignocellulose over the course of their evolutionary history. The microbes in these consortia evolved together to degrade lignocellulose, not simpler substrates. As exemplified in the xylan consortia of chapter 2, simple substrates may enrich smaller and metabolically less-diverse communities relative to consortia enriched on more complex substrates and to the source consortium. Fully understanding multi-strain metabolic strategies from nature that microbiome engineers wish to mimic in synthetic consortia, in order to develop and tune the system in a rational way, thus requires the collection of data from non-engineered consortia consuming substrates they evolved to consume. Studies conducted in ruminants come closest to accomplishing this goal, though as open systems in living organisms they permit the least amount of experimental control. Parallel enrichment on a panel of both simple and complex substrates always results in some degree of strain loss from the inoculum, but enables dissection of specific strategies

for the liberation and metabolism of specific carbohydrates and investigation of the conditions that change them, holistic evaluation of larger-scale strategies employed by larger subsets of the consortium, and observation of the behavior of microbes who occupy specific metabolic niches and who may consequently not be enriched on less complex substrates. We successfully discovered one of these strategies, partitioning carbohydrate-active enzyme labor between high-redundancy and high-diversity groups, by enriching a goat fecal microbiome on three different complex substrates. Even if no enrichment can perfectly recapitulate a source microbiome, enrichment is nonetheless a highly powerful and controllable tool for dissecting the dynamic nature of metabolic cooperation between microbes in a consortium.

A more complete dissection of a consortium's behavioral "program" may be made by resolving the transcriptional networks, signal transduction networks, and metabolite flux of its members. The enrichment consortia whose transcriptional activity on a panel of four biomass substrates was investigated in Chapter 2 were already characterized using metagenomic and metabolomic experimental approaches (Peng et al. 2021). The study presented in Chapter 2 added metatranscriptomic information to this set of analyses, adding one of a few possible links between the genomic repertoire of the members of the consortium and their metabolic output by explaining which genes each organism actually transcribed in response to the substrate stimulus. However, DNA, RNA, protein, metabolite, and similar data in and of themselves represent static snapshots of a system at a particular point in time, with little information by itself about how the consortium actually senses the substrate and mobilizes its metabolic repertoire to process it. Using this information to reconstruct a model of gene transcription, signal transduction, and metabolite flux in the organisms in a

consortium links metabolic repertoire to metabolic activity and then to metabolic outcomes with a mechanistic explanation of not just how the consortium responds to a stimulus, but also how the consortium responds to a variety of other stimuli and how those responses change with the changing environment. How did the consortia growing on alfalfa, bagasse, and reed canary grass marshal their specific repertoires of CAZymes together, and would different CAZymes have dominated with a sufficient change of substrate? Were the seemingly superfluous CAZymes that the xylan consortium mobilized to degrade xylan actually part of the same regulons as the enzymes that degraded xylan? As techniques to model consortia of organisms with genome-scale metabolic models, transcriptional regulation network models, and other model types become more sophisticated, they will demand more computational resources to faithfully model the metabolic complexity of ever larger consortia, so available computational resources must keep pace with the speed of modeling technique development.

4.1.2. Non-human primate gut microbiota link humanity to the evolutionary history of gut microbiota and are core sources of information about primate nutrition and health

A substantial body of research has expanded our understanding of how microbial communities in the human gut influence health and well-being (Clavel et al., 2022). Non-human primate (NHP) gut microbiota link us to the rest of the tree of life and help us understand why humans have the unique microbial communities associated with our species and diverse individual lifestyles. Our gut microbiota evolved alongside us while we diverged from our common ancestors with other great apes and became modern humans. Studies of gut microbiota in non-human primates are not only worthwhile through an anthropocentric perspective, but also through one that centers NHP species themselves – often vulnerable or

endangered, and frequently important forces that structure ecosystems in which they live (Chapman et al., 2013).

Outside of humans, primate gut mycobiota are poorly explored. The proportion of fungi unidentified at the phylum level in the primate gut may be high in an individual study; Barelli *et al*. failed to identify approximately 30% of the fungi at the phylum level in their samples from red colobus (*Procolobus gordonii*) (2020). In our investigation of *Colobus guereza* and *Gorilla gorilla gorilla* gut mycobiota, most fungal ASVs were unidentified at the phylum level. In addition, many of the identifiable fungal ASVs in enrichment consortia were previously undetected in studies of *C. guereza* and *G. gorilla gorilla* mycobiota, both captive and wild. Available data on the disparity between named fungi and unnamed environmental ASVs suggests that many of these unknown fungi are disproportionately likely to represent early-diverging fungi, defined as all fungi outside the clade Dikarya (Lazarus & James, 2015; Picard, 2017; Tedersoo et al., 2017, 2020). A renewed focus by the research community on NHP gut mycobiota as core community members in digestion and health, particularly in herbivorous species, is a critical step towards gaining a full understanding of primate microbiome function. The proportion of fungi identified in these samples, however, will be dependent on the rate of identification and characterization of the fungi associated with the taxonomic barcoding sequences recovered in primate microbiota, particularly with respect to more cryptic and/or difficult-to-culture isolates among early-diverging fungal clades. Enrichment, with the help of antibacterials and antiprokaryotics such as penicillin/streptomycin and chloramphenicol, on a variety of simple and complex substrates, is a critical tool to enable the detection of rare fungal taxa, which, with the rest of the mycobiome, are otherwise easily outcompeted by the much larger and more quickly

reproducing bacterial community in antibiotic-free batch culture (see Swift et al., 2021 for an example in which the bacterium *Fibrobacter* succinogenes UWB7 was cocultured with and outcompeted *Anaeromyces robustus* and *Caecomyces churrovis*, the latter of which was detected in *C. guereza* feces).

Selection and proper amplification of taxonomic barcoding regions that accurately and fully characterize the mycobiome, along with integrating shotgun metagenome sequencing where possible, is critical to improving surveys of the primate mycobiome. The primer sets used to amplify ITS1 and ITS2 regions from fungi often do not amplify these regions from many clades of fungi. A survey of 16 studies of primate gut microbiota from 2014 to the present which included fungal ITS amplicon sequence variant analyses indicated that most selected at least one primer known to fail in amplification of a portion of the fungal kingdom (**Appendix 3, Supplemental Table 5**) (Bellemain et al., 2010; Op De Beeck et al., 2014; Reynolds et al., 2022; Tedersoo et al., 2015; Tedersoo and Lindahl, 2016). Of the studies from this sample, 4 employed the primer set ITS1F/ITS2 to sequence the ITS1 region. The ITS1F/ITS2 primer pair is a popular selection for ITS1 surveys, but does not amplify ITS1 sequences from many fungal clades, including several clades of early-diverging fungi, and is strongly biased toward Dikarya ((Bellemain et al., 2010; Gardes & Bruns, 1993; Tedersoo et al., 2015)Tedersoo and Lindahl, 2016). We employed the primer set gITS7ngs/ITS9 (forward) and ITS4ngs (reverse), which Tedersoo and Lindahl endorsed as an optimized primer set covering the ITS2 region; gITS7ngs is a shortened version of the gITS7 primer with a similar melting temperature and further improvements in coverage of fungal taxa (2016). ITS9 has previously been employed by the US Department of Energy Joint Genome Institute for fungal ITS2 amplicon (“iTag”) sequencing services, and has been shown to

amplify ITS2 sequences from Neocallimastigomycota (Peng et al., 2021). The use of the D1/D2 region of the 28S large ribosomal subunit rDNA to classify fungi has been proposed as an alternative to use of the ITS region, and may aid in further resolution of unidentified fungi in ITS surveys. The 28S D1/D2 region is more uniform in length than the ITS1, and due to its lower degree of within-strain polymorphism in both length and sequence, requires substantially less manual curation for efficient taxonomic diagnosis (Callaghan et al., 2015; Edwards et al., 2019; Koetschan et al., 2014). Use of the 28S D1/D2 has been validated in the Ascomycota, Basidiomycota, and Neocallimastigomycota (Fell et al., 2000; Gade et al., 2017; Hanafy et al., 2020; Kwiatkowski et al., 2012; Scorzetti et al., 2002; Vu et al., 2019). Hanafy *et al.*'s work to develop the use of this region in the Neocallimastigomycota, and their subsequent discovery of several candidate genera, suggests that similar investigation and development of the use of this region for taxonomic diagnosis of other fungal taxa may uncover additional diversity among fungi. Validation of the 28S D1/D2 region and development of sequence databases for other fungal taxa in turn enables the re-analysis of existing mycobiome composition studies. Development of primers to amplify the 28S D1/D2 region in other fungal taxa would likely facilitate future amplicon sequence variant analyses of mycobiota. The primary drawback to the use of the 28S D1/D2 as a taxonomic barcode is the cost of sequencing; Hanafy et al.'s sequencing of this region in the Neocallimastigomycota found that it ranged in size from 740 to 767 bp, with a median length of 760 bp, which requires the use of long-read sequencing technologies (2020).

The opportunity to compare the fecal microbiota of two genetically closely related gorillas housed in the same enclosure allowed us to conduct a small case study of dysbiosis in captive gorillas. We discovered differences in taxonomic makeup and relative abundance

in bacterial populations between the brothers Nzinga, a gorilla suffering from gastric irritation, and Bangori, a healthy gorilla. These differences were consistent with other experimental results indicating that captivity humanizes primate microbiota and creates vulnerabilities to gastric illness (Amato et al., 2016; Frankel et al., 2019; Lukas, 1999; McKenzie et al., 2017; Strong et al., 2016; Zhu et al., 2018). Successful treatment and prevention of gastric dysbiosis among herbivorous non-human primates in captivity, as well as successful habitat remediation for wild populations, depends on understanding its etiology and contributing factors, its impact on the microbial community and its metabolism (which in turn necessitates understanding the community and its functions in its healthy state), and how different interventions alter the microbiome.

4.2. Overall conclusions

The degradation of lignocellulosic biomass by herbivore gut consortia is a complex, efficient, dynamic multi-stage process driven by metabolic interactions between hundreds of microbial taxa across multiple kingdoms of organisms, which collectively transform chemically complex biomass into a diverse array of small-molecule metabolites. We employed enrichment on complex substrates to dissect this cooperation through the lens of community structure and activity, and to understand how this cooperation is shaped by external forces and by the consortium's membership and genomic makeup.

The substrate on which a microbiome feeds shapes its composition and activity. To dissect how anaerobic lignocellulolytic consortia leverage their interdependence to degrade a variety of biomass substrates and how these substrates shape this process, we studied the metatranscriptome a microbiome derived from goat feces produced when it was challenged with three different lignocellulosic substrates and purified hemicellulose. We were able to

discern common broad patterns of activity across multiple substrates, as well as the extent to which substrate shaped this microbiome. We discovered that, even though substrate shaped the exact repertoire of carbohydrate-active enzymes (CAZymes) expressed as well as the degree to which different metabolic pathways were utilized, the cellulolytic bacteria in these consortia divided their labor into a two-portioned strategy combining more focused expression of a smaller set of CAZymes by most of the consortium with more broad-spectrum expression of a large set of CAZymes by a set of four or five MAGs that exhibited relatively consistent membership across lignocellulosic substrates. This highlighted how cellulolytic bacteria in these microbiota collectively utilize strategies to degrade complex lignocellulose substrates that persist across multiple substrates.

We then sought to investigate the role of the composition of herbivorous non-human primate fecal microbiota in lignocellulolysis and to investigate, in particular, the role of the mycobiome. We enriched fecal samples from the Eastern black and white colobus (*Colobus guereza*) and the Western lowland gorilla (*Gorilla gorilla gorilla*) on alfalfa and reed canary grass in an attempt to enrich anaerobic fungi of the Neocallimastigomycota. We failed to enrich members of the Neocallimastigomycota (and the one member of this clade we identified to the species level was present in fecal samples, but not enrichment cultures). However, we enriched a diverse community of microbial taxa, particularly fungi, not detectable in fecal samples, with roles in lignocellulolysis. To the best of our knowledge, this is the most detailed description of both the *C. guereza* and *G. gorilla gorilla* gut mycobiota. In addition, we conducted the first small case study of gut microbiome differences between two genetically closely related male *G. gorilla gorilla* kept in the same enclosure and fed similar diets, one of whom was afflicted with gastrointestinal health issues prevalent among

captive gorillas (Frankel et al., 2019; Popovich et al., 1997; Scott & Keymer, 1975; Strong et al., 2016).

The complex ecosystem of the herbivore gut microbiome is a rich source of strains, transporters, metabolites, enzymes, metabolic pathways, and regulatory mechanisms that have evolved from the need for a community of organisms to work together to degrade biomass. This incredible capacity to efficiently degrade biomass with such a diverse degradative arsenal has made the herbivore gut microbiome an attractive target for biotechnological development. We have employed the technique of enriching consortia on complex substrates to challenge a microbiome and assess its responses, as well as to survey rare taxa in lignocellulolytic mycobiota of the non-human primate gut in order to investigate the lignocellulolytic capacity of these microbiota and, in so doing, to further probe the microbiological substrate of herbivory in our closest relatives. This work stretches across and helps link together a variety of disciplines interested in exploring the herbivore gut microbiome for conservation, therapeutic, and biotechnological applications. Understanding a consortium's degradative "program" through reconstructing its metabolic, transcriptional, and signaling networks is a necessary next step from this point to generate a predictive model of how a consortium will behave in response to a stimulus, and a full picture of how these "programs" are enacted during lignocellulolysis will help generate a more nuanced look at how these microbiota shape, and are shaped by, their environments and evolutionary history.

References

Airgas USA LLC. (2018). *Carbon Dioxide, SDS No. 001013*. https://www.airgas.com/msds/001013.pdf

Allison, G. G., Morris, C., Lister, S. J., Barraclough, T., Yates, N., Shield, I., & Donnison, I. S. (2012). Effect of nitrogen fertiliser application on cell wall composition in switchgrass and reed canary grass. Biomass and Bioenergy, 40, 19–26. https://doi.org/10.1016/j.biombioe.2012.01.034

Alokika, Anu, Kumar, A., Kumar, V., & Singh, B. (2021). Cellulosic and hemicellulosic fractions of sugarcane bagasse: Potential, challenges and future perspective. International Journal of Biological Macromolecules, 169, 564–582. https://doi.org/10.1016/j.ijbiomac.2020.12.175

Amato, K. R. (2019). Missing Links: the Role of Primates in Understanding the Human Microbiome. *MSystems*, *4*(3), 1–4. https://doi.org/10.1128/msystems.00165-19

Amato, K. R., Clayton, J. B., & Hale, V. L. (2020). Colobine Gut Microbiota: New Perspectives on the Nutrition and Health of a Specialized Subfamily of Primates. In I. Matsuda, C. C. Grueter, & J. A. Teichroeb (Eds.), *The Colobines: Natural History, Behaviour, and Ecological Diversity*. Cambridge.

Amato, K. R., G. Sanders, J., Song, S. J., Nute, M., Metcalf, J. L., Thompson, L. R., Morton, J. T., Amir, A., J. McKenzie, V., Humphrey, G., Gogul, G., Gaffney, J., L. Baden, A., A.O. Britton, G., P. Cuozzo, F., Di Fiore, A., J. Dominy, N., L. Goldberg, T., Gomez, A., … R. Leigh, S. (2019). Evolutionary trends in host physiology outweigh dietary niche in structuring primate gut microbiomes. *ISME Journal*, *13*(3), 576–587. https://doi.org/10.1038/s41396-018-0175-0

Amato, K. R., Metcalf, J. L., Song, S. J., Hale, V. L., Clayton, J., Ackermann, G., Humphrey, G., Niu, K., Cui, D., Zhao, H., Schrenzel, M. D., Tan, C. L., Knight, R., & Braun, J. (2016). Using the gut microbiota as a novel tool for examining colobine primate GI health. *Global Ecology and Conservation*, *7*, 225–237. https://doi.org/10.1016/j.gecco.2016.06.004

Amato, K. R., Yeoman, C. J., Kent, A., Righini, N., Carbonero, F., Estrada, A., Rex Gaskins, H., Stumpf, R. M., Yildirim, S., Torralba, M., Gillis, M., Wilson, B. A., Nelson, K. E., White, B. A., & Leigh, S. R. (2013). Habitat degradation impacts black howler monkey (Alouatta pigra) gastrointestinal microbiomes. *ISME Journal*, *7*(7), 1344–1353. https://doi.org/10.1038/ismej.2013.16

Angelidaki, I., Karakashev, D., Batstone, D. J., Plugge, C. M., & Stams, A. J. M. (2011). Biomethanation and its potential. In *Methods in Enzymology* (1st ed., Vol. 494). Elsevier Inc. https://doi.org/10.1016/B978-0-12-385112-3.00016-0

Archibald, J. G., Barnes, H. D., Fenner, H., & Gersten, B. (1962). Digestibility of Alfalfa Hay and Reed Canary Grass Hay Measured by Two Procedures. Journal of Dairy Science, 45(7), 858–860. https://doi.org/10.3168/jds.S0022-0302(62)89510-1

Aung, H. W., Henry, S. A., & Walker, L. P. (2013). Revising the Representation of Fatty Acid, Glycerolipid, and Glycerophospholipid Metabolism in the Consensus Model of Yeast Metabolism. *Industrial Biotechnology*, *9*(4), 215–228. https://doi.org/10.1089/ind.2013.0013

Azad, E., Fehr, K. B., Derakhshani, H., Forster, R., Acharya, S., Khafipour, E., McGeough, E., & McAllister, T. A. (2020). Interrelationships of fiber-associated anaerobic fungi and bacterial communities in the rumen of bloated cattle grazing alfalfa. Microorganisms, 8(10), 1–17. https://doi.org/10.3390/microorganisms8101543

Balch, W. E., & Wolfe, R. S. (1976). New approach to the cultivation of methanogenic bacteria: 2-mercaptoethanesulfonic acid (HS-CoM) dependent growth of Methanobacterium ruminantium in a pressurized atmosphere. *Applied and Environmental Microbiology*, *32*(6), 781–791. https://doi.org/10.1128/aem.32.6.781-791.1976

Barelli, C., Albanese, D., Stumpf, R. M., Asangba, A., Donati, C., Rovero, F., & Hauffe, H. C. (2020a). The Gut Microbiota Communities of Wild Arboreal and Ground-Feeding Tropical Primates Are Affected Differently by Habitat Disturbance. *MSystems*, *5*(3), e00061-20.

Barelli, C., Albanese, D., Stumpf, R. M., Asangba, A., Donati, C., Rovero, F., & Hauffe, H. C. (2020b). The Gut Microbiota Communities of Wild Arboreal and Ground-Feeding Tropical Primates Are Affected Differently by Habitat Disturbance. *MSystems*, *5*(3), e00061-20. https://doi.org/10.1128/mSystems.00061-20

Basile, A., Campanaro, S., Kovalovszki, A., Zampieri, G., Rossi, A., Angelidaki, I., Valle, G., & Treu, L. (2020). Revealing metabolic mechanisms of interaction in the anaerobic digestion microbiome by flux balance analysis. *Metabolic Engineering*, *62*(August), 138–149. https://doi.org/10.1016/j.ymben.2020.08.013

Bauchop, T., & Martucci, R. W. (1968). Ruminant-like digestion of the langur monkey. *Science*, *161*(3842), 698–699. https://doi.org/10.1126/science.161.3842.698

Bauchop, T., & Mountfort, D. O. (1981). Cellulose fermentation by a rumen anaerobic fungus in both the absence and the presence of rumen methanogens. *Applied and*

Environmental Microbiology, *42*(6), 1103–1110. https://doi.org/10.1128/aem.42.6.1103-1110.1981

Bayané, A., & Guiot, S. R. (2011). Animal digestive strategies versus anaerobic digestion bioprocesses for biogas production from lignocellulosic biomass. *Reviews in Environmental Science and Biotechnology*, *10*(1), 43–62. https://doi.org/10.1007/s11157-010-9209-4

Beier, S., & Bertilsson, S. (2013). Bacterial chitin degradation-mechanisms and ecophysiological strategies. *Frontiers in Microbiology*, *4*(JUN), 1–12. https://doi.org/10.3389/fmicb.2013.00149

Belanche, A., Doreau, M., Edwards, J. E., Moorby, J. M., Pinloche, E., & Newbold, C. J. (2012). Shifts in the rumen microbiota due to the type of carbohydrate and level of protein ingested by dairy cattle are associated with changes in rumen fermentation. *Journal of Nutrition*, *142*(9), 1684–1692. https://doi.org/10.3945/jn.112.159574

Bellemain, E., Carlsen, T., Brochmann, C., Coissac, E., Taberlet, P., & Kauserud, H. (2010). ITS as an environmental DNA barcode for fungi: An in silico approach reveals potential PCR biases. *BMC Microbiology*, *10*, 1–9. https://doi.org/10.1186/1471-2180-10-189

Bellmann, S., Carlander, D., Fasano, A., Momcilovic, D., Scimeca, J. A., Waldman, W. J., Gombau, L., Tsytsikova, L., Canady, R., Pereira, D. I. A., & Lefebvre, D. E. (2015). Mammalian gastrointestinal tract parameters modulating the integrity, surface properties, and absorption of food-relevant nanomaterials. Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology, 7(5), 609–622. https://doi.org/10.1002/wnan.1333

Bharti, R., & Grimm, D. G. (2021). Current challenges and best-practice protocols for microbiome analysis. *Briefings in Bioinformatics*, *22*(1), 178–193. https://doi.org/10.1093/bib/bbz155

Bi, X., Liu, Y., Li, J., Du, G., Lv, X., & Liu, L. (2022). *Construction of Multiscale Genome-Scale Metabolic Models : Frameworks and Challenges*. 1–21.

Biddle, A., Stewart, L., Blanchard, J., & Leschine, S. (2013). Untangling the genetic basis of fibrolytic specialization by lachnospiraceae and ruminococcaceae in diverse gut communities. *Diversity*, *5*(3), 627–640. https://doi.org/10.3390/d5030627

Bilal, M., Vilar, D. S., Eguiluz, K. I. B., Ferreira, L. F. R., Bhatt, P., & Iqbal, H. M. N. (2021). Biochemical conversion of lignocellulosic waste into renewable energy. In *Advanced Technology for the Conversion of Waste into Fuels and Chemicals: Volume 1: Biological Processes* (First Edit, Vol. 1). Elsevier Ltd. https://doi.org/10.1016/B978-0-12-823139-5.00007-1

Biosafety in Microbiological and Biomedical Laboratories (BMBL) 6th Edition. (2020). In P. J. Meechan & J. Potts (Eds.), *Centers for Disease Control and Prevention, National Institutes of Health* (6th ed.). https://www.cdc.gov/labs/BMBL.html%0Ahttps://www.cdc.gov/labs/BMBL.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fbiosafety%2Fpublications%2Fbmbl5%2Findex.htm

Blair, E. M., Dickson, K. L., & O'Malley, M. A. (2021). Microbial communities and their enzymes facilitate degradation of recalcitrant polymers in anaerobic digestion. *Current Opinion in Microbiology*, *64*, 100–108. https://doi.org/10.1016/j.mib.2021.09.008

Blanco, C., Salomón, O., & Raga, J. A. (2001). Diet of the bottlenose dolphin (Tursiops truncatus) in the western Mediterranean Sea. *Journal of the Marine Biological Association of the United Kingdom*, *81*(6), 1053–1058. https://doi.org/10.1017/S0025315401005057

Blazeck, J., & Alper, H. (2010). Systems metabolic engineering : Genome-scale models and beyond. *Biotechnology Journal*, *5*, 647–659. https://doi.org/10.1002/biot.200900247

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bordenstein, S. R., & Theis, K. R. (2015). Host biology in light of the microbiome: Ten principles of holobionts and hologenomes. *PLoS Biology*, *13*(8), 1–23. https://doi.org/10.1371/journal.pbio.1002226

Borruso, L., Checcucci, A., Torti, V., Correa, F., Sandri, C., Luise, D., Cavani, L., Modesto, M., Spiezio, C., Mimmo, T., Cesco, S., Vito, M. Di, Bugli, F., Randrianarison, R. M., Gamba, M., Rarojoson, N. J., Zaborra, C. A., Mattarelli, P., Trevisi, P., & Giacoma, C. (2021). *I Like the Way You Eat It : Lemur (Indri indri) Gut Mycobiome and Geophagy*. 215–223.

Bradley, S. G., Antalis, T. M., & Bond, J. S. (2013). Proteases in the Mammalian Digestive System. In K. Brix & W. Stocker (Eds.), Proteases: Structure and Function (1st ed., pp. 373–393). Springer Vienna.

Brown, J. L., Perisin, M. A., Swift, C. L., Benyamin, M., Liu, S., Singan, V., Zhang, Y., Savage, E., Pennacchio, C., Grigoriev, I. V., & O'Malley, M. A. (2022). Co-cultivation of anaerobic fungi with Clostridium acetobutylicum bolsters butyrate and butanol

production from cellulose and lignocellulose. Journal of Industrial Microbiology and Biotechnology, 49(6), kuac024. https://doi.org/10.1093/jimb/kuac024

Brown, J. L., Swift, C. L., Mondo, S. J., Seppala, S., Salamov, A., Singan, V., Henrissat, B., Drula, E., Henske, J. K., Lee, S., Labutti, K., He, G., Yan, M., Barry, K., Grigoriev, I. V, & Malley, M. A. O. (2021). Co-cultivation of the anaerobic fungus Caecomyces churrovis with Methanobacterium bryantii enhances transcription of carbohydrate binding modules , dockerins, and pyruvate formate lyases on specific substrates. *Biotechnology for Biofuels*, 1–16. https://doi.org/10.1186/s13068-021-02083-w

Bryant, M. P. (1972). Commentary on the Hungate technique for culture of anaerobic bacteria. *The American Journal of Clinical Nutrition*, *25*(12), 1324–1328. https://doi.org/10.1093/ajcn/25.12.1324

Bushnell, B., Rood, J., & Singer, E. (2017). BBMerge – Accurate paired shotgun read merging via overlap. *PLoS ONE*, *12*(10), e0185056.

Callaghan, T. M., Podmirseg, S. M., Hohlweck, D., Edwards, J. E., Puniya, A. K., Dagar, S. S., & Griffith, G. W. (2015). Buwchfawromyces eastonii gen. nov., sp. nov.: A new anaerobic fungus (Neocallimastigomycota) isolated from buffalo faeces. *MycoKeys*, *9*, 11–28. https://doi.org/10.3897/mycokeys.9.9032

Canarini, A., Schmidt, H., Fuchslueger, L., Martin, V., Herbold, C. W., Zezula, D., Gündler, P., Hasibeder, R., Jecmenica, M., Bahn, M., & Richter, A. (2021). Ecological memory of recurrent drought modifies soil processes via changes in soil microbial community. *Nature Communications*, *12*(1), 1–14. https://doi.org/10.1038/s41467-021-25675-4

Carlos, C., Fan, H., Currie, C. R., Nicol, G. W., Carlos, C., & Currie, C. R. (2018). *Substrate Shift Reveals Roles for Members of Bacterial Consortia in Degradation of Plant Cell Wall Polymers*. *9*(March), 1–10. https://doi.org/10.3389/fmicb.2018.00364

Carreño, D., Toral, P. G., Pinloche, E., Belenguer, A., Yáñez-Ruiz, D. R., Hervás, G., McEwan, N. R., Newbold, C. J., & Frutos, P. (2019). Rumen bacterial community responses to DPA, EPA and DHA in cattle and sheep: A comparative in vitro study. *Scientific Reports*, *9*(1), 1–9. https://doi.org/10.1038/s41598-019-48294-y

Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C. A., Holland, T. A., Keseler, I. M., Kothari, A., Kubo, A., Krummenacker, M., Latendresse, M., Mueller, L. A., Ong, Q., Paley, S., Subhraveti, P., Weaver, D. S., Weerasinghe, D., Zhang, P., & Karp, P. D. (2014). The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway / Genome Databases. *Nucleic Acids Research*, *42*(2014), D459–D471. https://doi.org/10.1093/nar/gkt1103

Celiker, H., & Gore, J. (2012). Competition between species can stabilize public-goods cooperation within a species. Molecular Systems Biology, 8: 621. https://doi.org/10.1038/msb.2012.54

Chapman, C. A., Bonnell, T. R., Gogarten, J. F., Lambert, J. E., Omeja, P. A., Twinomugisha, D., Wasserman, M. D., & Rothman, J. M. (2013). Are Primates Ecosystem Engineers? *International Journal of Primatology*, *34*(1), 1–14. https://doi.org/10.1007/s10764-012-9645-9

Chappell, T. C., & Nair, N. U. (2017). Co-utilization of hexoses by a microconsortium of sugar-specific E. coli strains. *Biotechnology and Bioengineering*, *114*(10), 2309–2318. https://doi.org/10.1002/bit.26351

Chaucheyras, F., Fonty, G., Bertin, G., & Gouet, P. (1995). Effects of live Saccharomyces cerevisiae cells on zoospore germination, growth, and cellulolytic activity of the rumen anaerobic fungus, Neocallimastix frontalis MCH3. *Current Microbiology*, *31*(4), 201–205. https://doi.org/10.1007/BF00298373

Chen, L., Liu, J., Zhang, Y., Dai, B., An, Y., & Yu, L. (2015a). Structural, thermal, and anti-inflammatory properties of a novel pectic polysaccharide from Alfalfa (*Medicago sativa* L.) stem. Journal of Agricultural and Food Chemistry, 63(12), 3219–3228. https://doi.org/10.1021/acs.jafc.5b00494

Chen, L., Liu, J., Zhang, Y., Niu, Y., Dai, B., & Yu, L. (2015b). A Novel Alkaline Hemicellulosic Heteroxylan Isolated from Alfalfa (*Medicago sativa* L.) Stem and Its Thermal and Anti-inflammatory Properties. Journal of Agricultural and Food Chemistry, 63(11), 2970–2978. https://doi.org/10.1021/acs.jafc.5b00063

Cheng, Y. F., Edwards, J. E., Allison, G. G., Zhu, W. Y., & Theodorou, M. K. (2009). Diversity and activity of enriched ruminal cultures of anaerobic fungi and methanogens grown together on lignocellulose in consecutive batch culture. *Bioresource Technology*, *100*(20), 4821–4828. https://doi.org/10.1016/j.biortech.2009.04.031

Chivers, D. J. (1995). Functional anatomy of the gastrointestinal tract. In G. Davies & J. F. Oates (Eds.), *Colobine Monkeys: Their Ecology, Behaviour, and Evolution* (pp. 205–257).

Clavel, T., Horz, H. P., Segata, N., & Vehreschild, M. (2022). Next steps after 15 stimulating years of human gut microbiome research. *Microbial Biotechnology*, *15*(1), 164–175. https://doi.org/10.1111/1751-7915.13970

Clayton, J. B. (2015). *Associations Between Nutrition, Gut Microbial Communities, and Health in Nonhuman Primates* (Issue December). University of Minnesota.

Clayton, J. B., Gomez, A., Amato, K. R., Knights, D., Travis, D. A., Blekhman, R., Knight, R., Leigh, S., Stumpf, R., Wolf, T., Glander, K. E., Cabana, F., & Johnson, T. J. (2018). The gut microbiome of nonhuman primates: Lessons in ecology and evolution. *American Journal of Primatology*, *80*(6). https://doi.org/10.1002/ajp.22867

Clayton, J. B., Vangay, P., Huang, H., Ward, T., Hillmann, B. M., Al-Ghalith, G. A., Travis, D. A., Long, H. T., Van Tuan, B., Van Minh, V., Cabana, F., Nadler, T., Toddes, B., Murphy, T., Glander, K. E., Johnson, T. J., & Knights, D. (2016). Captivity humanizes the primate microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(37), 10376–10381. https://doi.org/10.1073/pnas.1521835113

Coyte, K. Z., Schluter, J., & Foster, K. R. (2015). The ecology of the microbiome: Networks, competition, and stability. *Science*, *350*(6261), 663–666. https://doi.org/10.1126/science.aad2602

Creevey, C. J., Kelly, W. J., Henderson, G., & Leahy, S. C. (2014). Determining the culturability of the rumen bacterial microbiome. *Microbial Biotechnology*, *7*(5), 467–479. https://doi.org/10.1111/1751-7915.12141

Crerar, M. M., & Rooks, N. E. (1987). The structure and expression of amylase genes in mammals: An overview. *Critical Reviews in Biotechnology*, *5*(3), 217–227. https://doi.org/10.3109/07388558709086980

Cruz, F., Faria, J. P., Rocha, M., Rocha, I., & Dias, O. (2020). A review of methods for the reconstruction and analysis of integrated genome-scale models of metabolism and regulation. *Biochemical Society Transactions*, *48*(5), 1889–1903. https://doi.org/10.1042/BST20190840

D'Souza, G., Shitut, S., Preussger, D., Yousif, G., Waschina, S., & Kost, C. (2018). Ecology and evolution of metabolic cross-feeding interactions in bacteria. Natural Product Reports, 35(5), 455–488. https://doi.org/10.1039/c8np00009c

da Lage, J. L., Maczkowiak, F., & Cariou, M. L. (2011). Phylogenetic distribution of intron positions in alpha-amylase genes of bilateria suggests numerous gains and losses. *PLoS ONE*, *6*(5). https://doi.org/10.1371/journal.pone.0019673

Dai, D. Q., Wijayawardene, N. N., Zhang, G. Q., Gao, Y., Wijayawardene, N. N., Hyde, K. D., Pem, D., Thiyagaraja, V., Hyde, K. D., Dong, W., Sánchez-García, M., Goto, B. T., Saxena, R. K., Erdoğdu, M., Selçuk, F., Rajeshkumar, K. C., Sruthi, O. P., Aptroot, A., Błaszkowski, J., … Thines, M. (2022). Outline of Fungi and fungus-like taxa – 2021. *Mycosphere*, *13*(1), 53–453. https://doi.org/10.5943/mycosphere/13/1/2

Davies, D. R., Theodorou, M. K., Brooks, A. E., & Trinci, A. P. J. (1993). Influence of drying on the survival of anaerobic fungi in rumen digesta and faeces of cattle. *FEMS*

Microbiology Letters, *106*(1), 59–63. https://doi.org/10.1111/j.1574-6968.1993.tb05935.x

de Groot, D. H., Van Boxtel, C., Planqué, R., Bruggeman, F. J., & Teusink, B. (2019). The number of active metabolic pathways is bounded by the number of cellular constraints at maximal metabolic rates. PLoS Computational Biology, 15(3), 1–24. https://doi.org/10.1371/journal.pcbi.1006858

Delgado, L. F., & Andersson, A. F. (2022). Evaluating metagenomic assembly approaches for biome-specific gene catalogues. *Microbiome*, *10*(1), 1–11. https://doi.org/10.1186/s40168-022-01259-2

Demment, M. W., & Soest, P. J. Van. (1985). A Nutritional Explanation for Body-Size Patterns of Ruminant and Nonruminant Herbivores. *The American Naturalist*, *125*(5), 641–672.

Deng, Y., & Wang, S. Y. (2016). Synergistic growth in bacteria depends on substrate complexity. *Journal of Microbiology*, *54*(1), 23–30. https://doi.org/10.1007/s12275-016-5461-9.Synergistic

Deng, Y., & Wang, S. Y. (2017). Complex carbohydrates reduce the frequency of antagonistic interactions among bacteria degrading cellulose and xylan. *FEMS Microbiology Letters*, *364*(5), fnx019. https://doi.org/10.1093/femsle/fnx019

Deter, H. S., & Lu, T. (2022). Engineering microbial consortia with rationally designed cellular interactions. *Current Opinion in Biotechnology*, *76*, 102730. https://doi.org/10.1016/j.copbio.2022.102730

Deublein, D., & Steinhauser, A. (2011). *Biogas from Waste and Renewable Resources* (2nd ed.). Wiley VCH. https://doi.org/10.1002/9783527632794

Doran-Sheehy, D., Mongo, P., Lodwick, J., & Conklin-Brittain, N. L. (2009). Male and female western gorilla diet: Preferred foods, use of fallback resources, and implications for ape versus old world monkey foraging strategies. *American Journal of Physical Anthropology*, *140*(4), 727–738. https://doi.org/10.1002/ajpa.21118

Drula, E., Garron, M., Dogan, S., Lombard, V., Henrissat, B., & Terrapon, N. (2021). The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Research*, *50*(D1), D571–D577.

Druzhinina, I. S., & Kubicek, C. P. (2012). Environmental and Microbial Relationships. In K. Esser (Ed.), *The Mycota: Environmental and Microbial Relationships: Vol. IV* (3rd ed.). Springer International Publishing.

Edwards, J. E., Forster, R. J., Callaghan, T. M., Dollhofer, V., Dagar, S. S., Cheng, Y., Chang, J., Kittelmann, S., Fliegerova, K., Puniya, A. K., Henske, J. K., Gilmore, S. P., O'Malley, M. A., Griffith, G. W., & Smidt, H. (2017). PCR and omics based techniques to study the diversity, ecology and biology of anaerobic fungi: Insights, challenges and opportunities. *Frontiers in Microbiology*, *8*(SEP). https://doi.org/10.3389/fmicb.2017.01657

Edwards, J. E., Hermes, G. D. A., Kittelmann, S., Nijsse, B., & Smidt, H. (2019). Assessment of the Accuracy of High-Throughput Sequencing of the ITS1 Region of Neocallimastigomycota for Community Composition Analysis. *Frontiers in Microbiology*, *10*(October), 1–11. https://doi.org/10.3389/fmicb.2019.02370

Elkhateeb, W. A. (2005). *Some mycological, phytopathological and physiological studies on mycobiota of selected newly reclaimed soils in Assiut Governorate, Egypt*. Assiut University.

Elton, C. S. (1927). Animal Ecology. Macmillan Publishers Limited. http://archive.org/details/animalecology00elto

Estrada, A., Garber, P. A., Rylands, A. B., Roos, C., Fernandez-Duque, E., Fiore, A. Di, Anne-Isola Nekaris, K., Nijman, V., Heymann, E. W., Lambert, J. E., Rovero, F., Barelli, C., Setchell, J. M., Gillespie, T. R., Mittermeier, R. A., Arregoitia, L. V., de Guinea, M., Gouveia, S., Dobrovolski, R., … Li, B. (2017). Impending extinction crisis of the world's primates: Why primates matter. *Science Advances*, *3*(1). https://doi.org/10.1126/sciadv.1600946

Estrela, S., Sánchez, Á., & Rebolleda-Gómez, M. (2021). Multi-Replicated Enrichment Communities as a Model System in Microbial Ecology. *Frontiers in Microbiology*, *12*, 657467. https://doi.org/10.3389/fmicb.2021.657467

Evans, P. N., Boyd, J. A., Leu, A. O., Woodcroft, B. J., Parks, D. H., Hugenholtz, P., & Tyson, G. W. (2019). An evolving view of methane metabolism in the Archaea. *Nature Reviews Microbiology*, *17*(4), 219–232. https://doi.org/10.1038/s41579-018-0136-7

Fell, J. W., Boekhout, T., Fonseca, A., Scorzetti, G., & Statzell-tallman, A. (2000). Biodiversity and systematics of basidiomycetous yeasts as determined by large-subunit rDNA D1/D2 domain sequence analysis. *International Journal of Systematic and Evolutionary Microbiology*, *50*, 1351–1371.

Feng, Y., Yu, Y., Wang, X., Qu, Y., Li, D., He, W., & Kim, B. H. (2011). Degradation of raw corn stover powder (RCSP) by an enriched microbial consortium and its community structure. *Bioresource Technology*, *102*(2), 742–747. https://doi.org/10.1016/j.biortech.2010.08.074

Feng, Z., Alén, R., & Pakkanen, H. (2002). Characterization of black liquors from soda-AQ pulping of reed canary grass (*Phalaris arundinacea* L.). Holzforschung, 56(3), 298–303. https://doi.org/10.1515/HF.2002.048

Filho, F. M., & Paula, A. (2021). A Systematic Strategy to Find Potential Therapeutic Targets for Pseudomonas aeruginosa Using Integrated Computational Models. *Frontiers in Molecular Biosciences*, *8*, 728129. https://doi.org/10.3389/fmolb.2021.728129

Fliegerova, K. O., Podmirseg, S. M., Vinzelj, J., Grilli, D. J., Kvasnová, S., Schierová, D., Sechovcová, H., Mrázek, J., Siddi, G., Arenas, G. N., & Moniello, G. (2021). The effect of a high-grain diet on the rumen microbiome of goats with a special focus on anaerobic fungi. *Microorganisms*, *9*, 157. https://doi.org/10.3390/microorganisms9010157

Fonty, G., & Grenet, E. (1994). Effects of Diet on the Fungal Population of the Digestive Tract of Ruminants. In D. O. Mountfort & C. G. Orpin (Eds.), *Anaerobic Fungi: Biology, Ecology, and Function, vol. 12* (pp. 229–239). Marcel Dekker, Inc.

Ford, J. E., Perry, K. D., & Briggs, C. A. (1958). Nutrition of lactic acid bacteria isolated from the rumen. *Journal of General Microbiology*, *18*(1), 273–284. https://doi.org/10.1099/00221287-18-1-273

Forsythe, S. J., & Parker, D. S. (1985). Nitrogen metabolism by the microbial flora of the rabbit caecum. Journal of Applied Bacteriology, 58(4), 363–369. https://doi.org/10.1111/j.1365-2672.1985.tb01475.x

Fouhse, J. M., Smiegielski, L., Tuplin, M., Guan, L. L., & Willing, B. P. (2017). Host immune selection of rumen bacteria through salivary secretory IgA. *Frontiers in Microbiology*, *8*(MAY), 1–9. https://doi.org/10.3389/fmicb.2017.00848

Frankel, J. S., Mallott, E. K., Hopper, L. M., Ross, S. R., & Amato, K. R. (2019). The effect of captivity on the primate gut microbiome varies with host dietary niche. *American Journal of Primatology*, *81*(12), 1–9. https://doi.org/10.1002/ajp.23061

Gade, L., Hurst, S., Balajee, S. A., Lockhart, S. R., & Litvintseva, A. P. (2017). Detection of mucormycetes and other pathogenic fungi in formalin fixed paraffin embedded and fresh tissues using the extended region of 28S rDNA. *Medical Mycology*, *55*(4), 385–395. https://doi.org/10.1093/mmy/myw083

GARDES, M., & BRUNS, T. D. (1993). ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. *Molecular Ecology*, *2*(2), 113–118. https://doi.org/10.1111/j.1365-294X.1993.tb00005.x

Garza, D. R., Van Verk, M. C., Huynen, M. A., & Dutilh, B. E. (2018). Towards predicting the environmental metabolome from metagenomics with a mechanistic model. *Nature Microbiology*, *3*(4), 456–460. https://doi.org/10.1038/s41564-018-0124-8

Gatesy, J., Geisler, J. H., Chang, J., Buell, C., Berta, A., Meredith, R. W., Springer, M. S., & McGowen, M. R. (2013). A phylogenetic blueprint for a modern whale. *Molecular Phylogenetics and Evolution*, *66*(2), 479–506. https://doi.org/10.1016/j.ympev.2012.10.012

Gilmore, S. P., Henske, J. K., & O'Malley, M. A. (2015). Driving biomass breakdown through engineered cellulosomes. *Bioengineered*, *6*(4), 204–208. https://doi.org/10.1080/21655979.2015.1060379

Gilmore, S. P., Lankiewicz, T. S., Wilken, S. E., Brown, J. L., Sexton, J. A., Henske, J. K., Theodorou, M. K., Valentine, D. L., & Malley, M. A. O. (2019). Top-Down Enrichment Guides in Formation of Synthetic Microbial Consortia for Biomass Degradation. *ACS Synthetic Biology*, *8*(9), 2174–2185. https://doi.org/10.1021/acssynbio.9b00271

Giri, S., Oña, L., Waschina, S., Shitut, S., Yousif, G., Kaleta, C., & Kost, C. (2021). Metabolic dissimilarity determines the establishment of cross-feeding interactions in bacteria. *Current Biology*, *31*(24), 5547-5557.e6. https://doi.org/10.1016/j.cub.2021.10.019

Giri, S., Waschina, S., Kaleta, C., & Kost, C. (2019). Defining Division of Labor in Microbial Communities. *Journal of Molecular Biology*, *431*(23), 4712–4731. https://doi.org/10.1016/j.jmb.2019.06.023

Gleason, F. H., Letcher, P. M., & McGee, P. a. (2007). Some aerobic Blastocladiomycota and Chytridiomycota can survive but cannot grow under anaerobic conditions. Australasian Mycologist, 26(2–3), 57–64.

Gleason, F. H., Kagami, M., Lefevre, E., & Sime-Ngando, T. (2008). The ecology of chytrids in aquatic ecosystems: roles in food web dynamics. *Fungal Biology Reviews*, *22*(1), 17–25. https://doi.org/10.1016/j.fbr.2008.02.001

Glöckner, F. O., Yilmaz, P., Quast, C., Gerken, J., Beccati, A., Ciuprina, A., Bruns, G., Yarza, P., Peplies, J., Westram, R., & Ludwig, W. (2017). 25 years of serving the community with ribosomal RNA gene reference databases and tools. *Journal of Biotechnology*, *261*(June), 169–176. https://doi.org/10.1016/j.jbiotec.2017.06.1198

Godon, J. J., Arcemisbéhère, L., Escudié, R., Harmand, J., Miambi, E., & Steyer, J. P. (2013). Overview of the Oldest Existing Set of Substrate-optimized Anaerobic Processes: Digestive Tracts. *Bioenergy Research*, *6*(3), 1063–1081. https://doi.org/10.1007/s12155-013-9339-y

Godon, J. J., Arulazhagan, P., Steyer, J. P., & Hamelin, J. (2016). Vertebrate bacterial gut diversity: Size also matters. *BMC Ecology*, *16*(1), 1–9. https://doi.org/10.1186/s12898-016-0071-2

Grajal, A., Park, W., & Seattle, Z. (1989). Foregut Fermentation in the Hoatzin , a Neotropical Leaf-Eating Bird. *Science*, *245*(4923), 1236–1238. https://doi.org/10.1126/science.245.4923.1236

Graur, D., & Higgins, D. G. (1994). Molecular evidence for the inclusion of cetaceans within the order Artiodactyla. *Molecular Biology and Evolution*, *11*(3), 357–364. https://doi.org/10.1093/oxfordjournals.molbev.a040118

Greene, L. K., McKenney, E. A., O'Connell, T. M., & Drea, C. M. (2018). The critical role of dietary foliage in maintaining the gut microbiome and metabolome of folivorous sifakas. *Scientific Reports*, *8*(1), 1–13. https://doi.org/10.1038/s41598-018-32759-7

Greub, G. (2012). Culturomics: A new approach to study the human microbiome. *Clinical Microbiology and Infection*, *18*(12), 1157–1159. https://doi.org/10.1111/1469-0691.12032

Groussin, M., Mazel, F., & Alm, E. J. (2020). Co-evolution and Co-speciation of Host-Gut Bacteria Systems. *Cell Host and Microbe*, *28*(1), 12–22. https://doi.org/10.1016/j.chom.2020.06.013

Gruninger, R. J., Ribeiro, G. O., Cameron, A., & McAllister, T. A. (2019). Invited review: Application of meta-omics to understand the dynamic nature of the rumen microbiome and how it responds to diet in ruminants. *Animal*, *13*(9), 1843–1854. https://doi.org/10.1017/S1751731119000752

Gruninger, Robert J., Puniya, A. K., Callaghan, T. M., Edwards, J. E., Youssef, N., Dagar, S. S., Fliegerova, K., Griffith, G. W., Forster, R., Tsang, A., Mcallister, T., & Elshahed, M. S. (2014). Anaerobic fungi (phylum Neocallimastigomycota): Advances in understanding their taxonomy, life cycle, ecology, role and biotechnological potential. *FEMS Microbiology Ecology*, *90*(1), 1–17. https://doi.org/10.1111/1574-6941.12383

Guo, W., Mishra, S., Wang, C., Zhang, H., & Ning, R. (2019). Comparative Study of Gut Microbiota in Wild and Captive Giant Pandas (Ailuropoda melanoleuca). *Genes*, *10*, 1–17.

Gupta, R. S., Nanda, A., & Khadka, B. (2017). Novel molecular, structural and evolutionary characteristics of the phosphoketolases from bifidobacteria and Coriobacteriales. *PLoS ONE*, *12*(2), 1–20. https://doi.org/10.1371/journal.pone.0172176

Haitjema, C. H., Gilmore, S. P., Henske, J. K., Solomon, K. V., De Groot, R., Kuo, A., Mondo, S. J., Salamov, A. A., LaButti, K., Zhao, Z., Chiniquy, J., Barry, K., Brewer, H. M., Purvine, S. O., Wright, A. T., Hainaut, M., Boxma, B., Van Alen, T., Hackstein, J. H. P., … O'Malley, M. A. (2017). A parts list for fungal cellulosomes revealed by

comparative genomics. *Nature Microbiology*, *2*(May), 1–8. https://doi.org/10.1038/nmicrobiol.2017.87

Haitjema, C. H., Solomon, K. V., Henske, J. K., Theodorou, M. K., & O'Malley, M. A. (2014). Anaerobic gut fungi: Advances in isolation, culture, and cellulolytic enzyme discovery for biofuel production. *Biotechnology and Bioengineering*, *111*(8), 1471–1482. https://doi.org/10.1002/bit.25264

Hale, V. L., Tan, C. L., Niu, K., Yang, Y., Knight, R., Zhang, Q., Cui, D., & Amato, K. R. (2018). Diet Versus Phylogeny: a Comparison of Gut Microbiota in Captive Colobine Monkey Species. *Microbial Ecology*, *75*(2), 1–13. https://doi.org/10.1007/s00248-017-1041-8

Hanafy, R. A., Johnson, B., Youssef, N. H., & Elshahed, M. S. (2020). Assessing anaerobic gut fungal diversity in herbivores using D1/D2 large ribosomal subunit sequencing and multi-year isolation. *Environmental Microbiology*, *00*, 1–26. https://doi.org/10.1111/1462-2920.15164

Haque, M. N. (2018). Dietary manipulation: A sustainable way to mitigate methane emissions from ruminants. *Journal of Animal Science and Technology*, *60*, 15. https://doi.org/10.1186/s40781-018-0175-7

Harper, E. G., Welch, R. P., Contreras Lara, D., Chang, J. S., & Calza, R. E. (1996). The effect of Aspergillus oryzae fermentation extract on the anaerobic fungi Neocallimastix frontalis EB 188, Piromyces communis DC 193 and Orpinomyces ssp. RW 206: Generalized effects and component analysis. *Applied Microbiology and Biotechnology*, *45*(6), 817–821. https://doi.org/10.1007/s002530050768

Hecker, N., Sharma, V., & Hiller, M. (2019). Convergent gene losses illuminate metabolic and physiological changes in herbivores and carnivores. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(8), 3036–3041. https://doi.org/10.1073/pnas.1818504116

Henrissat, B., Claeyssens, M., Tomme, P., Lemesle, L., & Mornon, J. P. (1989). Cellulase families revealed by hydrophobic cluster analysis. *Gene*, *81*(1), 83–95. https://doi.org/10.1016/0378-1119(89)90339-9

Henske, J. K., Gilmore, S. P., Knop, D., Cunningham, F. J., Sexton, J. A., Smallwood, C. R., Shutthanandan, V., Evans, J. E., Theodorou, M. K., & O'Malley, M. A. (2017). Transcriptomic characterization of Caecomyces churrovis: A novel, non-rhizoid-forming lignocellulolytic anaerobic fungus. *Biotechnology for Biofuels*, *10*(1), 1–12. https://doi.org/10.1186/s13068-017-0997-4

Henske, J. K., Wilken, S. E., Solomon, K. V., Smallwood, C. R., Shutthanandan, V., Evans, J. E., Theodorou, M. K., & O'Malley, M. A. (2018). Metabolic characterization of anaerobic fungi provides a path forward for bioprocessing of crude lignocellulose. *Biotechnology and Bioengineering*, *115*(4), 874–884. https://doi.org/10.1002/bit.26515

Hess, M., Paul, S. S., Puniya, A. K., & Giezen, M. Van Der. (2020). Anaerobic Fungi: Past, Present, and Future. *Frontiers in Microbiology*, *11*, 584893. https://doi.org/10.3389/fmicb.2020.584893

Heyer, R., Kohrs, F., Reichl, U., & Benndorf, D. (2015). Metaproteomics of complex microbial communities in biogas plants. *Microbial Biotechnology*, *8*(5), 749–763. https://doi.org/10.1111/1751-7915.12276

Hobson, P. N., & Stewart, C. S. (1997). *The Rumen Microbial Ecosystem*. Springer. https://doi.org/10.1007/978-94-009-1453-7

Hofmeyr, S., Egan, R., Georganas, E., Copeland, A. C., Riley, R., Clum, A., Eloe-Fadrosh, E., Roux, S., Goltsman, E., Buluç, A., Rokhsar, D., Oliker, L., & Yelick, K. (2020). Terabase-scale metagenome coassembly with MetaHipMer. *Scientific Reports*, *10*(1), 1–11. https://doi.org/10.1038/s41598-020-67416-5

Hooton, D., Lentle, R., Monro, J., Wickham, R., & Simpson, R. (2015). The Secretion and Action of Brush Border Enzymes in the Mammalian Small Intestine. In B. Nilius, T. Gudermann, R. Jahn, O. H. Petersen, & P. P. de Tombe (Eds.), Reviews of Physiology, Biochemistry and Pharmacology (168th ed., pp. 59–118). Springer Verlag.

Houtkamp, I., van Zijll Langhout, M., Bessem, M., Pirovano, W., & Kort, R. (2022). Multiomics characterization of the zoo-housed gorilla gut microbiome reveals loss of bacterial diversity besides abundant fungal cellulose-degrading and archaeal methanogenic activity. *BioRxiv*, *2022.11.15*. https://doi.org/10.1101/2022.11.15.516570

Hug, L. A., & Co, R. (2018). It Takes a Village: Microbial Communities Thrive through Interactions and Metabolic Handoffs. *MSystems*, *3*(2), 1–5. https://doi.org/10.1128/msystems.00152-17

Hume, I. D., & Warner, A. C. I. (1979). Evolution of microbial digestion in mammals. In Y. Ruckebusch & P. Thivend (Eds.), *Digestive Physiology and Metabolism in Ruminants: Proceedings of the 5th International Symposium on Ruminant Physiology, held at Clermont-Ferrand* (pp. 665–684). Springer Netherlands.

Hume, Ian D. (2013). Fermentation in the hindgut of mammals. In R. I. Mackie & B. A. White (Eds.), *Gastrointestinal Microbiology* (pp. 84–115). Chapman & Hall.

Hungate, R. E., & Macy, J. (1973). The roll-tube method for cultivation of strict anaerobes. *Bulletins from the Ecological Research Committee (Stockholm)*, *17*.

Illumina. (2013). *16S Metagenomic Sequencing Library Preparation*. http://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf

Isenman, L., Liebow, C., & Rothman, S. (1999). The endocrine secretion of mammalian digestive enzymes by exocrine glands. American Journal of Physiology, 276(Endocrinology and Metabolism, 39), E223-32. https://doi.org/10.1177/175114370400500315

IUCN. (2022). *The IUCN Red List of Threatened Species*. https://www.iucnredlist.org

James, S. A., Parker, A., Purse, C., Telatin, A., Baker, D., Holmes, S., Durham, J., Funnell, S. G. P., & Carding, S. R. (2022). The Cynomolgus Macaque Intestinal Mycobiome Is Dominated by the Kazachstania Genus and K. pintolopesii Species. *Journal of Fungi*, *8*, 1054. https://doi.org/10.3390/jof8101054

Janiak, M. C. (2016). Digestive enzymes of human and nonhuman primates. Evolutionary Anthropology, 25(5), 253–266. https://doi.org/10.1002/evan.21498

Jiang, D., Armour, C. R., Hu, C., Mei, M., & Tian, C. (2019). Microbiome Multi-Omics Network Analysis: Statistical Considerations, Limitations, and Opportunities. *Frontiers in Genetics*, *10*(November), 1–19. https://doi.org/10.3389/fgene.2019.00995

Jiménez, D. J., Julia, M., Brossi, D. L., Schückel, J., Kre, S., George, W., Willats, T., & Elsas, J. D. Van. (2016). Characterization of three plant biomass-degrading microbial consortia by metagenomics- and metasecretomics-based approaches. *Applied Microbiology and Biotechnology*, *100*, 10463–10477. https://doi.org/10.1007/s00253-016-7713-3

Jin, H., You, L., Zhao, F., Li, S., Ma, T., Kwok, L. Y., Xu, H., & Sun, Z. (2022). Hybrid, ultra-deep metagenomic sequencing enables genomic and functional characterization of low-abundance species in the human gut microbiome. *Gut Microbes*, *14*(1), e2021790. https://doi.org/10.1080/19490976.2021.2021790

Jin, S., Jeon, Y., Jeon, M. S., Shin, J., Song, Y., Kang, S., Bae, J., Cho, S., Lee, J. K., Kim, D. R., & Cho, B. K. (2021). Acetogenic bacteria utilize light-driven electrons as an energy source for autotrophic growth. In *Proceedings of the National Academy of Sciences of the United States of America* (Vol. 118, Issue 9, pp. 1–7). https://doi.org/10.1073/pnas.2020552118

Jin, W., Cheng, Y. F., Mao, S. Y., & Zhu, W. Y. (2011). Isolation of natural cultures of anaerobic fungi and indigenously associated methanogens from herbivores and their bioconversion of lignocellulosic materials to methane. *Bioresource Technology*, *102*(17), 7925–7931. https://doi.org/10.1016/j.biortech.2011.06.026

Joblin, K. N., & Naylor, G. E. (1993). Inhibition of the rumen anaerobic fungus Neocallimastix frontalis by fermentation products. *Letters in Applied Microbiology*, *16*(5), 254–256. https://doi.org/10.1111/j.1472-765X.1993.tb01412.x

Joblin, K. N., & Williams, A. G. (1991). Effect of cocultivation of ruminal chytrid fungi with Methanobrevibacter smithii on lucerne stem degradation and extracellular fungal enzyme activities. *Letters in Applied Microbiology*, *12*(4), 121–124. https://doi.org/10.1111/j.1472-765X.1991.tb00520.x

Jung, H. J. G., & Engels, F. M. (2002). Alfalfa stem tissues: Cell wall deposition, composition, and degradability. Crop Science, 42(2), 524–534. https://doi.org/10.2135/cropsci2002.5240

Jung, H. J. G., & Lamb, J. A. F. S. (2004). Prediction of cell wall polysaccharide and lignin concentrations of alfalfa stems from detergent fiber analysis. Biomass and Bioenergy, 27(4), 365–373. https://doi.org/10.1016/j.biombioe.2004.04.001

Jung, H. J. G., & Lamb, J. A. F. S. (2006). Stem morphological and cell wall traits associated with divergent in vitro neutral detergent fiber digestibility in alfalfa clones. Crop Science, 46(5), 2054–2061. https://doi.org/10.2135/cropsci2005.12.0470

Kaehler, B. D., Bokulich, N. A., McDonald, D., Knight, R., Caporaso, J. G., & Huttley, G. A. (2019). Species abundance information improves sequence taxonomy classification accuracy. *Nature Communications*, *10*(1), 1–10. https://doi.org/10.1038/s41467-019-12669-6

Karasov, W. H., & Douglas, A. E. (2013). Comparative digestive physiology. *Comprehensive Physiology*, *3*(2), 741–783. https://doi.org/10.1002/cphy.c110054

Karp, P. D., Midford, P. E., Billington, R., Kothari, A., Krummenacker, M., Latendresse, M., Ong, W. K., Subhraveti, P., Caspi, R., Fulcher, C., Keseler, I. M., & Paley, S. M. (2021). Pathway Tools version 23.0 update: software for pathway / genome informatics and systems biology. *Briefings in Bioinformatics*, *22*(1), 109–126. https://doi.org/10.1093/bib/bbz104

Kato, A., Azuma, J., & Koshijima, T. (1987). Isolation and identification of a new feruloylated tetrasaccharide from bagasse lignincarbohydrate complex containing phenolic acid. Agricultural and Biological Chemistry, 51(6), 1691–1693. https://doi.org/10.1080/00021369.1987.10868250

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., & Glöckner, F. O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Research*, *41*(1), 1–11. https://doi.org/10.1093/nar/gks808

Koetschan, C., Kittelmann, S., Lu, J., Al-Halbouni, D., Jarvis, G. N., Müller, T., Wolf, M., & Janssen, P. H. (2014). Internal transcribed spacer 1 secondary structure analysis reveals a common core throughout the anaerobic fungi (Neocallimastigomycota). *PLoS ONE*, *9*(3). https://doi.org/10.1371/journal.pone.0091928

Kohl, K. D., Weiss, R. B., Cox, J., Dale, C., & Dearing, M. D. (2014). Gut microbes of mammalian herbivores facilitate intake of plant toxins. *Ecology Letters*, *17*(10), 1238–1246. https://doi.org/10.1111/ele.12329

Konwar, K. M., Hanson, N. W., Pagé, A. P., & Hallam, S. J. (2013). MetaPathways: a modular pipeline for constructing pathway / genome databases from environmental sequence information. *BMC Bioinformatics*, *14*, 202.

Kouzuma, A., Kato, S., & Watanabe, K. (2015). Microbial interspecies interactions: Recent findings in syntrophic consortia. *Frontiers in Microbiology*, *6*(MAY), 1–8. https://doi.org/10.3389/fmicb.2015.00477

Kumamoto, C. A., Gresnigt, M. S., & Hube, B. (2020). The gut, the bad, and the harmless: Candida albicans as a commensal and opportunistic pathogen in the intestine. *Current Opinion in Microbiology*, *56*, 7–15. https://doi.org/10.1016/j.mib.2020.05.006.The

Kumar, S., Indugu, N., Vecchiarelli, B., & Pitta, D. W. (2015). Associative patterns among anaerobic fungi, methanogenic archaea, and bacterial communities in response to changes in diet and age in the rumen of dairy cows. *Frontiers in Microbiology*, *6*, 781. https://doi.org/10.3389/fmicb.2015.00781

Kwiatkowski, N. P., Babiker, W. M., Merz, W. G., Carroll, K. C., & Zhang, S. X. (2012). Evaluation of nucleic acid sequencing of the D1/D2 region of the large subunit of the 28S rDNA and the internal transcribed spacer region using smartgene idn software for identification of filamentous fungi in a clinical laboratory. *Journal of Molecular Diagnostics*, *14*(4), 393–401. https://doi.org/10.1016/j.jmoldx.2012.02.004

Laine, R. A. (1994). Invited commentary: A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05 × 10 structures for a reducing hexasaccharide: The Isomer Barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology*, *4*(6), 759–767. https://doi.org/10.1093/glycob/4.6.759

Lambert, J. E. (2011). Primate seed dispersers as umbrella species: A case study from Kibale National Park, Uganda, with implications for Afrotropical forest conservation. *American Journal of Primatology*, *73*, 9–24. https://doi.org/10.1002/ajp.20879

Langer, P. (2001). Evidence from the digestive tract on phylogenetic relationships in ungulates and whales. *Journal of Zoological Systematics and Evolutionary Research*, *39*(1–2), 77–90. https://doi.org/10.1046/j.1439-0469.2001.00158.x

Lavrinienko, A., Scholier, T., Bates, S. T., Miller, A. N., & Watts, P. C. (2021). Defining gut mycobiota for wild animals : a need for caution in assigning authentic resident fungal taxa. *Animal Microbiome*. https://doi.org/10.1186/s42523-021-00134-z

Lazarus, K. L., & James, T. Y. (2015). Surveying the biodiversity of the Cryptomycota using a targeted PCR approach. *Fungal Ecology*, *14*, 62–70. https://doi.org/10.1016/j.funeco.2014.11.004

Lazuka, A., Auer, L., Bozonnet, S., Morgavi, D. P., O'Donohue, M., & Hernandez-Raquet, G. (2015). Efficient anaerobic transformation of raw wheat straw by a robust cow rumen-derived microbial consortium. *Bioresource Technology*, *196*, 241–249. https://doi.org/10.1016/j.biortech.2015.07.084

Lazuka, A., Auer, L., O'Donohue, M., & Hernandez-Raquet, G. (2018). Anaerobic lignocellulolytic microbial consortium derived from termite gut: Enrichment, lignocellulose degradation and community dynamics. *Biotechnology for Biofuels*, *11*(1), 1–14. https://doi.org/10.1186/s13068-018-1282-x

Leary, N. A. O., Wright, M. W., Brister, J. R., Ciufo, S., Haddad, D., Mcveigh, R., Rajput, B., Robbertse, B., Smith-white, B., Ako-adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., … Pruitt, K. D. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Research*, *44*, D733–D745. https://doi.org/10.1093/nar/gkv1189

Lee, J. Y., Tsolis, R. M., & Bäumler, A. J. (2022). The microbiome and gut homeostasis. *Science*, *377*(6601). https://doi.org/10.1126/science.abp9960

Leggieri, P. A., Kerdman-Andrade, C., Lankiewicz, T. S., Valentine, M. T., & O'Malley, M. A. (2021). Non-destructive quantification of anaerobic gut fungi and methanogens in co-culture reveals increased fungal growth rate and changes in metabolic flux relative to mono-culture. *Microbial Cell Factories*, *20*, 199. https://doi.org/10.1186/s12934-021-01684-2

Leggieri, P. A., Liu, Y., Hayes, M., Connors, B., Seppälä, S., O'Malley, M. A., & Venturelli, O. S. (2021). Integrating Systems and Synthetic Biology to Understand and Engineer Microbiomes. *Annual Review of Biomedical Engineering*, *23*(1), 169–201. https://doi.org/10.1146/annurev-bioeng-082120-022836

Leng, L., Yang, P., Singh, S., Zhuang, H., Xu, L., Chen, W. H., Dolfing, J., Li, D., Zhang, Y., Zeng, H., Chu, W., & Lee, P. H. (2018). A review on the bioenergetics of anaerobic microbial metabolism close to the thermodynamic limits and its implications for digestion applications. *Bioresource Technology*, *247*(July 2017), 1095–1106. https://doi.org/10.1016/j.biortech.2017.09.103

Ley, R. E., Lozupone, C. A., Hamady, M., Knight, R., & Gordon, J. I. (2008). Worlds within worlds: Evolution of the vertebrate gut microbiota. *Nature Reviews Microbiology*, *6*(10), 776–788. https://doi.org/10.1038/nrmicro1978

Gleason, F. H., Letcher, P. M., & McGee, P. a. (2007). Some aerobic Blastocladiomycota and Chytridiomycota can survive but cannot grow under anaerobic conditions. Australasian Mycologist, 26(2–3), 57–64.

Li, C. Y., Li, X. L., Su, X. X., Yang, Y. W., & Li, H. L. (2021). Effects of Alpine Wetland Degradation on Soil Microbial Structure and Diversity on the Qinghai Tibet Plateau. Eurasian Soil Science, 54(S1), S33–S41. https://doi.org/10.1134/S1064229322030097

Li, D., Liu, C. M., Luo, R., Sadakane, K., & Lam, T. W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, *31*(10), 1674–1676. https://doi.org/10.1093/bioinformatics/btv033

Li, F., Bao, W., Ni, D. W., Huang, X., & Jun, G. (2019). A thermoset hybrid sol for the syntheses of zirconium carbide – silicon carbide foam via replica method. *Journal of Porous Materials*, *26*(2), 409–417. https://doi.org/10.1007/s10934-018-0618-z

Li, J., Zhou, P., Liu, H., Xiong, C., Lin, J., Xiao, W., Gong, Y., & Liu, Z. (2014). Synergism of cellulase, xylanase, and pectinase on hydrolyzing sugarcane bagasse resulting from different pretreatment technologies. Bioresource Technology, 155, 258–265. https://doi.org/10.1016/j.biortech.2013.12.113

Li, Yuqi, Sun, M., Li, Y., Cheng, Y., & Zhu, W. (2020). Co-cultured methanogen improved the metabolism in the hydrogenosome of anaerobic fungus as revealed by gas chromatography-mass spectrometry analysis. *Asian-Australasian Journal of Animal Sciences*, *33*(12), 1948–1956. https://doi.org/10.5713/ajas.19.0649

Liao, J. C., Mi, L., Pontrelli, S., & Luo, S. (2016). Fuelling the future: Microbial engineering for the production of sustainable biofuels. *Nature Reviews Microbiology*, *14*(5), 288–304. https://doi.org/10.1038/nrmicro.2016.32

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*(7), 923–930. https://doi.org/10.1093/bioinformatics/btt656

Liggenstoffer, A. S., Youssef, N. H., Couger, M. B., & Elshahed, M. S. (2010). Phylogenetic diversity and community structure of anaerobic gut fungi (phylum Neocallimastigomycota) in ruminant and non-ruminant herbivores. *ISME Journal*, *4*(10), 1225–1235. https://doi.org/10.1038/ismej.2010.49

Limon, J. J., Tang, J., Li, D., Wolf, A. J., Michelsen, K. S., Funari, V., Gargus, M., Nguyen, C., Sharma, P., Maymi, V. I., Iliev, I. D., Skalski, J. H., Brown, J., Landers, C., Borneman, J., Braun, J., Targan, S. R., McGovern, D. P. B., & Underhill, D. M. (2019). Malassezia Is Associated with Crohn's Disease and Exacerbates Colitis in Mouse Models. *Cell Host and Microbe*, *25*(3), 377-388.e6. https://doi.org/10.1016/j.chom.2019.01.007

Lindemann, S. R. (2020). A piece of the pie: engineering microbiomes by exploiting division of labor in complex polysaccharide consumption. *Current Opinion in Chemical Engineering*, *30*, 96–102. https://doi.org/10.1016/j.coche.2020.08.004

Liu, J. H., Bian, G. R., Zhu, W. Y., & Mao, S. Y. (2015). High-grain feeding causes strong shifts in ruminal epithelial bacterial community and expression of Toll-like receptor genes in goats. *Frontiers in Microbiology*, *6*(MAR), 1–10. https://doi.org/10.3389/fmicb.2015.00167

Liu, L., Huang, W. C., Liu, Y., & Li, M. (2021). Diversity of cellulolytic microorganisms and microbial cellulases. *International Biodeterioration and Biodegradation*, *163*(June), 105277. https://doi.org/10.1016/j.ibiod.2021.105277

Liu, S., Moon, C. D., Zheng, N., Huws, S., Zhao, S., & Wang, J. (2022). Opportunities and challenges of using metagenomic data to bring uncultured microbes into cultivation. *Microbiome*, *10*(1), 1–14. https://doi.org/10.1186/s40168-022-01272-5

Ljungdahl, L. G. (1986). The Autotrophic Pathway of Acetate Synthesis in Acetogenic Bacteria. *Annual Review of Microbiology*, *40*(89), 415–450.

Lobo, R. R., & Faciola, A. P. (2021). Ruminal Phages – A Review. *Frontiers in Microbiology*, *12*(December), 1–10. https://doi.org/10.3389/fmicb.2021.763416

Lukas, K. E. (1999). A review of nutritional and motivational factors contributing to the performance of regurgitation and reingestion in captive lowland gorillas (Gorilla gorilla gorilla). *Applied Animal Behaviour Science*, *63*(3), 237–249. https://doi.org/10.1016/S0168-1591(98)00239-1

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., … Wang, J. (2012). SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *GigaScience*, *1*, 18. https://doi.org/10.1186/s13742-015-0069-2

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., … Wang, J. (2015). Erratum to “SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler” [GigaScience, (2012), 1, 18]. *GigaScience*, *4*(1), 1. https://doi.org/10.1186/s13742-015-0069-2

Mackie, R. I. (2002). Mutualistic fermentative digestion in the gastrointestinal tract: Diversity and evolution. *Integrative and Comparative Biology*, *42*(2), 319–326.

Mall, A., Kasarlawar, S., & Saini, S. (2022). Limited Pairwise Synergistic and Antagonistic Interactions Impart Stability to Microbial Communities. *Frontiers in Ecology and Evolution*, *10*(February), 1–14. https://doi.org/10.3389/fevo.2022.648997

Mann, A. E., Mazel, F., Lemay, M. A., Morien, E., Billy, V., Kowalewski, M., Di, A., Andrés, F., Tony, L., Stacey, L. G., Andrea, T., Andres, L. B., Sauther, M. L., Cuozzo, F. P., Rice, G. A. O., Dominy, N. J., Stumpf, R., Lewis, R. J., Swedell, L., … Wegener, L. (2019). Biodiversity of protists and nematodes in the wild nonhuman primate gut. *The ISME Journal*, *14*, 609–622. https://doi.org/10.1038/s41396-019-0551-4

Mar Rodríguez, M., Pérez, D., Javier Chaves, F., Esteve, E., Marin-Garcia, P., Xifra, G., Vendrell, J., Jové, M., Pamplona, R., Ricart, W., Portero-Otin, M., Chacón, M. R., & Fernández Real, J. M. (2015). Obesity changes the human gut mycobiome. *Scientific Reports*, *5*, 14600. https://doi.org/10.1038/srep14600

Margulis, L. (1993). *Symbiosis in Cell Evolution*. W. H. Freeman.

Martín, R., Bermúdez-Humarán, L. G., & Langella, P. (2016). Gnotobiotic rodents: An in vivo model for the study of microbe-microbe interactions. *Frontiers in Microbiology*, *7*(MAR), 1–7. https://doi.org/10.3389/fmicb.2016.00409

Mason, P. M., & Stuckey, D. C. (2016). Biofilms, bubbles and boundary layers – A new approach to understanding cellulolysis in anaerobic and ruminant digestion. *Water Research*, *104*, 93–100. https://doi.org/10.1016/j.watres.2016.07.063

Matsuda, I., Chapman, C. A., & Clauss, M. (2019). Colobine forestomach anatomy and diet. *Journal of Morphology*, *July*, 1–9. https://doi.org/10.1002/jmor.21052

Matsuda, I., Clauss, M., Tuuga, A., Sugau, J., Hanya, G., Yumoto, T., Bernard, H., & Hummel, J. (2017). Factors Affecting Leaf Selection by Foregut-fermenting Proboscis Monkeys: New Insight from in vitro Digestibility and Toughness of Leaves. *Scientific Reports*, *7*(February). https://doi.org/10.1038/srep42774

Matsuda, I., Hashimoto, C., Ihobe, H., Yumoto, T., Baranga, D., Clauss, M., & Hummel, J. (2022). Dietary Choices of a Foregut-Fermenting Primate, Colobus guereza: A Comprehensive Approach Including Leaf Chemical and Mechanical Properties,

Digestibility and Abundance. *Frontiers in Ecology and Evolution*, *10*, 795015. https://doi.org/10.3389/fevo.2022.795015

Matsuda, I., Murai, T., Clauss, M., Yamada, T., Tuuga, A., Bernard, H., & Higashi, S. (2011). Regurgitation and remastication in the foregut-fermenting proboscis monkey (Nasalis larvatus). *Biology Letters*, *7*(5), 786–789. https://doi.org/10.1098/rsbl.2011.0197

Matsuda, I., Sha, J. C. M., Ortmann, S., Schwarm, A., Grandl, F., Caton, J., Jens, W., Kreuzer, M., Marlena, D., Hagen, K. B., & Clauss, M. (2015). Excretion patterns of solute and different-sized particle passage markers in foregut-fermenting proboscis monkey (Nasalis larvatus) do not indicate an adaptation for rumination. *Physiology and Behavior*, *149*, 45–52. https://doi.org/10.1016/j.physbeh.2015.05.020

Mccord, A. I., Chapman, C. A., Weny, G., Tumukunde, A., Hyeroba, D., Klotz, K., Koblings, A. S., Mbora, D. N. M., Cregger, M., White, B. A., Leigh, S. R., & Goldberg, T. L. (2014). 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. https://doi.org/10.1002/ajp.22238

McFall-Ngai, M., Hadfield, M. G., Bosch, T. C. G., Carey, H. V., Domazet-Lošo, T., Douglas, A. E., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S. F., Hentschel, U., King, N., Kjelleberg, S., Knoll, A. H., Kremer, N., Mazmanian, S. K., Metcalf, J. L., Nealson, K., Pierce, N. E., … Wernegreen, J. J. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(9), 3229–3236. https://doi.org/10.1073/pnas.1218525110

McKenney, E. A., Maslanka, M., Rodrigo, A., & Yoder, A. D. (2018). Bamboo Specialists from Two Mammalian Orders (Primates, Carnivora) Share a High Number of Low-Abundance Gut Microbes. *Microbial Ecology*, *76*(1), 272–284. https://doi.org/10.1007/s00248-017-1114-8

McKenzie, V. J., Song, S. J., Delsuc, F., Prest, T. L., Oliverio, A. M., Korpita, T. M., Alexiev, A., Amato, K. R., Metcalf, J. L., Kowalewski, M., Avenant, N. L., Link, A., Di Fiore, A., Seguin-Orlando, A., Feh, C., Orlando, L., Mendelson, J. R., Sanders, J., & Knight, R. (2017). The effects of captivity on the mammalian gut microbiome. *Integrative and Comparative Biology*, *57*(4), 690–704. https://doi.org/10.1093/icb/icx090

McMurdie, P. J., & Holmes, S. (2014). Waste Not, Want Not: Why Rarefying Microbiome Data Is Inadmissible. *PLoS Computational Biology*, *10*(4). https://doi.org/10.1371/journal.pcbi.1003531

Meili, C. H., Jones, A. L., Arreola, A. X., Habel, J., Pratt, C. J., Hanafy, R. A., Wang, Y., Yassin, A. S., TagElDein, M. A., Moon, C. D., Janssen, P. H., Shrestha, M., Rajbhandari, P., Nagler, M., Vinzelj, J. M., Podmirseg, S. M., Stajich, J. E., Goetsch, A. L., Hayes, J., … Elshahed, M. S. (2022). Patterns and determinants of the global herbivorous mycobiome. *BioRxiv*, *2022.11.21*. https://doi.org/10.1101/2022.11.21.517404

Melo-Silveira, R. F., Fidelis, G. P., Pereira Costa, M. S. S., Telles, C. B. S., Dantas-Santos, N., de Oliveira Elias, S., Ribeiro, V. B., Barth, A. L., Macedo, A. J., Leite, E. L., & Rocha, H. A. O. (2012). In vitro antioxidant, anticoagulant and antimicrobial activity and in inhibition of cancer cell proliferation by xylan extracted from corn cobs. International Journal of Molecular Sciences, 13(1), 409–426. https://doi.org/10.3390/ijms13010409

Miceli, M. H., Díaz, J. A., & Lee, S. A. (2011). Emerging opportunistic yeast infections. *The Lancet Infectious Diseases*, *11*(2), 142–151. https://doi.org/10.1016/S1473-3099(10)70218-8

Miller, T. L., & Wolin, M. J. (1974). A Serum Bottle Modification of the Hungate Technique for Cultivating Obligate Anaerobes. *Applied Microbiology*, *27*(5), 985–987. https://doi.org/10.1128/am.27.5.985-987.1974

Milton, K. (1999). Nutritional characteristics of wild primate foods: Do the diets of our closest living relatives have lessons for us? *Nutrition*, *15*(6), 488–498. https://doi.org/10.1016/S0899-9007(99)00078-7

Moore, K. J., & Hatfield, R. D. (2015). Carbohydrates and forage quality. In G. C. Fahey, L. E. Moser, D. R. Mertens, & M. Collins (Eds.), Forage Quality, Evaluation, and Utilization (pp. 229–280). American Society of Agronomy, Inc.; Crop Science Society of America, Inc.; Soil Science Society of America, Inc. https://doi.org/10.2134/1994.foragequality.c6

Morris, B. E. L., Henneberger, R., Huber, H., & Moissl-Eichinger, C. (2013). Microbial syntrophy: Interaction for the common good. *FEMS Microbiology Reviews*, *37*(3), 384–406. https://doi.org/10.1111/1574-6976.12019

Morrison, D. J., & Preston, T. (2016). Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. *Gut Microbes*, *7*(3), 189–200. https://doi.org/10.1080/19490976.2015.1134082

Morrison, G. A., Fu, J., Lee, G. C., Wiederhold, N. P., Cañete-Gibas, C. F., Bunnik, E. M., & Wickes, B. L. (2020). Nanopore sequencing of the fungal intergenic spacer sequence as

a potential rapid diagnostic assay. *Journal of Clinical Microbiology*, *58*(12), 1–16. https://doi.org/10.1128/JCM.01972-20

Mozaffari, M., Rosen, C. J., Russelle, M. P., & Nater, E. A. (2000). Chemical Characterization of Ash from Gasification of Alfalfa Stems: Implications for Ash Management. Journal of Environmental Quality, 29(3), 963–972. https://doi.org/10.2134/jeq2000.00472425002900030035x

Mukherjee, S., Seshadri, R., Varghese, N. J., Eloe-fadrosh, E. A., Meier-kolthoff, J. P., Göker, M., Coates, R. C., Hadjithomas, M., Pavlopoulos, G. A., Paez-espino, D., Yoshikuni, Y., Visel, A., Whitman, W. B., Garrity, G. M., Eisen, J. A., Hugenholtz, P., Pati, A., Ivanova, N. N., Woyke, T., … Kyrpides, N. C. (2017). 1,003 reference genomes of bacterial and archaeal isolates expand coverage of the tree of life. *Nature Biotechnology*, *35*(7), 676–683. https://doi.org/10.1038/nbt.3886

Mura, E., Edwards, J., Kittelmann, S., Kaerger, K., Voigt, K., Mrázek, J., Moniello, G., & Fliegerova, K. (2019). Anaerobic fungal communities differ along the horse digestive tract. *Fungal Biology*, *123*(3), 240–246. https://doi.org/10.1016/j.funbio.2018.12.004

Murphy, C. L., Youssef, N. H., Hanafy, R. A., Couger, M. B., Stajich, J. E., Wang, Y., Baker, K., Dagar, S. S., Griffith, G. W., Farag, I. F., Callaghan, T. M., & Elshahed, M. S. (2019). Horizontal gene transfer as an indispensable driver for Neocallimastigomycota evolution into a distinct gut-dwelling fungal lineage. *Applied and Environmental Microbiology*, *405*. https://doi.org/10.1128/aem.00988-19

Nair, R. R., Vasse, M., Wielgoss, S., Sun, L., Yu, Y. T. N., & Velicer, G. J. (2019). Bacterial predator-prey coevolution accelerates genome evolution and selects on virulence-associated prey defences. *Nature Communications*, *10*(1), 1–10. https://doi.org/10.1038/s41467-019-12140-6

Newbold, C. J., & Ramos-Morales, E. (2020). Review: Ruminal microbiome and microbial metabolome: Effects of diet and ruminant host. *Animal*, *14*(S1), S78–S86. https://doi.org/10.1017/S1751731119003252

Nijboer, J., & Clauss, M. (2006). The Digestive Physiology of Colobine Primates. In J. Nijboer (Ed.), *Fibre Intake and Faeces Quality in Leaf-Eating Monkeys* (Issue January, pp. 9–28). Utrecht University Repository. https://doi.org/10.5167/uzh-3520

Nilsson, R. Henrik, Wurzbacher, C., Bahram, M., Coimbra, V. R. M., Larsson, E., Tedersoo, L., Eriksson, J., Ritter, C. D., Svantesson, S., Sánchez-García, M., Ryberg, M., Kristiansson, E., & Abarenkov, K. (2016). Top 50 most wanted fungi. *MycoKeys*, *12*, 29–40. https://doi.org/10.3897/mycokeys.12.7553

Nilsson, Rolf Henrik, Larsson, K. H., Taylor, A. F. S., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., Kennedy, P., Picard, K., Glöckner, F. O., Tedersoo, L., Saar, I., Kõljalg, U., & Abarenkov, K. (2019). The UNITE database for molecular identification of fungi: Handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Research*, *47*(D1), D259–D264. https://doi.org/10.1093/nar/gky1022

Nishida, A. H., & Ochman, H. (2018a). *Rates of gut microbiome divergence in mammals*. *September 2017*, 1884–1897. https://doi.org/10.1111/mec.14473

Nishida, A. H., & Ochman, H. (2018b). Rates of Gut Microbiome Divergence in Mammals. *Molecular Ecology*, *27*(8), 1884–1897. https://doi.org/10.1111/ijlh.12426

Nishimura, H., Kamiya, A., Nagata, T., Katahira, M., & Watanabe, T. (2018). Direct evidence for α ether linkage between lignin and carbohydrates in wood cell walls. *Scientific Reports*, *8*(1), 1–11. https://doi.org/10.1038/s41598-018-24328-9

Noor, E., Flamholz, A., Bar-Even, A., Davidi, D., Milo, R., & Liebermeister, W. (2016). The Protein Cost of Metabolic Fluxes: Prediction from Enzymatic Rate Laws and Cost Minimization. PLoS Computational Biology, 12(11), 1–29. https://doi.org/10.1371/journal.pcbi.1005167

Novaes, R. F. V. (1986). Microbiology of Anaerobic Digestion. *Water Science and Technology*, *18*(12), 1–14.

Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P. A. (2017). MetaSPAdes: A new versatile metagenomic assembler. *Genome Research*, *27*(5), 824–834. https://doi.org/10.1101/gr.213959.116

Ochman, H., Worobey, M., Kuo, C. H., Ndjango, J. B. N., Peeters, M., Hahn, B. H., & Hugenholtz, P. (2010). Evolutionary relationships of wild hominids recapitulated by gut microbial communities. *PLoS Biology*, *8*(11), 3–10. https://doi.org/10.1371/journal.pbio.1000546

Op De Beeck, M., Lievens, B., Busschaert, P., Declerck, S., Vangronsveld, J., & Colpaert, J. V. (2014). Comparison and validation of some ITS primer pairs useful for fungal metabarcoding studies. *PLoS ONE*, *9*(6). https://doi.org/10.1371/journal.pone.0097629

Orpin, C. G., & Ho, Y. W. (1991). Ecology and function of the anaerobic rumen fungi. In *Recent Advances on the Nutrition of Herbivores* (pp. 163–170). Malaysian Society of Animal Production.

Peng, X., Swift, C. L., Theodorou, M. K., & O'Malley, M. A. (2018). Methods for genomic characterization and maintenance of anaerobic fungi. *Methods in Molecular Biology*, *1775*, 53–67. https://doi.org/10.1007/978-1-4939-7804-5_5

Peng, X., Wilken, S. E., Lankiewicz, T. S., Gilmore, S. P., Brown, J. L., Henske, J. K., Swift, C. L., Salamov, A., Barry, K., Grigoriev, I. V, Theodorou, M. K., Valentine, D. L., & Malley, M. A. O. (2021). Genomic and functional analyses of fungal and bacterial consortia that enable lignocellulose breakdown in goat gut microbiomes. *Nature Microbiology*, *6*, 499–511. https://doi.org/10.1038/s41564-020-00861-0

Pérez, J., Muñoz-Dorado, J., De La Rubia, T., & Martínez, J. (2002). Biodegradation and biological treatments of cellulose, hemicellulose and lignin: An overview. *International Microbiology*, *5*(2), 53–63. https://doi.org/10.1007/s10123-002-0062-3

Pérez, S., & Tvaroška, I. (2014). Carbohydrate-protein interactions: Molecular modeling insights. In D. Horton (Ed.), Advances in Carbohydrate Chemistry and Biochemistry (Vol. 71, pp. 9–136). Academic Press. https://doi.org/10.1016/B978-0-12-800128-8.00001-7

Perofsky, A. C., Lewis, R. J., Abondano, L. A., Difiore, A., & Meyers, L. A. (2017). Hierarchical social networks shape gut microbial composition in wild Verreaux's sifaka. *Proceedings of the Royal Society B: Biological Sciences*, *284*, 20172274. https://doi.org/10.1098/rspb.2017.2274

Picard, K. T. (2017). Coastal marine habitats harbor novel early-diverging fungal diversity. *Fungal Ecology*, *25*, 1–13. https://doi.org/10.1016/j.funeco.2016.10.006

Pierre, J. F., Phillips, G. J., Chandra, L. C., Rendina, D. N., Thomas-gosain, N. F., & Lubach, G. R. (2021). Lyticase Facilitates Mycobiome Resolution Without Disrupting Microbiome Fidelity in Primates. *Journal of Surgical Research*, *267*(267), 336–341. https://doi.org/10.1016/j.jss.2021.06.023

Podolsky, I. A., Seppälä, S., Lankiewicz, T. S., Brown, J. L., Swift, C. L., & O'Malley, M. A. (2019). Harnessing Nature's Anaerobes for Biotechnology and Bioprocessing. *Annual Review of Chemical and Biomolecular Engineering*, *10*, 105–128. https://doi.org/10.1146/annurev-chembioeng-060718-030340

Popovich, D. G., Jenkins, D. J. A., Kendall, C. W. C., Dierenfeld, E. S., Carroll, R. W., Tariq, N., & Vidgen, E. (1997). The Western Lowland Gorilla Diet Has Implications for the Health of Humans and Other Hominoids. *Human and Clinical Nutrition*, *10*, 2000–2005.

Pruitt, K., Brown, G., Tatusova, T., & Maglott, D. (2002). *Chapter 18 . The Reference Sequence (RefSeq) Database*.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D. R., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., … Zoetendal, E. (2010). A human gut microbial gene

catalogue established by metagenomic sequencing. *Nature*, *464*(7285), 59–65. https://doi.org/10.1038/nature08821

Raimondi, S., Amaretti, A., Gozzoli, C., Simone, M., Righini, L., Candeliere, F., Brun, P., Ardizzoni, A., Colombari, B., Paulone, S., Castagliuolo, I., Cavalieri, D., Blasi, E., Rossi, M., & Peppoloni, S. (2019). Longitudinal survey of fungi in the human gut: ITS profiling, phenotyping, and colonization. Frontiers in Microbiology, 10, 1575. https://doi.org/10.3389/fmicb.2019.01575

Rakoff-Nahoum, S., Foster, K. R., & Comstock, L. E. (2016). The evolution of cooperation within the gut microbiota. *Nature*, *533*(7602), 255–259. https://doi.org/10.1038/nature17626

Read, M. N., & Holmes, A. J. (2017). Towards an integrative understanding of diet-host-gut microbiome interactions. *Frontiers in Immunology*, *8*(MAY), 1–9. https://doi.org/10.3389/fimmu.2017.00538

Reese, A. T., & Dunn, R. R. (2018). Drivers of Microbiome Biodiversity: A Review of General Rules, Feces, and Ignorance. *MBio*, 1–14.

Reynolds, N. K., Jusino, M. A., Stajich, J. E., & Smith, M. E. (2022). Understudied, underrepresented, and unknown: Methodological biases that limit detection of early diverging fungi from environmental samples. *Molecular Ecology Resources*, *22*(3), 1065–1085. https://doi.org/10.1111/1755-0998.13540

Rivière, A., Selak, M., Lantin, D., Leroy, F., & De Vuyst, L. (2016). Bifidobacteria and butyrate-producing colon bacteria: Importance and strategies for their stimulation in the human gut. *Frontiers in Microbiology*, *7*(JUN). https://doi.org/10.3389/fmicb.2016.00979

Roberfroid, M., Gibson, G. R., Hoyles, L., McCartney, A. L., Rastall, R., Rowland, I., Wolvers, D., Watzl, B., Szajewska, H., Stahl, B., Guarner, F., Respondek, F., Whelan, K., Coxam, V., Davicco, M.-J., Léotoing, L., Wittrant, Y., Delzenne, N. M., Cani, P. D., … Meheust, A. (2010). Prebiotic effects: metabolic and health benefits. *British Journal of Nutrition*, *104*(S2), S1–S63. https://doi.org/10.1017/S0007114510003363

Saleem, F., Bouatra, S., Guo, A. C., Psychogios, N., Mandal, R., Dunn, S. M., Ametaj, B. N., & Wishart, D. S. (2013). The Bovine Ruminal Fluid Metabolome. *Metabolomics*, *9*(2), 360–378. https://doi.org/10.1007/s11306-012-0458-9

Sanders, J. G., Beichman, A. C., Roman, J., Scott, J. J., Emerson, D., McCarthy, J. J., & Girguis, P. R. (2015). Baleen whales host a unique gut microbiome with similarities to both carnivores and herbivores. *Nature Communications*, *6*, 1–8. https://doi.org/10.1038/ncomms9285

Saye, L. M. G., Navaratna, T. A., Chong, J. P. J., O'malley, M. A., Theodorou, M. K., & Reilly, M. (2021). The anaerobic fungi: Challenges and opportunities for industrial lignocellulosic biofuel production. *Microorganisms*, *9*(4), 1–28. https://doi.org/10.3390/microorganisms9040694

Scarborough, M. J., Lawson, C. E., Hamilton, J. J., Donohue, T. J., & Noguera, D. R. (2018). Metatranscriptomic and Thermodynamic Insights into Medium-Chain Fatty Acid Production Using an Anaerobic Microbiome. *MSystems*, *3*(6), e00221-18. https://doi.org/10.1128/msystems.00221-18

Schmidt, J. A., Albright, S., Calza, G. M., & Calza, R. E. (2004). Characterization of Aspergillus oryzae fermentation extract effects on the rumen fungus Neocallimastix frontalis, EB 188. Part 2. Carbon source utilization and effects on zoospore production. *Applied Microbiology and Biotechnology*, *63*(4), 431–437. https://doi.org/10.1007/s00253-003-1294-7

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., Chen, W., Bolchacova, E., Voigt, K., Crous, P. W., Miller, A. N., Wingfield, M. J., Aime, M. C., An, K. D., Bai, F. Y., Barreto, R. W., Begerow, D., Bergeron, M. J., Blackwell, M., … Schindel, D. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(16), 6241–6246. https://doi.org/10.1073/pnas.1117018109

Schulz, D., Qablan, M. A., Profousova-Psenkova, I., Vallo, P., Fuh, T., Modry, D., Piel, A. K., Stewart, F., Petrzelkova, K. J., & Fliegerová, K. (2018). Anaerobic Fungi in Gorilla (Gorilla gorilla gorilla) Feces: an Adaptation to a High-Fiber Diet? *International Journal of Primatology*, *39*(4), 567–580. https://doi.org/10.1007/s10764-018-0052-8

Scorzetti, G., Fell, J. W., Fonseca, A., & Statzell-Tallman, A. (2002). Systematics of basidiomycetous yeasts: A comparison of large subunit D1/D2 and internal transcribed spacer rDNA regions. *FEMS Yeast Research*, *2*(4), 495–517. https://doi.org/10.1016/S1567-1356(02)00128-9

Scott, G. B. D., & Keymer, I. F. (1975). Ulcerative colitis in apes: a comparison with the human disease. *Journal of Pathology*, *115*, 241–244.

Seppälä, S., Solomon, K. V., Gilmore, S. P., Henske, J. K., & O'Malley, M. A. (2016). Mapping the membrane proteome of anaerobic gut fungi identifies a wealth of carbohydrate binding proteins and transporters. *Microbial Cell Factories*, *15*(1), 1–14. https://doi.org/10.1186/s12934-016-0611-7

Seppälä, S., Wilken, S. E., Knop, D., Solomon, K. V., & O'Malley, M. A. (2017a). The importance of sourcing enzymes from non-conventional fungi for metabolic engineering and biomass breakdown. *Metabolic Engineering*, *44*(September), 45–59. https://doi.org/10.1016/j.ymben.2017.09.008

Seppälä, S., Wilken, S. E., Knop, D., Solomon, K. V., & O'Malley, M. A. (2017b). The importance of sourcing enzymes from non-conventional fungi for metabolic engineering and biomass breakdown. *Metabolic Engineering*, *44*(June), 45–59. https://doi.org/10.1016/j.ymben.2017.09.008

Seshadri, R., Leahy, S. C., Attwood, G. T., Teh, K. H., Lambie, S. C., Cookson, A. L., Eloe-Fadrosh, E. A., Pavlopoulos, G. A., Hadjithomas, M., Varghese, N. J., Paez-Espino, D., Perry, R., Henderson, G., Creevey, C. J., Terrapon, N., Lapebie, P., Drula, E., Lombard, V., Rubin, E., … Cerón Cucchi, M. (2018). Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. *Nature Biotechnology*, *36*(4), 359–367. https://doi.org/10.1038/nbt.4110

Shahab, R. L., Brethauer, S., Davey, M. P., Smith, A. G., Vignolini, S., Luterbacher, J. S., & Studer, M. H. (2020). A heterogeneous microbial consortium producing short-chain fatty acids from lignocellulose. *Science*, *369*(6507). https://doi.org/10.1126/science.abb1214

Shahab, R. L., Brethauer, S., Luterbacher, J. S., & Studer, M. H. (2020). Engineering of ecological niches to create stable artificial consortia for complex biotransformations. *Current Opinion in Biotechnology*, *62*, 129–136. https://doi.org/10.1016/j.copbio.2019.09.008

Sharma, A. K., Davison, S., Pafco, B., Clayton, J. B., Rothman, J. M., & Mclennan, M. R. (2022). The primate gut mycobiome-bacteriome interface is impacted by environmental and subsistence factors. *Npj Biofilms and Microbiomes*, *8*, 12. https://doi.org/10.1038/s41522-022-00274-3

Shoun, H., & Tanimoto, T. (1991). Denitrification by the fungus *Fusarium oxysporum* and involvement of cytochrome P-450 in the respiratory nitrite reduction. Journal of Biological Chemistry, 266(17), 11078–11082. https://doi.org/10.1016/s0021-9258(18)99130-1

Sigma-Aldrich. (2023). *Chloramphenicol, SDS No. C0378*. https://www.sigmaaldrich.com/US/pt/sds/sigma/c0378

Simeonidis, E., & Price, N. D. (2015). Genome-scale modeling for metabolic engineering. *Journal of Industrial Microbiology and Biotechnology*, *42*, 327–338. https://doi.org/10.1007/s10295-014-1576-3

Snelling, T. J., Auffret, M. D., Duthie, C.-A., Stewart, R. D., Watson, M., Dewhurst, R. J., Roehe, R., & Walker, A. W. (2019). Temporal stability of the rumen microbiota in beef cattle, and response to diet and supplements. *Animal Microbiome*, *1*, 16. https://doi.org/10.1186/s42523-019-0018-y

Solden, L. M., Naas, A. E., Roux, S., Daly, R. A., Collins, W. B., Nicora, C. D., Purvine, S. O., Hoyt, D. W., Schückel, J., Jørgensen, B., Willats, W., Spalinger, D. E., Firkins, J. L., Lipton, M. S., Sullivan, M. B., Pope, P. B., & Wrighton, K. C. (2018). Interspecies cross-feeding orchestrates carbon degradation in the rumen ecosystem. Nature Microbiology, 3(11), 1274–1284. https://doi.org/10.1038/s41564-018-0225-4

Solomon, R., Wein, T., Levy, B., Eshed, S., Dror, R., Reiss, V., Zehavi, T., & Furman, O. (2022). Protozoa populations are ecosystem engineers that shape prokaryotic community structure and function of the rumen microbial ecosystem. *ISME Journal*, *16*, 1187–1197. https://doi.org/10.1038/s41396-021-01170-y

Solomon, K. V, Brewer, H. M., Haitjema, C. H., Henske, J. K., Gilmore, S. P., Borges-rivera, D., Brewer, H. M., Purvine, S. O., Wright, A. T., Theodorou, M. K., Igor, V., Regev, A., Thompson, D. A., & Malley, M. A. O. (2016). Early-branching gut fungi possess a large, comprehensive array of biomass-degrading enzymes. *Science*, *351*(6278), 1192–1195. https://doi.org/10.1126/science.aad1431

Soverini, M., Quercia, S., Biancani, B., Furlati, S., Turroni, S., Biagi, E., Consolandi, C., Peano, C., Severgnini, M., Rampelli, S., Brigidi, P., & Candela, M. (2016). The bottlenose dolphin (Tursiops truncatus) faecal microbiota. *FEMS Microbiology Ecology*, *92*(4), 1–8. https://doi.org/10.1093/femsec/fiw055

Stewart, C. B., Schilling, J. W., & Wilson, A. C. (1987). Adaptive evolution in the stomach lysozymes of foregut fermenters. *Nature*, *330*(6146), 401–404. https://doi.org/10.1038/330401a0

Strong, V. J., Grindlay, D., Redrobe, S., Cobb, M., & White, K. (2016). A systematic review of the literature relating to captive great ape morbidity and mortality. *Journal of Zoo and Wildlife Medicine*, *47*(3), 697–710.

Sun, B., Gu, Z., Wang, X., Huffman, M. A., Garber, P. A., Sheeran, L. K., Zhang, D., Zhu, Y., Xia, D. P., & Li, J. hua. (2018). Season, age, and sex affect the fecal mycobiota of free-ranging Tibetan macaques (Macaca thibetana). *American Journal of Primatology*, *80*(7), 1–9. https://doi.org/10.1002/ajp.22880

Sun, B., Xia, Y., Garber, P. A., Amato, K. R., Gomez, A., Xu, X., Li, W., Huang, M., Xia, D., Wang, X., & Li, J. (2021). Captivity Is Associated With Gut Mycobiome

Composition in Tibetan Macaques (Macaca thibetana). *Frontiers in Microbiology*, *12*, 665853. https://doi.org/10.3389/fmicb.2021.665853

Sun, H., Wu, Y. M., Wang, Y. M., Liu, J. X., & Myung, K. H. (2014). Effects of aspergillus oryzae culture and 2-hydroxy-4-(methylthio)-butanoic acid on in vitro rumen fermentation and microbial populations between different roughage sources. *Asian-Australasian Journal of Animal Sciences*, *27*(9), 1285–1292. https://doi.org/10.5713/ajas.2013.13742

Sun, J. X., Sun, X. F., Sun, R. C., & Su, Y. Q. (2004). Fractional extraction and structural characterization of sugarcane bagasse hemicelluloses. Carbohydrate Polymers, 56(2), 195–204. https://doi.org/10.1016/j.carbpol.2004.02.002

Swift, C. L., Brown, J. L., Seppälä, S., & O'Malley, M. A. (2019). Co-cultivation of the anaerobic fungus Anaeromyces robustus with Methanobacterium bryantii enhances transcription of carbohydrate active enzymes. *Journal of Industrial Microbiology and Biotechnology*, *46*(9–10), 1427–1433. https://doi.org/10.1007/s10295-019-02188-0

Swift, C. L., Louie, K. B., Bowen, B. P., Hooker, C. A., Solomon, K. V., Singan, V., Daum, C., Pennacchio, C. P., Barry, K., Shutthanandan, V., Evans, J. E., Grigoriev, I. V., Northen, T. R., & O'malley, M. A. (2021). Cocultivation of anaerobic fungi with rumen bacteria establishes an antagonistic relationship. *MBio*, *12*(4). https://doi.org/10.1128/mBio.01442-21

Sypherd, P. S., Borgia, P. T., & Paznokas, J. L. (1979). Biochemistry of Dimorphism in the Fungus *Mucor*. Advances in Microbial Physiology, 18(C), 67–104. https://doi.org/10.1016/S0065-2911(08)60415-4

Tang, G. S., Liang, X. X., Yang, M. Y., Wang, T. T., Chen, J. P., Du, W. G., Li, H., & Sun, B. J. (2020). Captivity Influences Gut Microbiota in Crocodile Lizards (Shinisaurus crocodilurus). *Frontiers in Microbiology*, *11*(April), 1–12. https://doi.org/10.3389/fmicb.2020.00550

Tedersoo, L., Anslan, S., Bahram, M., Kõljalg, U., & Abarenkov, K. (2020). Identifying the 'unidentified' fungi: a global-scale long-read third-generation sequencing approach. *Fungal Diversity*, *103*(1), 273–293. https://doi.org/10.1007/s13225-020-00456-4

Tedersoo, L., Anslan, S., Bahram, M., Põlme, S., Riit, T., Liiv, I., Kõljalg, U., Kisand, V., Nilsson, R. H., Hildebrand, F., Bork, P., & Abarenkov, K. (2015). Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. *MycoKeys*, *10*, 1–43. https://doi.org/10.3897/mycokeys.10.4852

Tedersoo, L., Bahram, M., Puusepp, R., Nilsson, R. H., & James, T. Y. (2017). Novel soil-inhabiting clades fill gaps in the fungal tree of life. *Microbiome*, *5*(1), 1–10. https://doi.org/10.1186/s40168-017-0259-5

Tedersoo, L., & Lindahl, B. (2016). Fungal identification biases in microbiome projects. *Environmental Microbiology Reports*, *8*(5), 774–779. https://doi.org/10.1111/1758-2229.12438

Tedersoo, L., Tooming-Klunderud, A., & Anslan, S. (2018). PacBio metabarcoding of Fungi and other eukaryotes: errors, biases and perspectives. *New Phytologist*, *217*(3), 1370–1385. https://doi.org/10.1111/nph.14776

Terry, S. A., Badhan, A., Wang, Y., Chaves, A. V., & McAllister, T. A. (2019). Fibre digestion by rumen microbiota — a review of recent metagenomic and metatranscriptomic studies. *Canadian Journal of Animal Science*, *99*(4), 678–692. https://doi.org/10.1139/cjas-2019-0024

Theodorou, M. K., Brookman, J., & Trinci, A. P. J. (2005). Anaerobic fungi. In H. P. S. Makkar & C. S. McSweeney (Eds.), *Methods in Gut Microbial Ecology for Ruminants* (pp. 55–66). Springer Science and Business Media.

Trinci, A. P. J., Davies, D. R., Gulli, K., Lawrence, M. I., Nielsew, B. B., I, A. R., & Theodorou, M. K. (1994). Anaerobic fungi in herbivorous animals. *Mycology Research*, *98*(2), 129–152. https://doi.org/10.1016/S0953-7562(09)80178-0

Tsoi, R., Wu, F., Zhang, C., Bewick, S., Karig, D., & You, L. (2018). Metabolic division of labor in microbial systems. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(10), 2526–2531. https://doi.org/10.1073/pnas.1716888115

Tsuchida, S., Kakooza, S., Nguema, P. P. M., Wampande, E. M., & Ushida, K. (2018). Characteristics of gorilla-specific lactobacillus isolated from captive and wild gorillas. *Microorganisms*, *6*(3), 1–11. https://doi.org/10.3390/microorganisms6030086

Turnbaugh, P. J., Ridaura, V. K., Faith, J. J., Rey, F. E., Knight, R., & Gordon, J. I. (2009). The effect of diet on the human gut microbiome: A metagenomic analysis in humanized gnotobiotic mice. *Science Translational Medicine*, *1*(6). https://doi.org/10.1126/scitranslmed.3000322

Uhlenhaut, C., & Kracht, M. (2005). Viral infectivity is maintained by an RNA protection buffer. *Journal of Virological Methods*, *128*(1–2), 189–191. https://doi.org/10.1016/j.jviromet.2005.05.002

Underhill, D. M., & Iliev, I. D. (2014). The mycobiota: Interactions between commensal fungi and the host immune system. *Nature Reviews Immunology*, *14*(6), 405–416. https://doi.org/10.1038/nri3684

van Tilburg Bernardes, E., Pettersen, V. K., Gutierrez, M. W., Laforest-Lapointe, I., Jendzjowsky, N. G., Cavin, J. B., Vicentini, F. A., Keenan, C. M., Ramay, H. R., Samara, J., MacNaughton, W. K., Wilson, R. J. A., Kelly, M. M., McCoy, K. D., Sharkey, K. A., & Arrieta, M. C. (2020). Intestinal fungi are causally implicated in microbiome assembly and immune development in mice. *Nature Communications*, *11*(1), 1–16. https://doi.org/10.1038/s41467-020-16431-1

Vandepol, N., Liber, J., Desirò, A., Na, H., Kennedy, M., Barry, K., Grigoriev, I. V., Miller, A. N., O'Donnell, K., Stajich, J. E., & Bonito, G. (2020). Resolving the Mortierellaceae phylogeny through synthesis of multi-gene phylogenetics and phylogenomics. *Fungal Diversity*, *104*(1), 267–289. https://doi.org/10.1007/s13225-020-00455-5

Vet, S., Gelens, L., & Gonze, D. (2020). Mutualistic cross-feeding in microbial systems generates bistability via an Allee effect. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-63772-4

Vu, D., Groenewald, M., de Vries, M., Gehrmann, T., Stielow, B., Eberhardt, U., Al-Hatmi, A., Groenewald, J. Z., Cardinali, G., Houbraken, J., Boekhout, T., Crous, P. W., Robert, V., & Verkley, G. J. M. (2019). Large-scale generation and analysis of filamentous fungal DNA barcodes boosts coverage for kingdom fungi and reveals thresholds for fungal species and higher taxon delimitation. *Studies in Mycology*, *92*, 135–154. https://doi.org/10.1016/j.simyco.2018.05.001

Wang, L., York, S. W., Ingram, L. O., & Shanmugam, K. T. (2019). Simultaneous fermentation of biomass-derived sugars to ethanol by a co-culture of an engineered Escherichia coli and Saccharomyces cerevisiae. *Bioresource Technology*, *273*, 269–276. https://doi.org/10.1016/j.biortech.2018.11.016

Wang, M., Chen, X., Liu, X., Fang, Y., Zheng, X., Huang, T., Tang, Y. Q., Ackermann, M., Nie, Y., & Wu, X. L. (2022). Even allocation of benefits stabilizes microbial community engaged in metabolic division of labor. *Cell Reports*, *40*(13), 111410. https://doi.org/10.1016/j.celrep.2022.111410

Wang, Y., Youssef, N. H., Couger, M. B., Hanafy, R. A., Elshahed, M. S., & Stajich, J. E. (2019). Molecular Dating of the Emergence of Anaerobic Rumen Fungi and the Impact of Laterally Acquired Genes. *MSystems*, *4*(4), 1–18. https://doi.org/10.1128/msystems.00247-19

Waskom, M. L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, *6*(60), 3021. https://doi.org/10.21105/joss.03021

Watanabe, H., & Tokuda, G. (2001). Animal cellulases. *Cellular and Molecular Life Sciences*, *58*(9), 1167–1178. https://doi.org/10.1007/PL00000931

Waters, J. L., & Ley, R. E. (2019). The human gut bacteria Christensenellaceae are widespread, heritable, and associated with health. *BMC Biology*, *17*(1), 1–11. https://doi.org/10.1186/s12915-019-0699-4

Wheeler, M. L., Limon, J. J., Bar, A. S., Leal, C. A., Gargus, M., Tang, J., Brown, J., Funari, V. A., Wang, H. L., Crother, T. R., Arditi, M., Underhill, D. M., & Iliev, I. D. (2016). Immunological Consequences of Intestinal Fungal Dysbiosis. *Cell Host and Microbe*, *19*(6), 865–873. https://doi.org/10.1016/j.chom.2016.05.003

Whitcomb, D. C., & Lowe, M. E. (2007). Human pancreatic digestive enzymes. Digestive Diseases and Sciences, 52(1), 1–17. https://doi.org/10.1007/s10620-006-9589-z

Whitman, W. B., Woyke, T., Klenk, H. P., Zhou, Y., Lilburn, T. G., Beck, B. J., De Vos, P., Vandamme, P., Eisen, J. A., Garrity, G., Hugenholtz, P., & Kyrpides, N. C. (2015). Genomic encyclopedia of bacterial and archaeal type strains, phase III: The genomes of soil and plant-associated and newly described type strains. *Standards in Genomic Sciences*, *10*(MAY2015), 8–13. https://doi.org/10.1186/s40793-015-0017-x

Wilken, S. E., Monk, J. M., Leggieri, P. A., Lawson, C. E., Lankiewicz, T. S., Seppälä, S., Daum, C. G., Jenkins, J., Lipzen, A. M., Mondo, S. J., Barry, K. W., Grigoriev, I. V, Henske, J. K., Theodorou, M. K., Palsson, B. O., Petzold, L. R., & Malley, A. O. (2021). Experimentally Validated Reconstruction and Analysis of a Genome-Scale Metabolic Model of an Anaerobic Neocallimastigomycota Fungus. *MSystems*, *6*(1).

Wilken, S. E., Seppälä, S., Lankiewicz, T. S., Saxena, M., Henske, J. K., Salamov, A. A., Grigoriev, I. V., & O'Malley, M. A. (2020). Genomic and proteomic biases inform metabolic engineering strategies for anaerobic fungi. *Metabolic Engineering Communications*, *10*(August 2019). https://doi.org/10.1016/j.mec.2019.e00107

Wilken, S., Saxena, M., Petzold, L., & O'Malley, M. (2018). In Silico Identification of Microbial Partners to Form Consortia with Anaerobic Fungi. *Processes*, *6*(1), 7. https://doi.org/10.3390/pr6010007

Williams, C. L., Thomas, B. J., McEwan, N. R., Rees Stevens, P., Creevey, C. J., & Huws, S. A. (2020). Rumen Protozoa Play a Significant Role in Fungal Predation and Plant Carbohydrate Breakdown. *Frontiers in Microbiology*, *11*(April), 1–14. https://doi.org/10.3389/fmicb.2020.00720

Wolfensohn, S. (2004). Nutrient Requirements of Non-Human Primates. In *Laboratory Animals* (Vol. 38, Issue 1). https://doi.org/10.1258/00236770460734461

Wong, M. T., Wang, W., Lacourt, M., Couturier, M., Edwards, E. A., & Master, E. R. (2016). Substrate-driven convergence of the microbial community in lignocellulose-amended enrichments of gut microflora from the Canadian beaver (Castor canadensis) and North American moose (Alces americanus). *Frontiers in Microbiology*, *7*, 961. https://doi.org/10.3389/fmicb.2016.00961

Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N. N., Kunin, V., Goodwin, L., Wu, M., Tindall, B. J., Hooper, S. D., Pati, A., Lykidis, A., Spring, S., Anderson, I. J., Dhaeseleer, P., Zemla, A., Singer, M., Lapidus, A., … Eisen, J. A. (2009). A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature*, *462*(7276), 1056–1060. https://doi.org/10.1038/nature08656

Wu, G., Yan, Q., Jones, J. A., Tang, Y. J., Fong, S. S., & Koffas, M. A. G. (2016). Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends in Biotechnology*, *34*(8), 652–664. https://doi.org/10.1016/j.tibtech.2016.02.010

Xia, W., Liu, G., Wang, D., Chen, H., Zhu, L., & Li, D. (2022). Functional convergence of Yunnan snub-nosed monkey and bamboo-eating panda gut microbiomes revealing the driving by dietary flexibility on mammal gut microbiome. *Computational and Structural Biotechnology Journal*, *20*, 685–699. https://doi.org/10.1016/j.csbj.2022.01.011

Xinyun Qiu, Jingjing Ma, Chunhua Jiao, Xiaqiong Mao, Xiaojing Zhao, Meijiao Lu, Kai Wang, & Hongjie Zhang. (2017). Alterations in the mucosa-associated fungal microbiota in patients with ulcerative colitis. *Oncotarget*, *8*(64), 107577–107588.

Xiros, C., Shahab, R. L., & Studer, M. H. P. (2019). A cellulolytic fungal biofilm enhances the consolidated bioconversion of cellulose to short chain fatty acids by the rumen microbiome. *Applied Microbiology and Biotechnology*, *103*(8), 3355–3365. https://doi.org/10.1007/s00253-019-09706-1

Xu, B., Xu, W., Li, J., Dai, L., Xiong, C., Tang, X., Yang, Y., Mu, Y., Zhou, J., Ding, J., Wu, Q., & Huang, Z. (2015). Metagenomic analysis of the Rhinopithecus bieti fecal microbiome reveals a broad diversity of bacterial and glycoside hydrolase profiles related to lignocellulose degradation. *BMC Genomics*, *16*, 174. https://doi.org/10.1186/s12864-015-1378-7

Yadav, M., Joshi, C., Paritosh, K., Thakur, J., Pareek, N., Masakapalli, S. K., & Vivekanand, V. (2022). Organic waste conversion through anaerobic digestion: A critical insight into

the metabolic pathways and microbial interactions. *Metabolic Engineering*, *69*(November 2021), 323–337. https://doi.org/10.1016/j.ymben.2021.11.014

Yeoman, C. J., Fields, C. J., Lepercq, P., Ruiz, P., Forano, E., White, B. A., & Mosoni, P. (2021). In vivo competitions between fibrobacter succinogenes, ruminococcus flavefaciens, and ruminoccus albus in a gnotobiotic sheep model revealed by multi-omic analyses. *MBio*, *12*(2), 1–16. https://doi.org/10.1128/mBio.03533-20

Yildirim, S., Yeoman, C. J., Sipos, M., Torralba, M., Wilson, B. A., Goldberg, T. L., Stumpf, R. M., Leigh, S. R., White, B. A., & Nelson, K. E. (2010). Characterization of the fecal microbiome from non-human wild primates reveals species specific microbial communities. *PLoS ONE*, *5*(11). https://doi.org/10.1371/journal.pone.0013963

Yu, I. K. M., Chen, H., Abeln, F., Auta, H., Fan, J., Budarin, V. L., Clark, J. H., Parsons, S., Chuck, C. J., Zhang, S., Luo, G., & Tsang, D. C. W. (2021). Chemicals from lignocellulosic biomass: A critical comparison between biochemical, microwave and thermochemical conversion methods. *Critical Reviews in Environmental Science and Technology*, *51*(14), 1479–1532. https://doi.org/10.1080/10643389.2020.1753632

Zeng, Y., Pu, Y., Niu, L. L., Deng, J. B., Zeng, D., Amato, K. R., Li, Y., Zhou, Y., Lin, Y. C., Wang, J., Wu, L. Q., Chen, B. H., Pan, K. C., Jing, B., & Ni, X. Q. (2022). *Comparison of gastrointestinal microbiota in golden snub-nosed monkey (Rhinopithecus roxellanae), green monkey (Chlorocebus aethiops sabaeus), and ring-tailed lemur (Lemur catta) by high throughput sequencing*. *33*. https://doi.org/10.1016/j.gecco.2021.e01946

Zhang, F., Bunterngsook, B., Li, J.-X., Zhao, X.-Q., Champreda, V., Liu, C.-G., & Bai, F.-W. (2019). Regulation and production of lignocellulolytic enzymes from Trichoderma reesei for biofuels production. In Yebo Li & X. Ge (Eds.), *Advances in Bioenergy*, *vol. 4* (pp. 79–119). Elsevier Science. https://doi.org/10.1016/bs.aibe.2019.03.001

Zhang, S., Ning, R., Zeng, B., Deng, F., Kong, F., Guo, W., Zhao, J., & Li, Y. (2022). Gut Microbiota Composition and Metabolic Potential of Long-Living People in China. *Frontiers in Aging Neuroscience*, *14*(July), 1–14. https://doi.org/10.3389/fnagi.2022.820108

Zhang, X., & Figeys, D. (2019). Perspective and Guidelines for Metaproteomics in Microbiome Studies. *Journal of Proteome Research*, *18*(6), 2370–2380. https://doi.org/10.1021/acs.jproteome.9b00054

Zhu, H., Zeng, D., Wang, Q., Wang, N., Zeng, B., Niu, L., & Ni, X. (2018). Diarrhea-associated intestinal microbiota in captive Sichuan golden snub-nosed monkeys

(Rhinopithecus roxellana). *Microbes and Environments*, *33*(3), 249–256. https://doi.org/10.1264/jsme2.ME17163

Zoelzer, F., Burger, A. L., & Dierkes, P. W. (2021). Unraveling differences in fecal microbiota stability in mammals: from high variable carnivores and consistently stable herbivores. Animal Microbiome, 3, 77. https://doi.org/10.1186/s42523-021-00141-0

Zou, Y., Xue, W., Luo, G., Deng, Z., Qin, P., Guo, R., Sun, H., Xia, Y., Liang, S., Dai, Y., Wan, D., Jiang, R., Su, L., Feng, Q., Jie, Z., Guo, T., Xia, Z., Liu, C., Yu, J., … Wang, J. (2019). 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nature Biotechnology*, *37*, 179–185. https://doi.org/10.1038/s41587-018-0008-8

Appendix 1: Use of HEPA-Filtered Anaerobic Glove Bag for Anaerobic Fungal Isolation from Non-Human Primate Fecal Samples

This is a standard operating procedure (SOP) for use of an anaerobic glove bag with HEPA-filtered inflow and outflow, intended for work with fecal samples from NHPs, which must be handled at a minimum of Biosafety Level 2 (BSL-2) (Centers for Disease Control and Prevention and National Institutes of Health, 2020). This appendix has been reproduced, with reformatting for more general use, from the SOP with the same title on the O'Malley Laboratory's shared Google Drive, which should be consulted as the most up-to-date version of this SOP.

The Biological Use Authorization associated with this apparatus and procedure was approved by the UCSB Institutional Biosafety Committee on February 17, 2023.

A1.1. Safety

A1.1.1. Potential hazards and concerns

The primary source of risk to users of this SOP is from samples that must be handled using BSL-2 practices. NHP fecal samples must be handled using BSL-2 practices because of an increased risk of zoonotic infection, due to the close evolutionary relationship between NHPs and humans. Consequently, in the absence of additional information to the contrary, assume any organism capable of infecting and causing illness in a NHP can do the same in a human. Handle your samples with the appropriate care. Prioritize the use of fresh fecal samples from captive primates over wild primates if possible, due to the increased risk of transmission of unknown pathogens from wild primates. Always obtain health and nutrition information, if available, for every NHP individual producing samples.

The user is responsible for familiarizing themselves with potential hazards posed by pathogenic organisms from the gut microbiota of the NHP species from which their sample originates. See **A.1.2., "Specific hazardous chemicals and organisms"**, below for a list of common hazardous organisms in primate fecal samples that may be safely handled at BSL-2. In addition, this SOP was developed and approved for use in a laboratory otherwise approved only for research conducted under Biosafety Level 1 (BSL-1) practices. Ensure all users of this protocol have completed their institution's BSL-2 training prior to use, as well as bloodborne pathogens training if applicable. Always consult your institutional office of environmental health and safety prior to initiating new projects and/or working with new organisms, both animal and microbe, that require the use of this bag.

Proper maintenance of the glove bag is critical to maintaining a safe operating environment during this procedure. Avoid using sharp objects if possible. If a small puncture occurs, seal it with electric tape on the inside and outside of the bag, immediately after treating any injuries to yourself or others. Bring electric tape into the bag with the rest of your equipment and keep a supply of electric tape outside it during the procedure. If using sharps, bring a small sharps disposal container into the bag.

This glove bag (**Figure 40**) is designed to be used with two in-line Whatman HEPA-CAP 36 filters (manufacturer catalog number 6702-3600). These filters retain 99.97% of all particles $\geq$ 0.3 μm, and will generally filter out anything larger than a large virus. These filters must be replaced every 6 months. Designate one filter as the CO_2 inflow filter and one filter as the vacuum outflow filter; these should not be considered interchangeable. The maximum allowable pressure is 60 pounds per square inch (psi), over which the housing will burst; check the pressure from the CO_2 tank in use before beginning the procedure. Both

filters must be autoclaved at 121 ºC (131 ºC maximum) for 20 minutes after each use. Refer to the documentation included with the filter for more information. The HEPA-CAP 36 filters accept tubing from 1/4 to 3/8 inches in internal diameter. The 10 ml plastic serological pipette tubes (Corning, catalog number 4101) used in the original construction of this glove bag accept tubing 1/4 inches in internal diameter, and may be replaced by carefully breaking off the ends of one serological pipette, breaking the pipette in half, smoothing the broken-off ends by melting over a Bunsen burner flame, and sealing one end to the glove bag with electric tape.

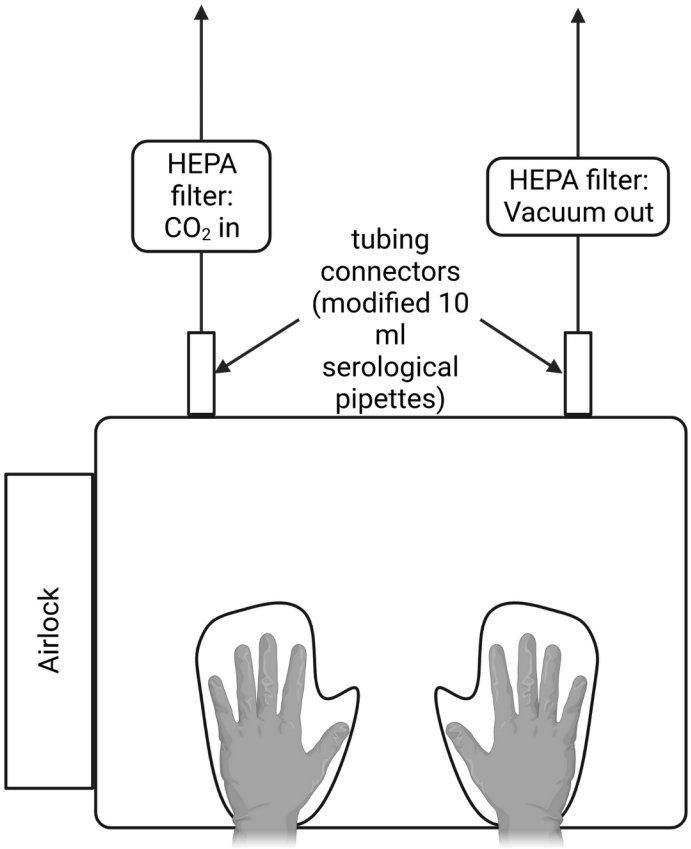


Figure 40. Anaerobic glove bag schematic. Not to scale. Made with Biorender.

A1.1.2. Specific hazardous chemicals and organisms

Chloramphenicol: MSDS Sigma Aldrich (2023). Final concentration 0.025 mg/ml in liquid or agar medium. Chloramphenicol is used here as a broad-spectrum antibacterial to enrich fungi. The PPE required for this SOP provide adequate protection against chloramphenicol exposure. Chloramphenicol is a carcinogen, and is moderately toxic after acute exposure, which may occur through ingestion, inhalation, and skin and eye contact. Acute digestive tract exposure can result in gastrointestinal irritation with nausea, vomiting, and diarrhea. Exposure may cause liver damage, digestive tract hemorrhage, and anemia and other blood abnormalities. **First aid** for chloramphenicol exposure depends on the route of exposure, and all exposures require immediate physician consultation in addition to specific advice for each route. For ingestion, drink 1-2 glasses of water; do not induce vomiting without medical advice. For eye exposure, immediately flush eyes with copious amounts of water for 15 minutes, preferably in an emergency eyewash. For skin exposure, immediately wash affected area with soap and copious amounts of water for 15 minutes.

Clarified rumen fluid (Medium C component): This substance is non-hazardous after autoclaving. The full recipe of Medium C may be found in Theodorou, Brookman, and Trinci's published anaerobic fungal isolation and culture methods (2005).

Pressurized CO_2 gas: MSDS Airgas USA, LLC (2018). CO_2 poses an asphyxiation risk. If asphyxiation occurs, ensure CO_2 and vacuum taps are closed, remove the victim to fresh air, and close the glove bag to prevent the escape of organisms within the glove bag.

Common hazardous organisms in fresh primate feces handled at BSL-2: Most of these pathogens are transmitted through the fecal-oral route.

- **Roundworms** (Phylum Nematoda)
- **Protozoans:** *Entamoeba histiolytica*, *Giardia lamblia*
- **Pathogenic fungi:** *Cryptosporidium* spp.
- **Pathogenic bacteria:** *Campylobacter* spp., *Shigella* spp., *Salmonella* spp.
- **Pathogenic viruses:**
    - Hepatitis A and B: great apes only, immunization available.
    - Simian immunodeficiency virus (SIV): Causes an asymptomatic and non-persistent infection in humans. BSL-2 is permitted for handling of clinical specimens, such as feces, containing SIV; more concentrated samples require BSL-3, or BSL-2 with BSL-3 practices. Captive primates are unlikely to carry this infection, but verify with the primates' keepers prior to obtaining samples. Prioritize using primate samples from captive animals over wild animals, if possible; if this is unavoidable and if you can treat them to eliminate the risk of SIV transmission, do so.

A1.1.3. Personal protective equipment

Use of the HEPA-filtered anaerobic glove bag requires standard personal protective equipment (PPE) for BSL-2 operations, including but not necessarily limited to the following:

- **Protective eyewear**: safety glasses at a minimum, or safety goggles.
- **Barrier laboratory coat**: made of hydrophobic fabric, with elastic cuffs around the wrist and snap closures in front. Fasten all snaps.
- **Nitrile gloves**: stretch glove cuffs over the cuffs of the barrier laboratory coat.

A1.2. Protocol

This protocol is designed to be used as a companion to Theodorou, Brookman, and Trinci's protocol (2005), or modifications thereof, detailing the isolation of anaerobic fungi from inoculum (digesta or feces).

A1.2.1. Materials

Inside bag:

- Small bag weights
- Small plastic bin to hold tools and 70% ethanol
- 70% ethanol in spray bottle
- Rack of 50 ml Falcon tubes containing 97-100% ethanol (if preservation of samples is desired)
- Paper towels
- Electric tape (a few pieces)
- Samples
- Rack of Hungate tubes containing Medium C (or other appropriate medium) with 0.025 mg/ml chloramphenicol and carbon substrate (additionally, if obtaining axenic cultures of isolates: roll tubes containing Medium C agar with 0.025 mg/ml chloramphenicol and glucose as a carbon source)
- Sterile loops
- Sealable waste container lined with BSL-2 rated biohazard bag

Outside bag: Electric tape (a few pieces)

A1.2.2. Procedure

Safety controls: If carrying out this procedure in a lab otherwise rated BSL-1, it is recommended to notify all laboratory personnel of when and where this protocol will be carried out, with appropriate signage posted if necessary. Avoid using this glove bag alone if possible, to facilitate a quick response to hazards both inside and outside the bag and to monitor gas pressure.

Decontamination: Completely wipe inside and outside of bag with 70% ethanol.

Plumbing: Ensure dedicated HEPA-filtered inflow is connected to the CO_2 supply and the HEPA-filtered outflow is connected to vacuum. Verify that arrows on inflow and outflow filters are oriented in the direction of gas flow.

Items to place in the bag before flushing: Small bag weights, a small waste container lined with a BSL-2 biohazard bag, and all experimental items other than cultures and items that must be kept warm; place all items in a small plastic bin.

Flushing: Prior to flushing, place weights in glove bag and on gloves to prevent them from inverting. Monitor bag to ensure gloves do not invert while bag is being flushed. Flush the bag, starting with vacuum, then completing 3 cycles of CO_2 and vacuum. Inflate the bag to at least 30 cm high each time.

Items to place in the bag after flushing: Live cultures and items that must be kept warm.

In use: Inoculate a small amount (< 1 g) of fecal sample into pre-warmed Hungate tubes containing medium, chloramphenicol, and carbon substrate. For published axenic isolation

procedure, refer to Theodorou, Brookman, and Trinci's protocol (Theodorou, Brookman, and Trinci, 2005). Open the zip lock as little as possible. Dispose of all biological waste in a BSL-2 biohazard bag.

At the end of experimental work: Shut off CO_2 and vacuum, and use large bag clips to clip gas inflow and outflow lines shut in front of the tape attaching the serological pipettes to the glove bag. With the bag still inflated, remove one arm from the bag glove and run your gloved hand over the outside surface of the bag, pressing to check for leaks. Inspect tubing, serological pipette hardware, and electrical tape for any holes or weak spots. Close all live culture containers. Flush the bag 3 times, starting with vacuum, then performing 3 cycles of CO_2 and vacuum. Inflate the bag to at least 30 cm high each time. Wipe down the outside of the culture containers and all equipment in bag with 70% ethanol. Completely wipe inside and outside of bag with 70% ethanol. Seal biohazard waste container and remove all cultures and equipment from glove bag. After storing bag, dispose of biohazard waste as appropriate and wash hands and forearms with antibacterial soap after doffing PPE.

A1.2.3. Hazardous chemical and organism disposal

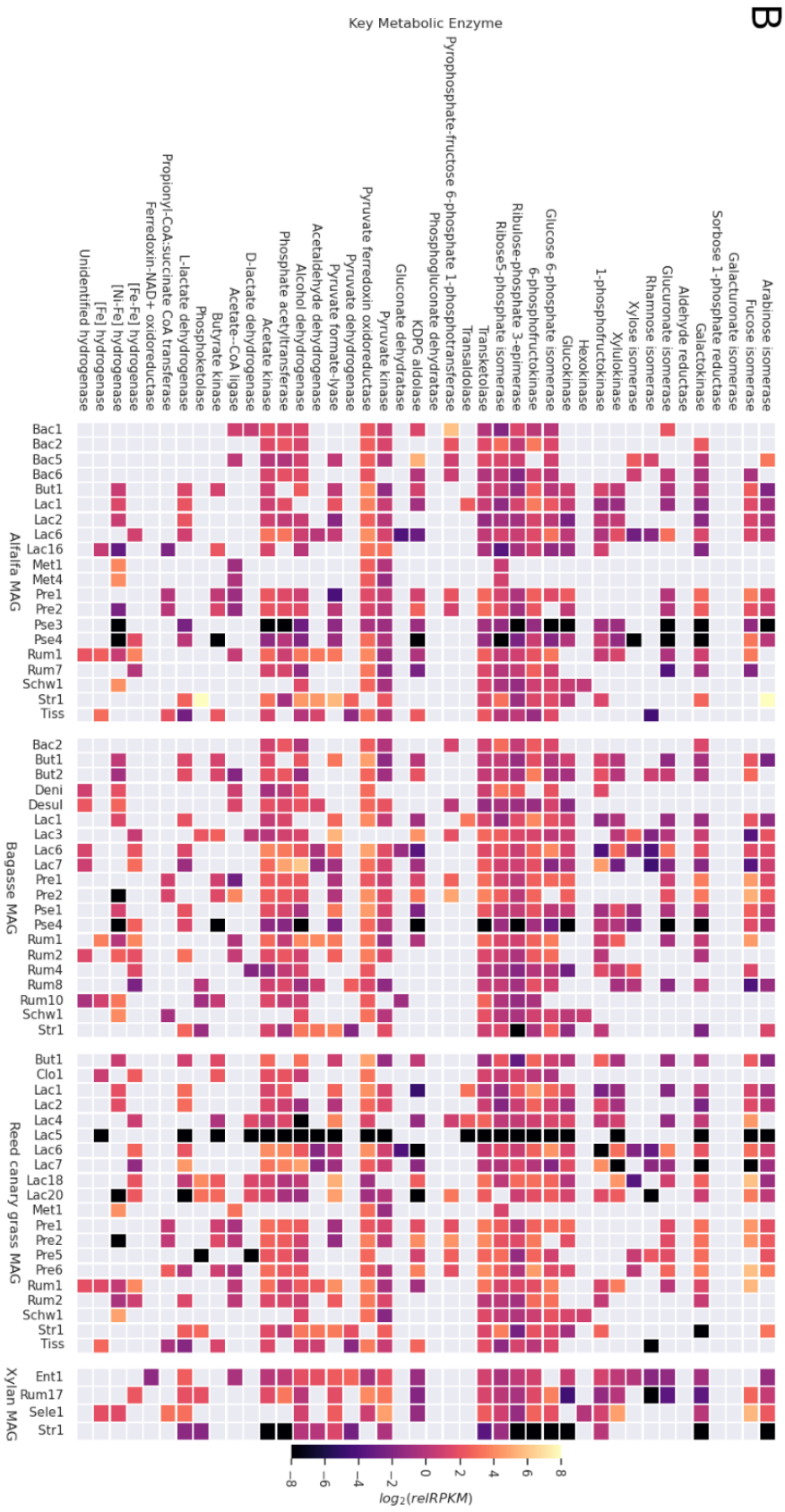
Chloramphenicol (0.025 mg/ml in medium): All volumes used are < 1 L. Decontaminate equipment by wiping with an ethanol-soaked paper towel. For spills in the glove bag, there is no splash hazard if the glove bag is closed, but dispose of excess spilled medium as listed under "liquid medium containing organisms" entry under **Organisms**, then wipe inside of bag with an ethanol-soaked paper towel. Medium containing dilute concentrations of chloramphenicol is considered non-hazardous and does not require special disposal other than that listed below under **Organisms**.

Clarified rumen fluid: See chloramphenicol entry above for disposal as part of chloramphenicol-containing media.

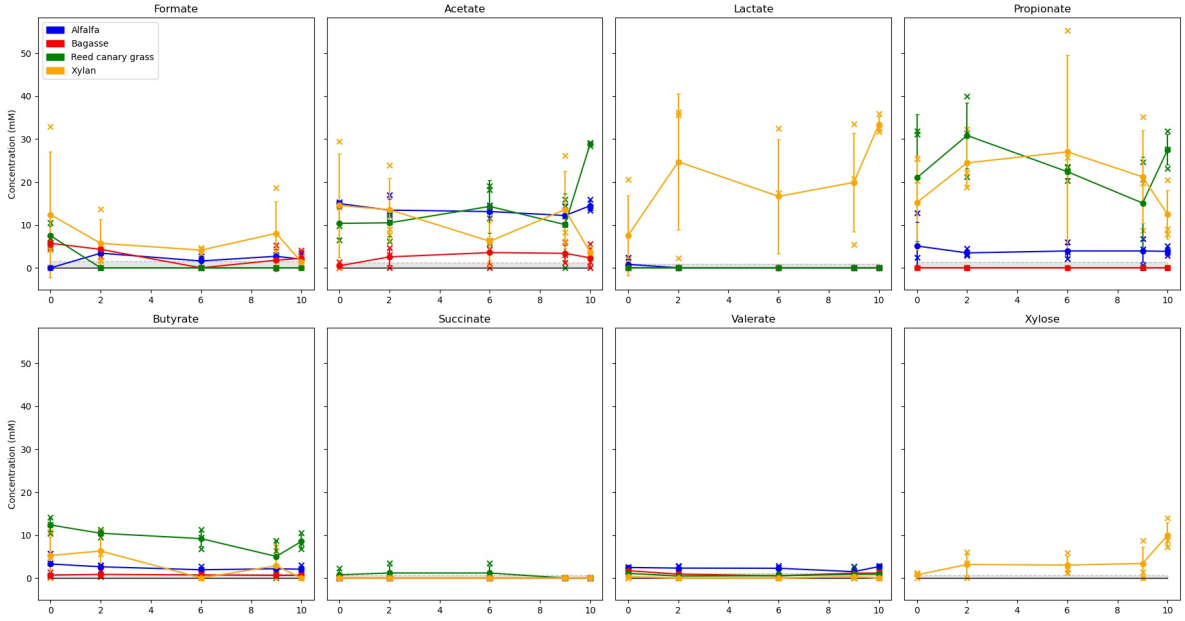
CO_2: Contaminated gas in the glove bag is disposed of through the HEPA-filtered vacuum line during the flushing process at the end of the procedure.

Organisms in fresh feces: If not preserving for future use, place in a BSL-2 rated biohazard waste bag, seal, and place in secondary containment. Sterilize in an autoclave rated for BSL-2 waste at a minimum of 121 ºC for at least 30 minutes. If preserving in 97-100% ethanol, freeze overnight at -80 ºC before continuing use. Preservation in RNAlater™, or similar solutions intended to preserve RNA, will not sufficiently disinfect the sample due to insufficient deactivation of viral pathogens (Uhlenhaut & Kracht, 2005); consult your institutional environmental health and safety office for further information.

Appendix 2: Chapter 2 Supplementary Material


B
Key Metabolic Enzyme
Arabinose isomerase
Fucose isomerase
Galacturonate isomerase
Sorbose 1-phosphate reductase
Galactokinase
Aldehyde reductase
Glucuronate isomerase
Rhamnose isomerase
Xylose isomerase
Xylulokinase
1-phosphofructokinase
Hexokinase
Glucokinase
Glucose 6-phosphate isomerase
6-phosphofructokinase
Ribulose-phosphate 3-epimerase
Ribose5-phosphate isomerase
Transketolase
Transaldolase
Phosphogluconate dehydratase
KDPG aldolase
Gluconate dehydratase
Pyruvate kinase
Pyruvate ferredoxin oxidoreductase
Pyruvate dehydrogenase
Pyruvate formate-lyase
Acetaldehyde dehydrogenase
Alcohol dehydrogenase
Phosphate acetyltransferase
Acetate kinase
D-lactate dehydrogenase
Acetate--CoA ligase
Butyrate kinase
Phosphoketolase
L-lactate dehydrogenase
Propionyl-CoA:succinate CoA transferase
Ferredoxin-NAD+ oxidoreductase
[Fe-Fe] hydrogenase
[Ni-Fe] hydrogenase
[Fe] hydrogenase
Unidentified hydrogenase
Pyrophosphate-fructose 6-phosphate 1-phosphotransferase
Alfalfa MAG
Bac1
Bac2
Bac5
Bac6
But1
Lac1
Lac2
Lac6
Lac16
Met1
Met4
Pre1
Pre2
Pse3
Pse4
Rum1
Rum7
Schw1
Str1
Tiss
Bagasse MAG
Bac2
But1
But2
Deni
Desul
Lac1
Lac3
Lac6
Lac7
Pre1
Pre2
Pse1
Pse4
Rum1
Rum2
Rum4
Rum8
Rum10
Schw1
Str1
Reed canary grass MAG
But1
Clo1
Lac1
Lac2
Lac4
Lac5
Lac6
Lac7
Lac18
Lac20
Met1
Pre1
Pre2
Pre5
Pre6
Rum1
Rum2
Schw1
Str1
Tiss
Xylan MAG
Ent1
Rum17
Sele1
Str1
−8
−6
−4
−2
0
2
4
6
8
log2(relRPKM)

Supplemental Figure 1 (preceding two pages). Heatmaps of relative (A) CAZyme and (B) metabolic enzyme expression in key MAGs across the carbohydrate substrates alfalfa, bagasse, reed canary grass, and xylan. log_2(relRPKM) = log_2[(RPKM for designated ORF)/(median of all expressed RPKMs for that MAG in the sample)]. Cell values are means across all biological samples.



Supplemental Figure 2. High-pressure liquid chromatography (HPLC) of short-chain fatty acids produced by all consortia. Generation number is plotted on the x-axis. Error bars are standard deviation; individual points are biological replicates.

Appendix 3: Chapter 3 Supplementary Material

Primer	Region amplified	Forward/ reverse	Sequence (5'-3')
S-D-Bact-0341-b-S-17	16S V3/V4	forward	CCTACGGGNGGCWGCAG
S-D-Bact-0785-a-A-21	16S V3/V4	reverse	GACTACHVGGGTATCTAATCC
gITS7ngs	ITS2	forward	GTGARTCATCRARTYTTTG
ITS4ngs	ITS2	reverse	CCTSCSCTTANTDATATGC
ITS9	ITS2	forward	GAACGCAGCRAAIIGYGA
Forward Illumina adapter overhang	N/A	forward	TCGTCGGCAGCGTCAGATGTG TATAAGAGACAG
Reverse Illumina adapter overhang	N/A	reverse	GTCTCGTGGGCTCGGAGATGT GTATAAGAGACAG

Supplemental Table 3. Primers and adapter overhangs used for 16S and ITS2 amplicon sequencing.

Replicate	Species	Individual	Amplicon	Shannon entropy
CT1R1-16S_S22_L001	*Colobus guereza*	Pooled sample	16S	7.25865171197553
CT1R2-16S_S23_L001	*Colobus guereza*	Pooled sample	16S	7.16577210562216
CT1R3-16S_S24_L001	*Colobus guereza*	Pooled sample	16S	7.78477348022494
CT2R1-16S_S25_L001	*Colobus guereza*	Pooled sample	16S	7.79802741297074
CT2R2-16S_S26_L001	*Colobus guereza*	Pooled sample	16S	7.34890499889484
CT2R3-16S_S27_L001	*Colobus guereza*	Pooled sample	16S	7.57225788548139
CT3R1-16S_S28_L001	*Colobus guereza*	Pooled sample	16S	7.56669486933197
CT3R2-16S_S29_L001	*Colobus guereza*	Pooled sample	16S	7.82536410197377
CT3R3-16S_S30_L001	*Colobus guereza*	Pooled sample	16S	7.99218898432717
CT1R1-ITS2_S84_L001	*Colobus guereza*	Pooled sample	ITS2	3.42497320433506
CT1R2-ITS2_S85_L001	*Colobus guereza*	Pooled sample	ITS2	3.18558870389873
CT1R3-ITS2_S86_L001	*Colobus guereza*	Pooled sample	ITS2	3.35250669544849
CT2R1-ITS2_S87_L001	*Colobus guereza*	Pooled sample	ITS2	2.76217955217194
CT2R2-ITS2_S88_L001	*Colobus guereza*	Pooled sample	ITS2	2.36979083560054
CT2R3-ITS2_S89_L001	*Colobus guereza*	Pooled sample	ITS2	2.43511063022068
CT3R1-ITS2_S90_L001	*Colobus guereza*	Pooled sample	ITS2	2.22406998357486
CT3R2-ITS2_S91_L001	*Colobus guereza*	Pooled sample	ITS2	3.05307228797791
CT3R3-ITS2_S92_L001	*Colobus guereza*	Pooled sample	ITS2	1.5549688657657
BT2R1-16S_S4_L001	*Gorilla gorilla gorilla*	Bangori	16S	7.84145795078618
BT2R2-16S_S5_L001	*Gorilla gorilla gorilla*	Bangori	16S	7.81310244543922
BT2R3-16S_S6_L001	*Gorilla gorilla gorilla*	Bangori	16S	7.95126178576695

NT4R1-16S_S1_L001	*Gorilla gorilla gorilla*	Nzinga	16S	7.33294937616163
NT4R2-16S_S2_L001	*Gorilla gorilla gorilla*	Nzinga	16S	7.38799351896448
NT4R3-16S_S3_L001	*Gorilla gorilla gorilla*	Nzinga	16S	7.37147736306886
BT2R1-ITS2_S66_L001	*Gorilla gorilla gorilla*	Bangori	ITS2	3.74949186200393
BT2R2-ITS2_S67_L001	*Gorilla gorilla gorilla*	Bangori	ITS2	4.66354620991559
BT2R3-ITS2_S68_L001	*Gorilla gorilla gorilla*	Bangori	ITS2	2.01538559113712
NT4R1-ITS2_S63_L001	*Gorilla gorilla gorilla*	Nzinga	ITS2	4.70560483016711
NT4R2-ITS2_S64_L001	*Gorilla gorilla gorilla*	Nzinga	ITS2	2.59691818114889
NT4R3-ITS2_S65_L001	*Gorilla gorilla gorilla*	Nzinga	ITS2	3.71598397853788

Supplemental Table 4. Table of Shannon entropies for each primate fecal sample. All colobine samples of each amplicon are normalized to other colobine samples with that amplicon, and all gorilla samples of each amplicon are normalized to other gorilla samples with that amplicon (both Bangori and Nzinga combined).

Citation	Species	Region(s) sequenced	Primers used for fungi (forward/reverse)
Mar Rodriguez et al 2015, Scientific Reports 5: 14600	*Homo sapiens*	ITS1+2	Primer set A; H1SeqF GTCATTTAGAGGAAGTAAAA GTCGTAACAAGG and H1SeqRb GCTRYGTTCTTCATCGDTGC. Primer set B; H2SeqFb GCA TCG ATG AAG AAC RYA GC and primerH2SeqRb TTC TTT TCC TCC GCT TAT TGA TAT GC.
Schulz et al 2018, Int Journal of Primatology 39(4): 567-580	*Gorilla gorilla gorilla*	ITS1	ITS1F/NeoQ PCR R
Pierre et al 2021, Journal of Surgical Research 267: 336-341	*Macaca mulatta*	16S, ITS1+2	“Next Generation Sequencing"
Raimondi et al 2019, Frontiers in Microbiology 10: 1575	*Homo sapiens*	ITS1	18SF/5.8S1R
Barelli et al 2020,	*Procolobus gordonorum,*	16S,	5'-

mSystems 5(3): e00061-20	*Papio cynocephalus Alouatta seniculus, A. pigra, Lagothrix lagotricha, Ateles hybridus, At. belzebuth, Gorilla gorilla, Pan troglodytes, Papio hamadryas, P. anubis, Theropithecus gelada, Cercopithecus ascanius, Colobus guereza, Piliocolobus badius, Propithecus verreauxi, Eulemur rubriventer*	ITS1+2	GCATCGATGAAGAACGCAGC-3'/5'-TCCTCCGCTTATTGATATGC-3'
Mann et al 2019, The ISME Journal 14: 609-622		18S	E572F/E1009R
Schei et al 2017, Microbiome 5: 107	*Homo sapiens*	ITS1	ITS1F/ITS2
James et al 2022, Journal of Fungi 8: 1054	*Macaca fascicularis*	ITS1	ITS1F/ITS2
Taylor et al 2018, PeerJ 6: e4612	*Macaca mulatta*	ITS1	ITS1F/ITS2
Sun et al 2021, Frontiers in Microbiology 12: 730477	*Macaca thibetana*	ITS1	ITS1F/ITS2
Murillo et al 2022, ISME Communications 2:3	*Eulemur rufifrons*	16S, 18S, ITS2	ITS3_KYO2/ITS4
Sharma et al 2022, npj Biofilms and Microbiomes 8: 12	*Homo sapiens, Pan troglodytes schweinfurthii, Gorilla gorilla gorilla, G. gorilla beringei, Cercocebus agilis*	16S, ITS2	ITS3/ITS4
Auchtung et al 2018, mSphere 3(2): e00092-18	*Homo sapiens*	ITS2	ITS3F/ITS4R
Sawaswong et al 2020, Fungal Genetics and Biology 144: 103468	*Macaca fascicularis*	ITS2	ITS86F/ITS4R

Hamad et al 2014, Scientific Reports 4: 6417	*Gorilla gorilla gorilla*	18S rRNA, ITS1+2	NSI/FRI, Euk1A/Euk516r, ITS F/ITS4R, E528F/Univ1492RE, 121F/1147R, FunF/FunR
Barelli et al 2021, Scientific Reports 11: 21569	*Procolobus gordonorum, Papio cynocephalus*	16S, ITS1+2	unclear
Scanlan and Marchesi 2008, ISME Journal 2(12): 1183-1193	*Homo sapiens*	ITS1+2	ITS1F/ITS4R

Supplemental Table 5. Table of ITS primers from 16 studies of primate gut microbiota including ITS amplicon sequence variant analysis from 2014-present.