UC Santa Barbara
UC Santa Barbara Electronic Theses and Dissertations

Title

Deciphering the Functions of Natural Products from Anaerobic Fungi for Applications in Biotechnology

Permalink

https://escholarship.org/uc/item/76k6g0bp

Author

Swift, Candice Lee

Publication Date

2020

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Santa Barbara

Deciphering the Functions of Natural Products from Anaerobic Fungi for Applications in

Biotechnology

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Chemical Engineering

by

Candice Lee Swift

Committee in charge:

Professor Michelle A. O'Malley, Chair

Professor Siddharth Dey

Professor Song-I-Han

Professor David Low

December 2020

The dissertation of Candice Lee Swift is approved.

_____

Siddharth Dey

_____

Song-I Han

_____

David Low

_____

Michelle A. O'Malley, Committee Chair

December 2020

ACKNOWLEDGEMENTS

I would like to thank my advisor, Michelle O'Malley, for the opportunity to work in her lab and achieve the work described in this dissertation. I would also like to thank all O'Malley lab members, both past and present. I am extremely grateful for the training I received from others in the lab and for the foundation that previous lab members laid in isolating anaerobic fungi and sequencing their genomes and transcriptomes. I also want to thank Nikola Malinov, the diligent undergraduate researcher whose help in lab was invaluable to me and who contributed to the experiments in Chapter five.

I would also like to thank our collaborators at the Joint Genome Institute, the Environmental Molecular Sciences Laboratory, and both the Keller lab at the University of Wisconsin-Madison as well as the Solomon lab at Purdue for their experimental contributions and insight. Dr. Jennifer Smith at the Biological Nanostructures Laboratory at UCSB was particularly helpful to me during the library preparation steps of Chapter five.

I am grateful to professors Jennifer Reed, Nicholas Abbott, and Betty Chewning for writing letters of recommendation for me when I applied to graduate school after working for five years in industry.

Finally, I want to thank my family and friends for their encouragement and support, as well as my husband and son.

VITA OF CANDICE LEE SWIFT

December 2020

EDUCATION

2015-2020   Doctor of Philosophy in Chemical Engineering
             University of California – Santa Barbara

2006-2010   Bachelor of Science in Chemical Engineering
             University of Wisconsin – Madison

PUBLICATIONS

1. **C.L. Swift**, J.L. Brown, S. Seppälä, M.A. O'Malley, "Co-cultivation of the anaerobic fungus *Anaeromyces robustus* with *Methanobacterium bryantii* enhances transcription of carbohydrate active enzymes," Journal of Industrial Microbiology and Biotechnology. 46, 1427-1433, 2019.
2. T.A. Rush, V. Puech-Pagès, A. Bascaules, P. Jargeat, F. Maillet, A. Haouy, A. QuyManh Maës, C.C. Carriel, D. Khokhani, M. Keller-Pearson, J. Tannous, K.R. Cope, K. Garcia, J. Maeda, C. Johnson, B. Kleven, Q.J. Choudhury, J. Labbé, **C.L. Swift**, M.A. O'Malley, J.W. Bok, S. Cottaz, S. Fort , V. Poinsot, M.R. Sussman, C. Lefort, J. Nett, N.P. Keller, G. Bécard, J Ané, "Lipo-chitooligosaccharides as regulatory signals of fungal growth and development," *Nature Communications*, 11(1), 2020.
3. I.A. Podolsky, S. Seppala, T.S. Lankiewicz, J.L. Brown, **C.L. Swift,** M.A. O'Malley, "Harnessing Nature's Anaerobes for Biotechnology and Bioprocessing," *Annual Review of Chemical and Biomolecular Engineering* 13(27), 6.1-6.24, 2019.
4. S.E. Wilken, **C.L. Swift,** I.A. Podolsky, T.S. Lankiewicz, S. Seppala, M.A. O'Malley, "Linking 'omics' to function unlocks the biotech potential of non-model fungi," *Current Opinion in Systems Biology* 14, 9-17, 2019.
5. X. Peng, **C.L. Swift**, M.K. Theodorou, M.A. O'Malley, in 53–67 (Humana Press, New York, NY, 2018). doi:10.1007/978-1-4939-7804-5_5
6. R.J. Carlton, J.K. Gupta, **C.L. Swift,** N.L. Abbott, "Influence of Simple Electrolytes on the Orientational Ordering of Thermotropic Liquid Crystals at Aqueous Interfaces," *Langmuir* 28(1), 31-36, 2012.

PLANNED PUBLICATIONS

1. X. Peng, X., S.E. Wilken, T.S. Lankiewicz, S.P. Gilmore1, J.L. Brown, J.K. Henske, **C.L. Swift**, A. Salamov, K.Barry, I.V. Grigoriev, M.K. Theodorou5, D.L. Valentine6, M.A. O'Malley, "Sculpting gut microbial communities alters fermentation products and methane release," *Nature Microbiology*, **In press 2020.**

2. C.L. Swift, K.B. Louie, B.P. Bowen, H.M. Olson, S.O. Purvine, A. Salamov, S.J. Mondo, K.V. Solomon, A.T. Wright, T.R. Northen, I.V. Grigoriev, N.P. Keller, M.A. O'Malley, "Anaerobic gut fungi are an untapped reservoir of natural products," **In progress.**
3. C.L. Swift, N.G. Malinov, S.J. Mondo, A. Salamov, I.V. Grigoriev, M.A. O'Malley, "Anaerobic fungal genomes encode stress response genes similar to early eukaryotic parasites," **In progress.**
4. C.L Swift, K.B. Louie, B.P. Bowen, C.A. Hooker, K.V. Solomon, T.R. Northen, I.V. Grigoriev, M.A. O'Malley, "Co-cultivation of anaerobic fungi with rumen bacteria establishes an antagonistic relationship," **In progress.**
5. 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

Swift CL, O'Malley MA. "Discovery of natural products from anaerobic gut fungi," oral and poster presentations at the 30th Fungal Genetics Conference; 2019 Mar 14; Pacific Grove, California.

Swift CL, Louie KB, Bowen BP, Bingol K, Heyman HM, Northen TR, O'Malley MA. "Enabling natural products discovery from anaerobic fungi via integrated 'omics' approaches," oral presentation at the 255th ACS National Meeting, Division of Biochemical Technology (BIOT); 2018 Mar 22; New Orleans, Louisiana.

Swift CL, O'Malley MA. "Deciphering the role of fungal secondary metabolites in anaerobic microbial communities," oral presentation at Joint Genome Institute 2017 Microbial and Plant Systems Modulated by Secondary Metabolites Meeting; 2017 July; Walnut Creek, California.

Swift CL, O'Malley MA, Louie KB, Northen TR. "Identifying and characterizing the secondary metabolites of anaerobic fungi," poster presentation at 253rd ACS National meeting, Division of Biochemical Technology (BIOT); 2017 Apr; San Francisco, California.

SELECTED AWARDS

Graduate Division Dissertation Fellowship (2020)
Connie Frank Fellowship (2019)
MIT CHEME Rising Stars (2019)
National Science Foundation Graduate Research Fellowships Program (2016-2019)
Genetics Society of America member
Best poster at Clorox-Amgen Graduate Student Symposium (October 2018)
Runner-up for best poster at the 50th anniversary of the UCSB Chemical Engineering department
Undergraduate scholarships (2006-2010): American Chemical Society Scholar, Chancellor's Scholar, Leaders in Engineering Excellence and Diversity

PROFESSIONAL EMPLOYMENT

2014-2015   Veolia Water Solutions & Technologies

2013-2014   Valero Energy Corporation

2010-2012   Honeywell UOP

ABSTRACT

Deciphering the Functions of Natural Products from Anaerobic Fungi for Applications in

Biotechnology

by

Candice Lee Swift

Natural products, or secondary metabolites, are small molecules produced primarily by

bacteria, fungi, and plants. Their chemical diversity has conferred many usefuiiul

bioactivities, such as antibacterial, antifungal, cholesterol-lowering, and

immunosuppressant. In addition to their value as medicinal drugs, these molecules also serve

important ecological roles in their native environments, by mediating interactions between

microorganisms. Natural products are crucial to future microbiome engineering efforts for

their natural ability to modulate microbial communities. The digestive tracts of large

herbivores, including the rumen, harbor complex microbial communities, consisting of

fungi, bacteria, protozoa, and methanogenic archaea. This vast reservoir of chemical

diversity is almost completely untapped with regards to the discovery of natural products.

These microbiomes are of interest both for their potential to yield new drug candidates and

as a microbial platform for chemical production from renewable feedstocks.

In this work, we explore the biosynthetic potential of anaerobic fungi native to the

digestive tracts of herbivores to synthesize natural products. The genomes, transcriptomes,

proteomes, and metabolomes of anaerobic fungi reveal that these underexplored fungi synthesize novel natural products, some of which are used to compete with rumen bacteria. Groundbreaking dual transcriptomics of rumen fungi co-cultured with rumen bacteria provides evidence that the relationship between these two organisms is antagonistic and that the presence of bacteria stimulates the expression of biosynthetic genes encoding enzymes that synthesize potential antibiotics. By sequencing the active genes of microbiomes inoculated from a source microbiome (goat fecal pellet), we demonstrated that the bacterial and fungal biosynthetic genes of natural products were active in sequential cultures passaged *in vitro*. We found that fungal biosynthetic genes were upregulated at later generations of batch passaging. This finding revealed that the expression of fungal biosynthetic genes is dynamic and suggested that natural products may function in the stability of the microbial community. Overall this dissertation points to the potential of natural products sourced from the rumen as drug candidates, as well as their importance in future microbiome engineering efforts.

TABLE OF CONTENTS

1 Introduction

1.1 Motivation

Natural products have been the basis for many medicinal drugs over the past century. Analysis of FDA-approved new molecular entities (NMEs) from 1931 to 2013 revealed that natural products and their derivatives consistently supplied roughly 30-50% of NMEs for each decade[1]. During this time period, more than half of these natural product-based NMEs were sourced from bacteria and fungi, with fungal NMEs comprising 23% of the total[1]. Most antibiotics in use today were discovered from the secondary metabolites of microorganisms, such as actinomycete soil bacteria[2]. Despite this history of importance in drug discovery, natural products research by pharmaceutical companies has diminished because of a high rate of rediscovery of compounds with the desired bioactivities[1,3].

Next-generation sequencing of microbial genomes strongly supports that that the capacity for discovery of novel natural products is far from being reached[4]. In fact, some microbial environments, such as the digestive tracts of large herbivores, have been almost completely neglected by natural products research until recently. The rumen microbiome consists of a complex membership of bacteria, fungi, methanogenic archaea, and protozoa[5]. Rumen fungi, also called anaerobic gut fungi, are outnumbered by bacteria in the consortia by four orders of magnitude[5], but they are indispensable contributors to the breakdown of biomass consumed by the host using a mechanism of biomass-breakdown distinct from bacteria[6,7]. By mining the genomes of anaerobic gut fungi, we detected the presence of genes encoding the biosynthetic enzymes of natural products, which suggested that anaerobic gut fungi are capable of synthesizing natural products of unknown function.

In addition to their potential application as antimicrobials, natural products facilitate interactions between microbes and thus contribute to microbial community structure and dynamics[8]. Understanding how microbial consortia are assembled and maintained in nature is a key milestone towards engineering the performance of microbiomes. The rumen microbiome is of particular interest for biotechnological applications for its ability to efficiently degrade non-food lignocellulosic biomass into sugars and then ferment them into short chain fatty acids, carbon dioxide, and methane[5]. Resolving the roles of natural products in microbial interactions within the rumen is a necessary step to enable manipulation of the system to achieve desired outcomes, such as improved host health, biomass degradation, or product selectivity. More broadly, natural products may be essential in the design of stable synthetic microbial consortia.

1.2   *Historical and current significance of natural products*

Natural products are generally considered to be small (less than 2000 Da), often nonpolar molecules produced by organisms such as plants, bacteria, and fungi, but not strictly necessary for the growth, survival and viability of the producer[9]. Natural products may be of medicinal value when unmodified, as a semisynthetic derivative, or as the conceptual basis for a synthetic structure[1]. Humans have benefited from natural products for thousands of years, with the Ebers Papyrus recording the medicinal use of over 700 plant-based products in Egypt as early as 2900 B.C.E.[10]. Examples of compounds with medicinal use and their natural basis are given in Table 1.1. Since the discovery of penicillin from the fungus *Penicillium notatum* by Alexander Fleming in 1929[10], natural products have been one of the main sources of antibiotics. The diverse classes of natural products include polyketides, non-ribosomal peptides, terpenes, bacteriocins, and hybrids of these classes[11]. Certain classes of

natural products, such as polyketides, have been exceptionally successful in medicinal application because of their extensive chemical diversity. During the period from 2005-2007, more than a third of approved drugs derived from natural products were polyketides, resulting in annual sales for pharmaceutical companies of $20 million[12].

Table 1.1: Examples of medicinal applications of compounds derived from natural sources, discussed by Berkov and colleagues in *Biotechnology Advances*[13].

Compound	Use	Natural basis
Penicillin	Antibiotic	Fungus
Lovastatin	Cholesterol-lowering	Fungus
Paclitaxel (Taxol®)	Anticancer	Plant
Trabectedin (Yondelis®)	Anticancer	Tunicate
Ingenol mebutate	Anticancer	Plant
Elsamicin A	Anticancer	Bacteria

1.2.1   *Mitigating the problem of antibiotic resistance through the discovery of novel antimicrobial compounds from the rumen microbiome*

Natural products have contributed significantly towards human health, especially in their application as antibiotics. However, antibiotic resistance is one of the greatest challenges facing humanity today. Mitigating the problem of antibiotic resistance requires a multi-faceted approach that includes many components, such as (1) reduction of agriculture use and inappropriate clinical use through legislation and education[14], (2) improved clinical antimicrobial susceptibility testing that better mimics *in vivo* infection[15], and (3) discovery of antibiotics with new modes of action or even dual modes of action[16]. One strategy to discover novel molecules with potential antimicrobial activity is to explore the secondary metabolism of microorganisms in previously untapped environments.

The rumen microbiome has been neglected as a source of novel antibiotics, with the exception of the discovery of a few antimicrobial peptides. The bacteriocins butyrivibriocin OR79A[17,18], butyrivibriocin AR10[17], and the peptide Lynronne-1, which decreased the

bacterial count of methicillin-resistant *Staphylococcus aureus* (MRSA) in a mouse model[19], were discovered from the rumen microbiome (discussed further in Chapter six).

To date, no fungal antibiotics or antimicrobials have been discovered from the rumen microbiome. Some fungal antibiotics are toxic to mammals and therefore not useful in clinical application[20]. However, rumen fungi are attractive as sources of novel antibiotics, since whatever antibiotics they may produce cannot be toxic to their mammalian host, at least in the native concentrations. Secondary metabolites from rumen fungi might also have an immunosuppressive effect on the host.

1.2.2 *Anaerobic fungi and the rumen microbiome as a platform for natural product discovery*

Fungi are prodigious producers of diverse classes of chemical compounds. Fungal natural products have served humanity as antibiotics, anti-tumor agents, insecticides and could even serve as drop-in biofuels[21,22]. The genome of a single species such as *Aspergillus oryzae* encodes 30 PKS genes and 18 NRPS genes[21], although fungal genomes in other clades such as Microsporidia encode almost no predicted biosynthetic enzymes of natural products (see the MycoCosm portal[23] for the natural product clusters predicted in each fungal phylum). In some fungal phyla, the low number of predicted biosynthetic enzymes may be due to a lack of sequenced genomes. However, the average size of the fungal genome may also determine the biosynthetic potential of the phylum, since it has been demonstrated for bacteria that genomes less than 2 Mbp typically do not encode certain classes of biosynthetic enzymes[24]. Despite the wealth of valuable natural products already derived from fungi, the capacity for discovery is far from being reached. Certain fungal genera, such as *Aspergillus*, are disproportionately studied, although they represent only a

fraction of sequenced fungal genomes[25]. In contrast, anaerobic fungi are understudied microorganisms that are completely untapped with regards to their potential to synthesize natural products. Anaerobic fungi thrive in a competitive microbial environment despite being outnumbered in the rumen by bacteria[26]. They also have the ability to degrade biomass through unique enzymatic assemblies called cellulosomes[27]. Compared to aerobic fungi, the genomes of anaerobic fungi encode a greater diversity of enzymes, including hemicellulases[6].

Advancements in next-generation sequencing (NGS), such as PacBio Single Molecule Real Time sequencing platforms, facilitated the genomic sequencing of anaerobic fungi. These non-model organisms were historically difficult to sequence by conventional, short-read NGS techniques due to their highly repetitive and AT-rich genomes[28]. High-quality genomes as well as transcriptomes are now available for *Anaeromyces robustus*, *Caecomyces churrovis*, *Neocallimastix californiae*, and *Piromyces finnis*[6,7,29]. The high resolution genomes acquired for these organisms have revealed numerous biosynthetic genes for the synthesis of natural products (discussed further in Chapter two).

The rumen microbiome in general is a large source of secondary metabolites, since some anaerobic bacteria synthesize natural products[30]. However, since the genomes of anaerobic bacteria are small (roughly 2-3 Mb)[30], the number of biosynthetic genes encoded in each genome is generally less than fungi. Nevertheless, the great diversity of bacteria found within the rumen makes rumen bacteria an attractive source of novel natural products as well. Recent sequencing efforts have assembled thousands of metagenome-assembled genomes from 12 distinct phyla in the rumen[31].

1.3 Biosynthesis of natural products

Natural products are the results of secondary metabolism and thus are also referred to as secondary metabolites. Primary metabolism supplies the precursors necessary for secondary metabolism[32,33]. Secondary metabolite production is often synced with certain stages in morphological differentiation[33]. Historically, it was thought that secondary metabolites were only produced during stationary or secondary phase, when nutrients were limiting[33]. However, recent transcriptomic data[34] suggests that secondary metabolites are also produced during mid-log phase, even in the absence of induction by abiotic or biotic stimuli such as heat shock[35,36] or co-cultivation with another organism[37].

Even among eukaryotes, where clustering of genes is less common than prokaryotes, the pathways of secondary metabolism are co-localized on the genome[38]. In fungi, gene clustering of secondary metabolic pathways and even some primary metabolic pathways is more common that other eukaryotes[39]. The biosynthetic genes and other genes associated with the production of a particular secondary metabolite, such as genes encoding transporters, product tailoring enzymes, and regulatory proteins are clustered together (located adjacently) on the genome[21,39], as illustrated in Figure 1.1. The core or backbone gene of the biosynthetic gene cluster (BGC) can fall into many different categories, but the canonical classes are polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs). These enzymes are huge (on the order of 100 kDa), multi-modular complexes that encode all of the necessary domains for biosynthesis of complex molecules. These enzymes can be classified as modular, in which case each module acts only once, or iterative[40], in which case each module can act more than once[12].
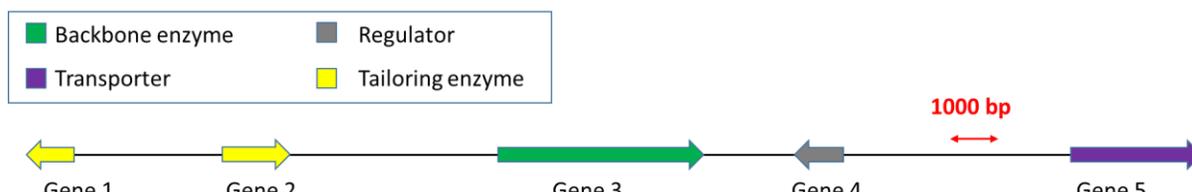
Figure 1.1: Depiction of a typical biosynthetic gene cluster consisting of genes co-localized on a chromosome. Arrows indicate genes and their direction of transcription and color denotes function.

1.3.1 Polyketide synthases and nonribosomal peptide synthetases are multi-modular enzymes that function like molecular assembly lines

Polyketides range extensively in both size and complexity, but they are all synthesized by polyketide synthases (PKSs) in a stepwise mechanism of chain extensions, similar to fatty acid formation[12]. The synthesis is commonly referred to as an "assembly-line"[41]. However, there are exceptions to this analogy, such as synthases in which modules or domains are skipped or act iteratively[42]. The minimal domains required for a complete elongation cycle in a PKS are ketosynthase (KS), acyl transferase (AT), and acyl carrier protein (ACP)[43], but other domains such ketoreductase (KR) and thioesterase (TE) can also be present (Figure 1.2). The KS domain is responsible for the key carbon-carbon bond formation via Claisen condensation between a malonyl thioester and an acyl thioester[44]. During polyketide biosynthesis, the ACP domain serves to tether a growing product chain[44]. The AT domain transfers acyl groups from CoA to the KS or ACP domains[44]. Fungal PKSs are classified as non-reducing, partially reducing, and highly reducing[44]. Examples are orsellinic acid, 6-methylsalicylic acid, and lovastatin, respectively[44]. Highly reducing PKSs bear the most similarity to fatty acid synthases[44].

In a manner similar to PKSs, NRPSs also function like molecular assembly lines. The building blocks that an NRPS uses to synthesize its product are amino acids, including nonproteinogenic amino acids such as D-amino acids[45]. As of 2007, more than 422

7

substrates were known for NRPSs[46], which suggests the wide chemical diversity of nonribosomal peptides. Essential domains catalyzing the elongation of peptidic intermediates include adenylation, peptidyl-carrier-protein (PCP), and condensation[47] (Figure 1.2). The adenylation domain is responsible for substrate recognition, the PCP domain is necessary for tethering the growing product chain, and the condensation domain catalyzes peptide bond formation[47].

(A)

(B)
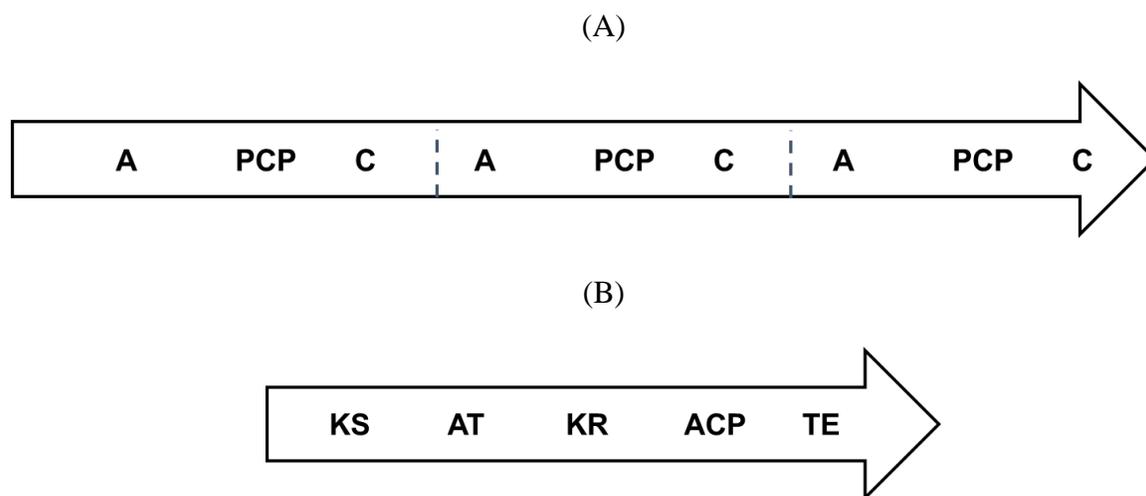
Figure 1.2: Depiction of the domain architecture of a modular NRPS (A) and an iterative PKS (B). A=adenylation, C=condensation, PCP=peptidyl-carrier-protein. AT=acyltransferase, KR=ketoreductase, ACP=acyl carrier protein, TE=thioesterase. Dashed lines indicate separate elongation modules.

Polyketides and nonribosomal peptides have furnished some of the most well-known antibiotics in history. The polyketide antibiotic erythromycin is synthesized by a well-studied modular Type I PKS and its biosynthetic mechanism became the paradigm for other Type I PKSs[42]. Penicillin, a nonribosomal peptide[48] synthesized by *Penicillium* spp., is arguably the most famous of all fungal secondary metabolites for both ushering in the golden age of antibiotics as well as for its impact on the outcome of World War II[21]. Although it was anticipated that chemical synthesis would replace cultivation as the means of production, fungi were able to produce remarkably high titers due to advancements in industrial

fermentation processes, which made fungal fermentation more advantageous than chemical synthesis[21]. Despite their renown as antibiotics, the bioactivities of polyketides and nonribosomal peptides are not limited to antimicrobial. For example, the polyketide lovastatin is used is a cholesterol-lowering agent[49].

1.3.2 *Ribosomally synthesized and post-translationally modified peptides (RiPPs) are chemically diverse natural products*

In contrast to the multi-modular enzymes that synthesize polyketides and nonribosomal peptides, ribosomally synthesized and post-translationally modified peptides (RiPPs) gain chemical complexity from post-translational modifications to a ribosomally synthesized precursor peptide, as their name suggests. Bacteriocins, or antimicrobial peptides typically produced by bacteria, are ribosomally produced peptides without post-translational modifications[50], but they are sometimes classified as RiPPs. Although their biosynthesis differs from polyketides and nonribosomal peptides, some commonalities exist between RiPPs and these classes of molecules. The biosynthetic genes of RiPPs are also clustered together on the genome. In addition, the chemical diversity of RiPPs may also confer useful properties, such as antifungal, antibacterial, or antiviral. Specific examples are discussed in the following paragraphs.

RiPPs are divided into several families according to their biosynthesis[51], but for all families the synthesis starts from a peptide precursor with an amino-N-terminal leader sequence necessary for recognition and subsequent post-translational modifications to the carboxy-terminal, core region of the peptide[51]. In the final steps of synthesis, the leader peptide is cleaved by peptidases and in some cases cyclization occurs[51]. RiPP families include lanthipeptides, sactipeptides, thiopeptides, lasso peptides, and microcins.

Lanthipeptides (sometimes referred to as lantipeptides) are an expanded family of RiPPs that originated from lantibiotics[52]. Lanthipeptides are characterized by the presence of the non-proteinogenic amino acids lanthionine and 3-methyllanthionine bonded by a thioether linkage[53]. Nisin is an antimicrobial lanthipeptide with a history of application in the food industry to prevent spoilage for more than 40 years[54]. The mechanisms of action for nisin are membrane permeabilization and disruption of cell wall biosynthesis[54]. Other important antibiotic lanthipeptides include actagardine, which is used to treat *Clostridium difficile* infection. The native function of lanthipeptides is not always antimicrobial. The lanthipeptides SapB and SapT are implicated in the formation of aerial mycelium by the filamentous bacteria *Streptomyces coelicolor*[55] and *Streptomyces tendae*[56], respectively.

Sactipeptides often demonstrate antimicrobial activity and thus are sometimes called "sactibiotics"[57]. Sactipeptides are readily identified by the presence of a specialized radical SAM enzyme within the gene cluster that is responsible for catalyzing a disulfide bridge between a cysteine and the α-carbon of another amino acid[53]. The mode of action for sactipeptides is not well-understood. The sactipeptide subtilosin A produced by *Bacillus subtilis* demonstrates antimicrobial activity against other Gram-positive bacteria[58]. The structure of sactipeptides is hairpin-like, with hydrophobic ends[58]. Solid-state NMR studies indicate that subtilosin A partially buries itself in lipid bilayers, disrupting the bilayer and causing permeabilization[59].

Thiopeptides are particularly complex macrocyclic natural products that contain multiple thioazole rings, a six-membered nitrogen-containing ring, as well as a side chain of dehydrated amino acid residues[53]. Thiopeptides have demonstrated antibacterial[60] and antifungal[61] activities. Some thiopeptides inhibit bacterial protein synthesis by binding to the ribosome and preventing conformational changes[60]. However, thiopeptides are not effective

against Gram-negative bacteria because they cannot transverse the outer bacterial membrane[58]. The native function of thiopeptides from streptomycetes is thought to be signaling[62].

Lasso peptides are named for their lasso-like structure, which is formed by an 8- or 9-membered macrolactam ring[53]. The structure of lasso peptides facilitates their activities as receptor antagonists and enzyme inhibitors[58]. The rigidity of their structure also provides resistance to peptidases and abiotic stressors such as heat and chemicals[53]. Several lasso peptides have medicinal applications: siamycin I as an anti-HIV agent[63], and lariatins A and B as anti-mycobacterial activity[64].

Microcins were historically considered to be ribosomally synthesized antimicrobial peptides produced by *Escherichia* spp.[58]. Some microcins are post-translationally modified, whereas others are not[58]. Microcins have potent, narrow-spectrum antimicrobial activity against Gram-negative bacteria, especially Enterobacteria[58]. Their mechanism of action involves commandeering the target cell's nutrient import machinery, such as siderophore receptors, as well as ABC transporters in some cases, to cross the double-membrane barrier[58]. Once inside the cell, microcins disrupt enzymes such as RNA polymerase or DNA gyrase[58]. The structure of some microcins, such as microcin J25, is lasso-like[65,66], and therefore they are often considered to be a subfamily of lasso peptides.

Letzel and colleagues characterized the RiPP portfolio of anaerobic bacteria[53] by mining publically available genomes. Although the BGCs were absent from more than half of anaerobic bacterial genomes, 18% encoded both PKS and NRPS gene clusters in addition to RiPPs and 10% encoded only RiPPs[53]. The Firmicutes phylum contained the largest number of strains encoding PKS/NRPS or RiPP gene clusters, which could in part be due to the large number of sequenced strains from Clostridia available for mining compared to other

anaerobic bacteria[53]. Bacterial strains native to animal environments comprised only ~10% of sactipeptide clusters. No strains containing lanthipeptides, thiopeptides, or lasso peptides originated from an animal environment. This work highlighted the capability of isolated anaerobic bacteria to synthesize RiPPs with potentially useful bioactivities, but also the need to sequence genomes from animal environments, such as the rumen, to gauge the biosynthetic potential of anaerobic bacteria from these environments. In 2018, Seshadri and colleagues mined the genomes from an extensive collection of 410 cultivated cow rumen microbes and revealed extensive representation of bacteriocins and lanthipeptides, as well as PKS and NRPS gene clusters, among these genomes (see Supplementary Figure 5 from their publication)[67].

Although the mechanisms of synthesis and examples of bioactive RiPPs discussed above have been limited to bacteria, fungi are also capable of producing RiPPs. However, characterization of fungal RiPPs is still at an early stage, since the biosynthesis of the first fungal RiPP, the phytotoxin ustiloxin B produced by *Ustilaginoideavirens*, was only elucidated in 2016. Nevertheless, many bioinformatics tools are available to predict both fungal and bacterial RiPPs[68], which facilitates the advancement of this field.

1.3.3   *Proposed mechanisms for the acquisition of biosynthetic gene clusters*

The clustering of genes involved in the synthesis of natural products, although convenient for bioinformatic detection, must have also an underlying biological significance. Several hypotheses have been proposed to explain biosynthetic gene clusters. One hypothesis is that the genes are clustered in order to facilitate co-regulation[21,69]. Others have attributed clustering to gene duplication and subfunctionalization or horizontal gene transfer (HGT) from bacterial operons[70,71].

HGT has been suggested to occur between a wide variety of eukaryotes and prokaryotes, including between fungi[72], as well as between fungi and bacteria[73]. Khaldi and colleagues suggested that horizontal transfer of a hybrid PKS-NRPS cluster occurred from *Magnaporthe grisea* to an ancestor of *Aspergillus clavatus*, arguing that HGT was the more parsimonious explanation for the presence of similar clusters in *Magnaporthe grisea*, *Aspergillus clavatus*, *Chaetomium globosum*, and *Stagonospora nodorum*[72] compared to multiple gene duplication and gene loss events. Evidence for HGT from bacteria to fungi includes the fungal isopenicillin synthase N gene cluster, where codon usage is more typical of bacteria than fungi[74]. Further evidence for HGT from bacteria to fungi includes the observation of genes with a scarcity of introns, such as is the case for certain PKS genes in lichen[73]. In some instances, the ketosynthase domain for a Type I fungal PKS will nest within bacterial domains[75] (example depicted in Figure 1.3), whereas in other cases the domains from bacterial PKS genes nest within fungal genes[24]. It has not been possible to ascertain the direction of HGT, but it is generally accepted that some amount of HGT has occurred between fungi and bacteria.
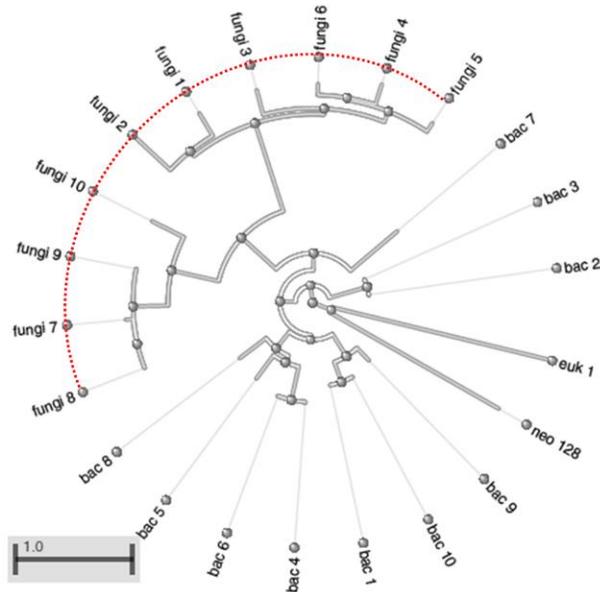
Figure 1.3: Phylogenetic tree showing fungal genes nested within bacterial genes provides evidence for horizontal gene transfer. Fungal genes (fungi 1-8) are marked by a dotted red half-circle) nested. Bac=bacterial gene, neo=Neocallimastigomycete fungus, euk=other eukaryote.

1.3.4 Regulation of secondary metabolism

The regulation of secondary metabolism is well-established in some organisms, such as filamentous fungi from the phylum Ascomycota[76]. Genes within BGCs are co-regulated, meaning that they are either simultaneously expressed or repressed[76]. Many BGCs contain a gene encoding a transcription factor that is necessary for the expression of the rest of the cluster[76]. Common transcription factor types in BGCs include zinc binuclear proteins as well as Cys2His2 zinc finger proteins[76]. However, in some cases other transcription factors are possible, such as bANK transcription factors, which are named for a bZIP basic region and ankyrin repeats[77–79].

In addition to pathway-specific regulators, some BGCs are known to be globally regulated in response to environmental stimuli. Global regulators of secondary metabolism upregulate some clusters and downregulate others. Examples include PacC, which

modulates BGC expression in responds to pH fluctuations[80,81], CreA that responds to different carbon sources[82], and AreA that responds to nitrogen sources[83].

Similarly, the regulation of secondary metabolism is also linked to fungal development. For example, mycotoxins, which are natural products from filamentous fungi with toxic effects on vertebrates, are produced at the onset of sporulation[84]. In some cases, secondary metabolites are pigments that protect spores[84]. Deletion or disruption of key regulatory genes results in simultaneous abatement of secondary metabolite production and fungal development. Examples include conidiation and melanin expression[85], as well as asexual sporulation and mycotoxin production[86].

Global regulators of secondary metabolism that act through chromatin-remodeling have been identified in ascomycetes. One such regulator is LaeA, first discovered in *Aspergillus nidulans*, but later identified in in other *Aspergillus* spp.[87] as well as other fungi, such as the industrially-relevant *Trichoderma reesei*[88]. Deletion of LaeA reduced secondary metabolite production through transcriptional repression, whereas overexpression increased production of the secondary metabolites penicillin and lovastatin[87]. LaeA encodes a putative protein methyltransferase with a SAM domain that is conserved with histone and arginine methyltransferases[87]. DNA microarray analysis revealed that LaeA-deficient mutants of *Aspergillus fumigatus* differentially expressed ~1000 genes, of which 10% were part of BGCs[89]. Thirteen out of 22 gene clusters identified in *A. fumigatus* were regulated by LaeA, and the majority of these were downregulated. This work demonstrated the utility of sequencing experiments in deciphering the regulation of secondary metabolism

A classic example of the coordination of secondary metabolism with fungal development is the VelB/VelA/LaeA complex[90]. *Aspergillus nidulans* can reproduce both sexually, via fruiting bodies, or asexually, via conidia[91]. In wild-type *A. nidulans*, red-light induced

conidiation and the protein VeA (named for its velvet domain) is localized to the cytoplasm[90]. In darkness, VeA is localized to the nucleus[90]. However, in *veA*-deletion strains, sexual fruiting bodies could not be formed and conversely, overexpression of *veA* caused constant fruiting body formation without regard to light[90]. Furthermore, *veA* deletion strains were unable to produce the secondary metabolites sterigmatocystin or penicillin[92]. Together, these studies revealed that VeA forms a complex localized in the nucleus with LaeA and another velvet domain-containing protein, VeB, resulting in a light-responsive connection between secondary metabolite production and the formation of sexual fruiting bodies[90]. The link between fungal development and secondary metabolism through the velvet regulatory protein was further established in numerous other filamentous fungi[91].

Several studies have pointed to chromatin remodeling as a mechanism of regulation for secondary metabolism. Since LaeA is a putative methyltransferase with homology to histone methyltransferases, it was hypothesized that LaeA controls chromatin accessibility[93]. Many of the gene clusters regulated by LaeA are located at subtelomeric regions[89], which is consistent with the epigenetic regulation of subtelomeric regions in other eukaryotic genomes[94–96]. The sterigmatocystin gene cluster in *A. niculans* has been shown to be inactivated by histone 3 lysine 9 trimethylation (H3K9me3) marks and the binding of heterochromatin protein-1[97]. LaeA was demonstrated to counteract heterochromatin formation at the promoter of the pathway-specific activator AflR[97]. Similarly, heterochromatin protein-1 and H3K9me3 marks were both involved in the regulation of multiple BGCs in *Fusarium graminearum* [98] Knowledge of epigenetic regulation of fungal secondary metabolism has informed strategies to enhance the production of known secondary metabolites and to aid in the discovery of novel secondary metabolites[99].

1.4 Strategies to discover novel natural products

1.4.1 'Omics' techniques revolutionize the field of natural products

Genomics, transcriptomics proteomics, and metabolomics have the potential to further speed the process of discovering novel natural products with therapeutic possibility. Genomics allows comparative analysis of BGCs between strains. For example, Ziemiert and colleagues performed comparative analyses of the BGCs of 75 strains of *Salinispora*, a marine actinomycete, and characterized the extent of their biosynthetic diversity[100]. Transcriptomics comprehensively quantifies the transcription level of all BGCs and thus allows for the determination of active clusters as well as comparisons between active and silent BGCs in different strains[34]. Proteomics provides additional validation of which BGCs are active, since the presence of the biosynthetic enzymes themselves can be detected by mass spectrometry.

Metabolomics and tools such as molecular networking enable quantification of metabolite abundance via mass spectrometry as well as the partitioning structurally related features into clusters. In addition, molecular networks can be used to visually compare experimental conditions in order to readily detect activating stimuli. These techniques have been proposed[101] and used[102] to discover novel natural products when combined with classic bioactivity-guided approaches.

Although methodologies for integrated interpretation of 'omics' datasets are still in development[103,104], even when used separately these techniques provide unprecedented insight into secondary metabolism, as depicted in Figure 1.4.

DNA
Genomics
Biosynthetic
Gene Clusters

mRNA
Transcriptomics
Active
Biosynthetic
Genes

Proteins
Proteomics
Biosynthetic
Enzymes

Natural Products
Metabolomics
Metabolic
Products
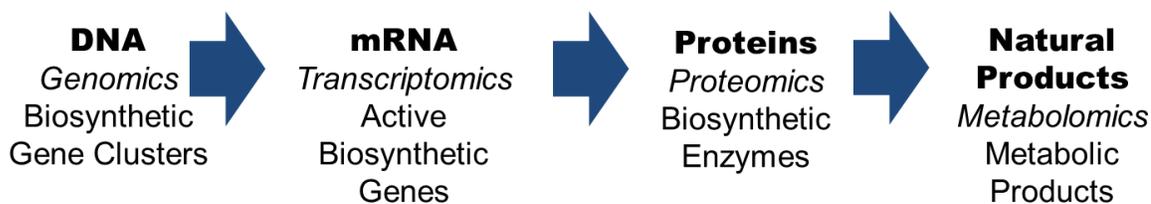
Figure 1.4: 'Omics' techniques allow the characterization of secondary metabolism at all stages, from biosynthetic gene clusters to metabolites.

1.4.2   Genome mining and strategies to awaken silent biosynthetic gene clusters

Next-generation sequencing has permitted unprecedented insight into the biosynthetic potential of organisms. Surprisingly, many organisms that are known to produce only a few secondary metabolites encode numerous uncharacterized gene clusters[4]. Bioinformatic software such as antiSMASH[11,105–108] (antibiotics and Secondary Metabolites Analysis Shell) and SMURF[109] (Secondary Metabolite Unknown Regions Finder) were developed to accurately predict the BGCs of various types of natural products, including PKSs and NRPSs. In parallel, strategies were developed to activate clusters not expressed during standard microbial cultivation. These strategies can be divided into two categories: abiotic and biotic. Examples of successful abiotic strategies include the application of ethanol or heat shock to elicit production of the polyketide antibiotic jadomycin B from *Streptomyces venezuelae* [35,36]. Co-cultivation with other microorganisms has proven to be an especially successful biotic strategy[37,110,111].

1.4.3   Co-cultivation to induce expression of silent biosynthetic gene clusters

Co-cultivation of bacteria with other bacteria, fungi with other fungi, and bacteria with fungi have all elicited the production of previously undetectable secondary metabolites[112]. In the majority of these cases, the mechanism of induction is not known. There is at least one instance in which the introduction of bacterial cell lysates was sufficient to induce expression of a BGC[113]. However, in most cases co-cultivation is the method of choice.

18

Potential mechanisms of elicitation that have been suggested for actinomycetes in particular include: (1) physical (cell to cell), (2) small molecule mediated, (3) catalytic activation of precursor molecules, and (4) horizontal gene transfer[114].

Studies of bacterial induction of fungal natural products have focused predominantly on fungi belonging to the Dikarya subkingdom[112]. Successful examples include the induction of a novel antimicrobial polyketide from *Aspergillus nidulans* by *Escherichia coli*[37], as well as the induction of a novel cytotoxic alkaloid from *Aspergillus fumigatus* by *Streptomyces peucetius*[115]. A mechanism of induction was not proposed in the latter case. However, in the case of *A. nidulans* and *E. coli*, dialysis experiments (in which the bacteria and fungus were separated by dialysis tubing with a membrane permeable to small molecules) suggested that a physical mechanism of induction was necessary.

Although rumen fungi have frequently been co-cultivated with methanogenic archaea[116–120] and rumen bacteria[121–125], these studies have all focused on biomass degradation. Dehority and Tirabasso demonstrated that fungal cellulose degradation *in vitro* in a secondary fermentation was inhibited by an initial bacterial fermentation period and subsequent autoclaving[124]. The degree of inhibition correlated with the length of the primary bacterial fermentation. They concluded that a heat-stable inhibitory factor was produced by the bacteria. They found that the factor was resistant to proteolytic enzymes. This work was the first evidence presented for antibiosis between rumen bacteria and fungi.

1.5   *The role of natural products in natural and synthetic microbial consortia*

1.5.1   *Fungal and bacterial secondary metabolites shape natural microbial communities*

Fungal secondary metabolites confer a selective advantage to the producer in their native environment: they can be used for defend or competition. For instance, endosymbiont fungi

are known to produce secondary metabolites that defend against arthropods[126], and saprotrophic fungi produce secondary metabolites in order to compete with prokaryotes in decomposer systems[20]. These examples illustrate how fungal secondary metabolites can influence the microbial communities where their producers reside.

The soil environment is well-studied in terms of the effect of secondary metabolites on community structure and dynamics. Experiments have demonstrated that the production of antimicrobial secondary metabolites by certain soil bacteria facilitates their use as biocontrol agents to prevent crop infection by pathogenic bacteria or fungi[8]. Volatile Organic Compounds (VOCs), of which many are secondary metabolites, have been proposed as a mechanism for communication between soil microbes, which are dispersed over a large surface area in the soil[8]. The observation that the VOC profile of soil microorganisms depends on environmental parameters such as nutrient availability and temperature is consistent with this hypothesis[8]. Furthermore, a sequence of experiments in which soil bacterial isolates were pairwise coupled through an atmospheric connection with one of four different fungi demonstrated discrete, species-specific pairwise interactions that were either inhibitory or stimulatory[127]. Together, these findings suggested that secondary metabolites in the soil serve to transmit environmental stimuli through the microbial community as well as to inhibit or stimulate the growth of microorganisms. In another study, Kellner and Zak also hypothesized that fungal secondary metabolites mediate ecological interactions in the forest soil between microbes based on the expression of Type I PKS genes by ascomycete fungi measured from environmental soil samples[128]. Using advanced mass spectrometry techniques coupled to molecular networking, more recent work has highlighted that species-level interactions *in vitro* can result in distinct chemical profiles that include secondary metabolites[110].

These studies point to the role of fungal and bacterial secondary metabolites as antimicrobials or signaling compounds, and hence as mediators of microbial community structure and dynamics. However, it should be noted that *in vitro* experiments may not accurately reflect either microbial or secondary metabolite concentrations representative of the natural environments. Nevertheless, combining studies of pairwise interactions with environmental measurements is a powerful approach to formulate and test hypotheses regarding the function of microbial secondary metabolites.

1.5.2 *Leveraging natural products in microbiome engineering*

Since natural products often serve important ecological roles in the interactions between microorganisms, they should be considered and leveraged in microbiome engineering efforts. Natural products may confer resistance to foreign microorganisms that would otherwise disrupt the stability and function of the microbiome. They may also control the population dynamics between different microbial community members, either directly, such as by narrow-spectrum antimicrobial activity against another specific microorganism, or indirectly, such as by conferring tolerance to environmental stressors or the ability to uptake scarce nutrient (e.g. siderophores that scavenge iron[129]). Natural products may serve to balance membership at the kingdom-level, through the production of antifungal and antibacterial compounds.

With all of these potential implications in community structure and dynamics, it is evident that the design of a stable microbiome to achieve specific aims, should incorporate the functions of natural products. The application of native natural products to microbiomes is a potential engineering strategy. Methods to engineer microbiomes *in situ* can be broadly evaluated based on their ability to control perturbation magnitude and the specificity.[130]. The

21

application of natural products to microbiomes affords both control of the perturbation magnitude as well as specificity, since natural product bioactivities can be broad- or narrow-spectrum. Although some natural products are already applied to microbiomes as prebiotics and antibiotics[130,131], there is room for improvement in the rational application of native natural products to manipulate community structure.

Even more desirable than the exogenous application of natural products, which would require total chemical synthesis or isolation from the native source, would be to control the expression of specific BGCs within target microbial populations. Combined with gene editing techniques to selectively activate the expression of BGCs, natural products could be used to perturb specific population within the microbiome by the production of narrow-spectrum antimicrobials. Knowledge of naturally-occurring secondary metabolites synthesized by community members is a necessary step towards achieving this goal.

1.6   Organization of the Dissertation

This thesis is comprised of seven chapters. Chapter two broadly describes the secondary metabolism of anaerobic fungi, establishing foundational 'omics' datasets that characterize and compare the biosynthetic enzymes of natural products. Chapters three and four use co-cultivation and transcriptomics to characterize the interactions of anaerobic fungi with other microbes: methanogenic archaea (chapter three) or rumen bacteria (chapter four). Chapter five explores how anaerobic fungi rely on conserved pathways (the unfolded protein response and heat shock response) rather than specialized metabolism in order to cope with other forms of environmental stress. Chapter six extends what we learned in chapter four about the role of fungal secondary metabolites more broadly, from the pairwise interaction of a fungus with a single bacterial strain, to the interactions of complex microbial communities

of multiple fungi and bacteria enriched from the goat fecal microbiome. Finally, chapter seven offers concluding remarks, future directions, and perspectives as to how secondary metabolites derived from the rumen microbiome can be leveraged for biotechnology.

2 Anaerobic fungi are an untapped reservoir of natural products

2.1 Introduction

Secondary metabolites, or natural products, have inspired many medicinal drugs, including antibiotics, antitumor agents, and immunosuppressants[132]. In addition to pharmaceuticals, natural products have also found use as valuable bio-based products such as drop-in biofuels and renewable polymers[22,133]. Across microbial diversity, fungi are especially prolific secondary metabolite producers, with a single strain such as *Aspergillus nidulans* FGSC A4 producing 15 compounds characterized in the Minimum Information about Biosynthetic Gene cluster (MIBiG) database[134] or *Aspergillus fumigatus* with over 18 characterized metabolites[135]. Despite the wealth of valuable polyketides and other secondary metabolites already derived from fungi, the capacity for discovery is far from realization. Certain fungal genera, such as *Aspergillus*, are disproportionately studied, although they represent only a fraction of sequenced fungal genomes[25]. Rediscovery of natural products like antibiotics has proved problematic, requiring innovative approaches to silence known antibiotic-producing genes[136], or alternatively, investigation of rarely explored microbiomes, such as the rumen, for sequence-divergent biosynthetic genes. Examples of antibiotics discovered from unusual environments include lugdunin, which was discovered from a commensal bacteria of the human microbiome [137], as well as teixobactin, which was discovered from a screen of previously uncultured bacteria[138]. Both lugdunin and teixobactin were active against *Staphylococcus aureus,* and teixobactin was active without detectable resistance.

Anaerobic gut fungi (class Neocallimastigomycetes) are understudied organisms that thrive as members of a consortium of archaea, bacteria, and protozoa in the digestive tracts of large herbivores[5,26,139]. In these habitats, fungi are vastly outnumbered by prokaryotic microorganisms by several orders of magnitude[5,26,139]. For instance, rumen bacteria are estimated at 10^{10} cells per gram rumen contents[5,26,139] whereas fungi are estimated at 10^6 per gram[5,26,139]. These fungi are of recent biotechnological interest due to their array of biomass-degrading enzymes, but through genome and transcriptome sequencing it has become evident that they also have a range of biosynthetic enzymes for natural products[6,7,140]. We hypothesize that anaerobic gut fungi synthesize natural products to compete with other microbes for survival in their native environment. Natural products are known to serve a variety of functions to their producers in other environments, including oxidative stress tolerance[141], fungal development[142], and antibiosis[143]. Similarly, the natural products of anaerobic gut fungi may serve directly (by antibiosis) or indirectly (by conferring environmental stress tolerance) to allow the fungi to persist despite being outnumbered by other members of the rumen community.

Here, we take an integrated approach combining genomics, transcriptomics, proteomics, and metabolomics to develop a pipeline to identify and characterize natural products from anaerobic gut fungi. By using antiSMASH (antibiotics and Secondary Metabolites Analysis Shell)[106], we classify the types of biosynthetic enzymes present in the fungal genomes of representative Neocallimastigomycetes and quantify the homology between strains as well as to other organisms. Transcriptomics and proteomics are used to validate and refine these predictions. Finally, we demonstrate by metabolomics and molecular networking that anaerobic gut fungi produce the polyketide-related compound baumin, as well as at least three groups of potentially novel natural products.

2.2   Results and Discussion

2.2.1   The genomes of anaerobic gut fungi encode diverse biosynthetic enzymes for natural products and antimicrobial peptides

Previously, we isolated and sequenced the genomes of four species anaerobic gut fungi from the early-branching fungal class Neocallimastigomycotes[7] (Brown et al., manuscript in preparation). Using antiSMASH version 3.0[106] to mine the genomes of *Anaeromyces robustus*, *Caecomyces churrovis*, *Neocallimastix californiae*, and *Piromyces finnis*, we uncovered 146 genes encoding enzymes responsible for the synthesis of various classes of secondary metabolites (Figure 2.1). These enzymes include canonical classes such as polyketide synthases (PKSs) and non-ribosomal peptide synthetases (NRPSs), as well as putative classes based on the ClusterFinder[144] algorithm. The number of predicted biosynthetic gene clusters in all four strains as a proportion of total genes is commensurate with the prolific secondary metabolite-producing Aspergilli, which contain roughly 50-70 backbone genes per strain, as quantified by the SMURF[109]-derived algorithm of the Secondary Metabolite Clusters feature in MycoCosm[23]. By the same metric, fungi of class Neocallimastigomycetes surpass other members of Chytridiomycota by an order of magnitude.
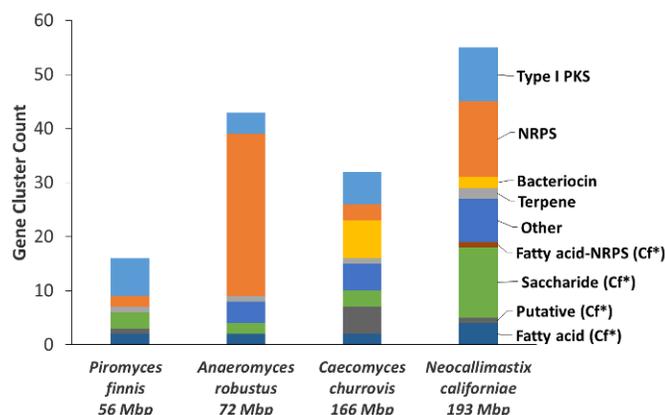
Figure 2.1: Anaerobic fungal genomes reveal putative natural products of many different types. The genomes of anaerobic fungi[6] (Brown et al., manuscript in preparation) were mined for biosynthetic gene clusters and cluster types by antiSMASH 3.0[106] with the ClusterFinder option. *Cf=gene clusters identified by ClusterFinder. The ClusterFinder[144] algorithm extends the secondary metabolite search to include biosynthetic gene clusters of unknown classes based on the occurrence of common protein family domains inside and outside of the cluster. PKS=polyketide synthase, NRPS=nonribosomal peptide synthetase.

Surprisingly, antiSMASH identified nine bacteriocins, or antimicrobial peptides (AMPs) typically produced by bacteria[145], in the genomes of anaerobic fungi: two predicted peptides were found in *N. californiae* and four unique peptide sequences were found in *C. churrovis*. We validated the antiSMASH predictions against other tools specifically designed to identify AMPs from sequence data. Namely, we queried the six unique bacteriocn amino acid sequences using APD3: Antimicrobial Peptide Calculator and Predictor[146] and CAMPSign[147]. Although none of the sequences belonged to the 45 AMP families in CAMPSign, subsequent BLAST to the AMP Databases indicated that all sequences shared at least 31.6% identity to putative bacteriocins or lactococcin 972[148]. The bacteriocins located on *C churrovis* scaffolds 90 and 616 and *N. californiae* scaffold 388 were transcribed, but not detected in the proteome. Taken together, these results indicate that both *C. churrovis* and *N. californiae* genomes encode potential AMPs in addition to an arsenal of PKSs and NRPSs.

27

To further affirm the biosynthetic genes of anaerobic gut fungi, we compared the antiSMASH predictions to the secondary metabolite genes predicted by the SMURF (Secondary Metabolite Unknown Regions Finder) algorithm[109] in the MycoCosm portal[23]. For all strains except *A. robustus*, antiSMASH, with the ClusterFinder algorithm enabled, predicted more biosynthetic genes because it detected a wider array of natural product classes than SMURF, including bacteriocins and putative classes such as fatty acid and saccharide derivatives. For *A. robustus*, SMURF predicted an additional five PKS-like biosynthetic genes. Despite these differences, the majority of the regions on each scaffold predicted by antiSMASH or SMURF to harbor biosynthetic genes were the same. Ninety percent of the backbone genes predicted by SMURF in each fungal strain were located on the same or overlapping scaffold region where antiSMASH also identified biosynthetic genes.

2.2.2 *The biosynthetic gene clusters of anaerobic gut fungi are isolated or cluster with non-conventional genes*

The biosynthetic enzymes of fungal secondary metabolism are typically, but not always, encoded by genes locally clustered on the chromosome with other genes in the biological pathway, such as genes that encode tailoring enzymes, transporters, self-resistance genes, and transcription factors[39,149]. AntiSMASH predicted cluster accessory genes based on GlimmerHMM[150] and up to 20 kbp intergenic distance for the outermost gene[11]. Based on RNA-seq data, antiSMASH was a poor predictor of the accessory genes. In order to delineate the accessory genes of each cluster, we relied on a variety of gene prediction models, including GeneMark[151,152] and fgenesh[153], and only included genes that were validated via RNA sequencing (see Methods). The curated gene clusters are presented in the

28

Secondary Metabolite Clusters feature of the MycoCosm portal[23]. Approximately 60% of the backbone genes with RNA-seq support are located in clusters of two or more genes, and 40% of the backbone genes are isolated (neighbored by genes greater than 10 kbp apart or by genes with poor RNA coverage). For the backbone genes of anaerobic fungi that are located in clusters, some of the neighboring genes are not typically found in either bacterial or other fungal biosynthetic gene clusters. Many of the neighboring genes encode hypothetical proteins or lack any homology-based annotations. However, in some cases the neighboring genes include solute transporters and enzymes responsible for post-translational modifications (e.g. phosphorylation and palmitoylation), which are more commonly observed in biosynthetic gene clusters. Notably, only the PKS-like gene cluster of *P. finnis* located on scaffold 39 (core gene MycoCosm Protein Id 358210) includes a putative transcription factor (414496). However, the *A. robustus* PKS located on scaffold 258 and *N. californiae* PKS-like gene cluster on scaffold 59 both include proteins with ankyrin repeats (*A. robustus* 270780 and *N. californiae* 668532) that may be bANK family transcription factors found in several other fungi[77,79,154]. The predicted gut fungal proteins do not match the motif of basic amino acids found in other bANK proteins[77–79], but this motif is not required for transcription factor activity[78]. Non-conventional neighboring genes that are present in more than one gene cluster include C-type lectins (*N. californiae* Protein Ids 502167 and 674020, and *P. finnis* 349079), peptidases (*C. churrovis* 519541, *P. finnis* 241287), and calmodulin-related proteins of the EF-hand superfamily (*A. robustus* 27040 and *C. churrovis* 200925). Although the functions of these genes are unknown, it is possible that they may be self-resistance genes. Self-resistance genes have been observed in both bacterial and fungal biosynthetic gene clusters[155,156]. Another candidate self-resistance gene is *C churrovis* 17006, encoding ribosomal protein L19e (specific to eukaryotes and archaea),

which suggests that the backbone enzyme, encoded by Protein Id 17094, may synthesize a compound with activity against another eukaryote. It has been suggested, but not proven, that variant copies of the ribosomal L11 protein may be self-resistance genes for *Bacillus cereus* ATCC 14579, which is a producer of thiocillin[157]. The function of these non-conventional neighboring genes in the gene clusters of anaerobic fungi and whether they have a role in gut fungal secondary metabolism remains to be determined.

2.2.3   *Biosynthetic gene sequences support horizontal gene transfer from other rumen microbes as a mechanism of acquisition*

We compared the similarity of the genes encoding core biosynthetic enzymes to other organisms to deduce the novelty and phylogenetic origin of the genes. The top-scoring BLAST+[158] hits (Figure 2.2) for 20% of the total core biosynthetic genes were other fungal genes (protein sequence identity>30%, coverage>25%, and E-value<$1x10^{-8}$). The majority of the homologous genes were hypothetical or uncharacterized proteins from other early-diverging fungi like chytrids, with a few instances of genes from higher-order fungi. 80% of the homologous genes from higher-order fungi were from basidiomycetes, possibly due to the ancestral intake of basidiomycete fungi with forage by the herbivore hosts and subsequent horizontal gene transfer (HGT). However, the top hits for 63% of the gut fungal core biosynthetic genes appeared to be of bacterial origin rather than fungal.
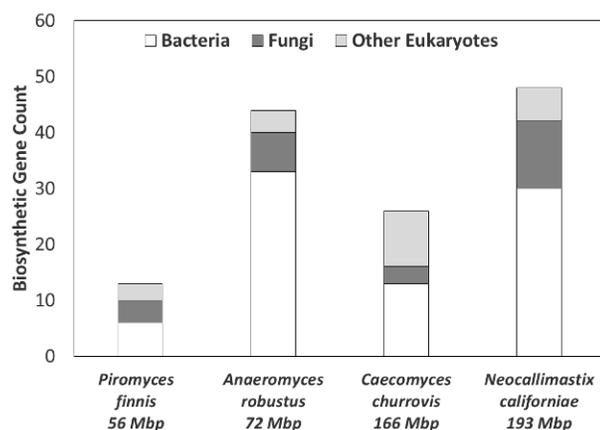
Figure 2.2: Biosynthetic genes from anaerobic fungi show the greatest similarity t bacteria. Core biosynthetic genes with at least three domains identified by antiSMASH were queried against NCBI's non-redundant protein database using BLAST+. Top hits (largest bitscore) with E-value less than 1×10^{-8}, greater than 30% identity, and greater than 25% coverage were classified for each biosynthetic gene according to taxonomy.

Due to the high level of HGT with bacteria characterized from the carbohydrate active enzymes of Neocallimastigomycetes[7], we probed whether any core biosynthetic genes may have also arisen in these fungi via HGT. Phylogenetic trees of the PKS ketosynthase domains and NRPS condensations domains were constructed. These domains append additional subunits to the growing product chain, and they have been shown to be a good proxy for the entire biosynthetic gene when constructing phylogenies[75,100,159,160]. HGT with bacteria was not supported for the PKS genes, since no ketosynthase domains nested within bacterial sequences and only 10% of domains were sister to bacteria. However, 44% of the fungal NRPS condensation domains were sister to or nested within domains from bacterial genes. Many of these bacteria are native to the rumen, such as *Clostridium cellulovorans*, thus supporting the hypothesis that some of the biosynthetic genes were likely horizontally transferred from bacteria. The majority of the transcribed NRPS genes were not part of a gene cluster. Therefore, at least some of the fungal NRPS genes may have been acquired by HGT of a single bacterial gene or transfer of an operon and subsequent loss of neighboring

31

genes. Similarly, phylogenies were constructed for the bacteriocins (Supplementary Figures 8.1.1-8.1.6). The putative bacteriocins were sister to Ciliophora, Firmicutes, or Actinobacteria (Supplementary Figures 8.1.1, 8.1.4-8.1.6), with the exception of *C. churrovis* bacteriocins on scaffolds 90 and 616, which were most closely related to eukaryotes from Rhizaria (Supplementary Figures 8.1.2 and 8.1.3). Therefore, the gut fungal bacteriocins may have been acquired from protozoa or bacteria in the rumen. There are many instances where HGT has been supported as the mechanism of acquisition of biosynthetic gene clusters (BGCs) between bacteria and fungi[24,71,73–75] as well as between fungi[72,161], though such BGC HGT mechanisms have not been previously characterized in the rumen. However, the case for HGT is not as definitive for the bacteriocins as for the NRPSs, since in order to identify homologs it was necessary to relax the E-value threshold to 0.1 and expand the search databases to include MMETSP[162] (see Methods). Nevertheless, it is clear that some of the genetic potential for natural products present in gut fungal genomes may be due to the complex microbial community in which they evolved.

Since 20% of the total biosynthetic genes were similar in sequence to other fungi, we investigated whether similarities existed in the regulation of secondary metabolism. Velvet regulatory proteins, which are characterized by a velvet domain approximately 150 amino acids long, are known to coordinate development with secondary metabolism in other fungi, typically in complex with the methyltransferase LaeA or and other velvet proteins[90,91]. Homologs of the developmental regulator *vosA* gene of *A. nidulans* [163,164], which contains a velvet domain at the N-terminus of the protein, were present in the *C. churrovis (*MycoCosm protein Ids 623244 and 624976), *N. californiae* (112212), and *P. finnis* (179530) genomes. These proteins have a primary region of homology centered at the velvet domain, with some conserved amino acids distal to the velvet domain. Genes containing the velvet domain have

been found in the genomes of other species of Chytridiomycota, such as the frog pathogen *Batrachochytrium dendrobatidis*[91]. However, the anaerobic gut fungi are unique among the chytrids in that their genomes contain not only velvet homologs, but also the biosynthetic machinery for secondary metabolism. Thus, the development and secondary metabolism of anaerobic gut fungi may be regulated by a velvet domain-containing protein acting in concert with other proteins. At present, it is not known whether the velvet proteins form a complex with a LaeA-like methyltransferase, similar to filamentous fungi.

2.2.4   *Polyketide synthases are conserved between genera of anaerobic gut fungi*

Although some of the core biosynthetic genes of anaerobic gut fungi are homologous to bacteria or higher fungi, the majority of PKSs are unique to the anaerobic gut fungi. On average, the PKS genes only share 34% amino acid identity to their top-scoring homolog, excluding Neocallimastigomycetes, and the highest similarity was only 39% (*C. churrovis* PKS on scaffold 118). We hypothesize that PKS genes present in multiple strains of gut fungi have important biological functions that confer fitness to anaerobic gut fungi, either by promoting their unique life cycle or distancing microbial competitors. A total of 23 iterative type I PKS genes of four or more enzymatic domains were identified by antiSMASH across all four fungal strains. These 23 PKS genes group into six PKS families by OrthoFinder[165]. All of the families have genes in three or more strains, and PKS families 1, 2, and 4 are represented across all four strains. The corresponding gene clusters of the PKSs contain orthologous neighboring genes (Figure 2.3), which suggests that the polyketides in each family may serve a common function. A phylogenetic tree of the PKS genes of *A. robustus*, *C. churrovis*, *N. californiae*, and *P. finnis*, shown in Supplementary Figure 8.1.7, affirms the

close evolutionary relationships between the PKS genes of different fungal genera. All of the

PKS genes are transcribed during the *in vitro* cultivation described previously[6,29].

Figure 2.3: Many PKS families are conserved across Neocallimastigomycetes. 23 total PKS genes predicted by antiSMASH across four fungal strains (vertical axis) can be represented in six PKS families (horizontal axis) as grouped by OrthoFinder[165]. Only one member per species is depicted in the figure. Neighboring orthologous genes in each cluster are defined as bidirectional top-scoring BLASTp hits from filtered model proteins between genomes with E-value threshold of 10[-5] and are indicated by matching colors in each PKS family. No color signifies the gene lacks a corresponding ortholog in the cluster. Triangle: PKS; Circle: post-translational modification enzyme; asterisk: multiple genes depicted as a single gene; diamond: transporter; square=all other genes. For full gene annotations, see additional dataset S3.

To probe the conservation of neighboring genes in PKS gene clusters, we compared the

PKS gene clusters from PKS family 1 (Figure 2.4). Six genes were conserved between *C*

churrovis and *A. robustus* and three genes were conserved between *N. californiae* and *P.*

finnis. Across all four strains, the PKS gene and a gene of unknown function containing a

WD 40 repeat were conserved. The PKS gene was present in two copies in the *N. californiae*

genome: one copy on scaffold 26 and the second on scaffold 182 (not shown in Figure 2.4).

The gene cluster on scaffold 182 included Protein Id 705610, a Rap1-GTPase-activating

protein, which was an ortholog of *A. robustus* protein Id 283391. All constituent genes in

these gene clusters were transcriptionally active. It is possible that these gene clusters

constitute a polyketide biosynthesis pathway, but these gene clusters could also be an artifact

of regions of genomic synteny between fungal strains. Regardless of whether the PKS gene

products independently synthesize a polyketide or are part of a more complex biosynthetic

pathway, the conservation of the PKS gene suggests that the polyketide may be biologically

important. One possibility is that the polyketide regulates the complex life cycle of anaerobic gut fungi. In the life cycle of anaerobic gut fungi, motile zoospores encyst into plant biomass and grow into a vegetative state, which develops reproductive sacs called sporangia that bear many zoospores[5]. Secondary metabolites are known to regulate morphology and differentiation in other fungi, especially sporulation in ascomycetes[142].
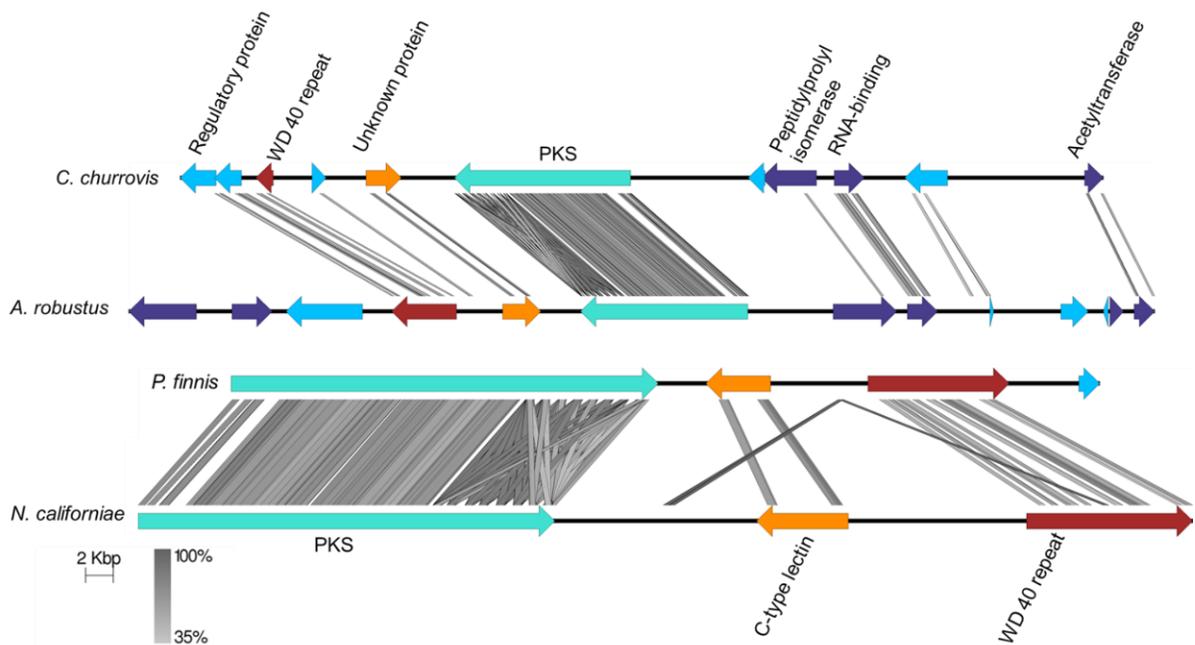
Figure 2.4: A PKS gene cluster is conserved among four strains of anaerobic gut fungi. Regions 50 bp and larger of at least 35% identity are highlighted in gray between genes. The turquoise PKS gene and red gene of unknown function containing a WD 40 repeat are shared among all four strains. Figure was generated using Easyfig[166].

2.2.5 *Transcriptomics, proteomics, and N6-methylation indicate that many of the biosynthetic genes of anaerobic gut fungi are active during standard laboratory cultivation*

Following the establishment of the presence of biosynthetic genes in gut fungal genomes, we probed what proportion of these genes were expressed. We demonstrate here through a combination of transcriptomics, epigenetics, and proteomics that anaerobic gut fungi

transcribe and translate a substantial portion of their biosynthetic genes. Out of 131 total

biosynthetic genes of three or more catalytic domains (e.g. adenylation) across all four

fungal strains, 34 are actively transcribed at mid-log phase during the standard laboratory

growth conditions described previously[6,29], whereas the remainder are silent (Figure 2.5).

The proportion of transcribed genes varied between 22 and 31% across all four strains of

anaerobic gut fungi. Using five different media formulations with varied nutrient complexity

and availability (Table S4), 13 of the backbone genes of *N. californiae* were differentially

regulated (Figure 2.6). The presence of mRNA and its regulation are promising indicator that

some secondary metabolite genes are active even when the anaerobic gut fungi are cultivated

outside of their native environment. These results also support the recent finding by Amos

and colleagues[34] that many biosynthetic genes are actively transcribed during mid-log phase,

not only during stationary phase. However, it is possible that more genes are expressed

during late stationary phase, but this was not tested by transcriptomics due to prevalence of

highly degraded mRNA from cultures harvested at that phase.
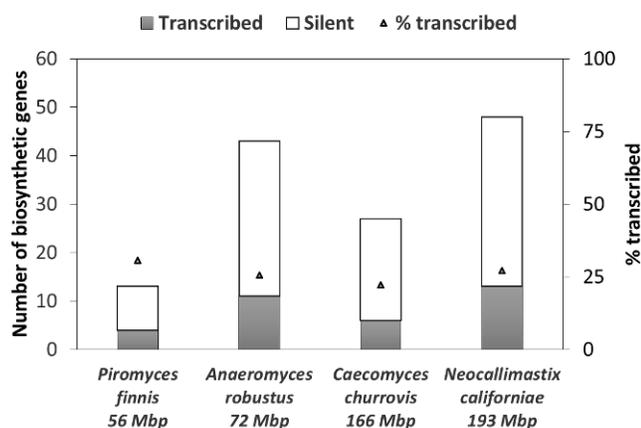
Figure 2.5: Many core biosynthetic genes of anaerobic gut fungi are transcribed during standard laboratory cultivation. Transcriptomes were previously acquired from anaerobic fungi cultivated on both grasses and soluble sugars[6,29]. The number of biosynthetic genes represented in the transcriptome is indicated by the gray bars, and the number of genes absent from the transcriptome (silent) are represented by empty bars. The percentage of transcribed genes is presented by the black triangles (secondary axis).
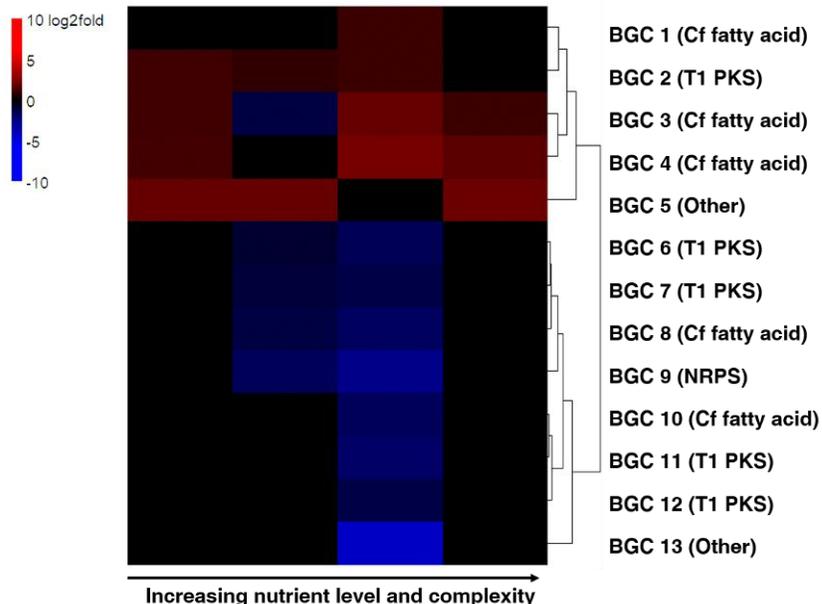
Figure 2.6: Nutrient availability and complexity regulates the expression of 13 core biosynthetic genes of *N. californiae*. BGC=biosynthetic gene cluster. The heatmap shows the log2fold change of the transcript abundance of *N. californiae* grown in nutrient-poor media formulations relative to highly complex media. Only statistically significant, differentially regulated core biosynthetic genes are shown (log2fold change≥1, $P \leq 0.01$, absolute). *N. californiae* cultures were grown to early stationary phase in quadruplicate in a minimal media supplemented with nutrients of increasing level and complexity. Columns of heatmap from left to right supplemented with (1) no supplement, (2) yeast extract and BactoTM Casitone, (3) rumen fluid, (4) yeast extract, BactoTM Casitone, and rumen fluid.

Another indicator of active genes in early-diverging fungi is the presence of adenine N6-methylation marks on the promoter regions[167]. Dense methylated adenine clusters (MACs) were observed within 500 bp of the transcription start site of 6 out of 13 core biosynthetic genes for *P. finnis*, 8 out of 14 for *C. churrovis*, and 3 out of 46 core genes for *A. robustus*. In addition, neighboring genes in the gene clusters were also marked by MACs: 34 neighboring genes in *A. robustus* clusters, 33 in *P. finnis* clusters, and 43 in *C. churrovis* clusters. The Type 1 PKS genes identified by antiSMASH were highly methylated: 5 out of 7 PKS genes from *P. finnis* and 5 out of 6 from *C. churrovis*. These data corroborate the

transcriptomic evidence that anaerobic gut fungi actively transcribe a significant portion of their backbone genes and associated gene clusters during standard laboratory cultivation.

Finally, we searched for detectable proteins in both the membrane-bound and cytosolic fractions of fungal intracellular proteins. Proteomics confirmed that that at least 30% of the total biosynthetic enzymes from all four strains are translated into protein, thus increasing the likelihood that the secondary metabolism of anaerobic gut fungi is functionally active during laboratory cultivation. Notably, all copies of PKSs belonging to family 1 were expressed. In addition, PKS families 5 and 7, which were represented in the genomes of *C. churrovis*, *N. californiae*, and *P. finnis* were also expressed in these three strains. Across all lines of evidence (transcriptomics, N6-methylation, and proteomics), 53% of core biosynthetic genes were active by at least one metric.

2.2.6   *Three groups of novel natural products from A. robustus and N. californiae are visualized via molecular networking of MS/MS spectra*

To further validate that anaerobic gut fungi synthesize natural products, we analyzed the nonpolar metabolites of *A. robustus*, *C. churrovis*, *N. californiae*, and *P. finnis* by LC-MS/MS. We first built molecular networks using the Global Natural Products Social Molecular Networking (GNPS) platform[168] to distinguish groups of natural products based on LC-MS/MS datasets collected for *A. robustus* and *N. californiae*. To discriminate between compounds secreted by *A. robustus* or *N. californiae* and compounds already present in the complex growth medium or released from autoclaving the reed canary grass growth substrate, we constructed a molecular network showing separate conditions for secreted nonpolar metabolites from *A. robustus* or *N. californiae* and compounds from a control of complex growth medium. The majority of nodes in three clusters of the network

(Figure 2.7) were only present in the anaerobic fungal strains. None of the nodes matched the spectral libraries in GNPS. Similarly, we constructed a molecular network of the nonpolar metabolites of *C. churrovis* and *P. finnis* and observed a cluster of 12 nodes present in the fungal supernatant, but absent from the control and spectral libraries in GNPS. These findings support the hypothesis that the anaerobic gut fungi produce novel, strain-specific as well as conserved secondary metabolites. Since the cultivation of anaerobic fungi requires the use of complex medium that contains a small amount of clarified rumen fluid, which is expected to harbor a low concentration of background secondary metabolites secreted by native microbes, there may be additional natural products that are present in both fungal supernatant and the growth medium.

2.2.7 *Anaerobic gut fungi produce the polyketide-related antioxidant baumin*

Among the 72 compounds detected from *A. robustus*, one had fragmentation spectra and exact mass consistent with the styrylpyrone baumin. Styrylpyrones are found in mushrooms, especially medicinal mushrooms, and are thought to have roles similar to those of flavonoids in plants, such as antioxidants. Baumin itself was first detected as a product of the fungus *Phellimus baumii* (now renamed *Sanghuangporus baumii*) from the distantly related fungal phylum Basidiomycota[169,170]. This compound, putatively identified as baumin, was also produced by *C. churrovis*, *N. californiae*, and *P. finnis.* In all strains, it was observed at 10-fold or greater intensity in the supernatant of fungal cultures compared to the growth medium.

We also used SIRIUS 4.0 [171] and CANOPUS[172] to predict the structure and class of the observed compound. Rather than baumin, SIRIUS predicted a flavonoid, whereas CANOPUS predicted a hydroxyflavonoid. However, upon inspection of the metabolic

39

pathways annotated in the MycoCosm portal for *A. robustus*, we found that *A. robustus* lacked any genes encoding the biosynthetic enzymes of flavonoids. Furthermore, sequence alignment of flavonoid biosynthetic enzymes in higher-order fungi to the predicted proteins from *A. robustus* identified no homologs. Therefore, baumin remained the top candidate for the unknown compound. The putative baumin is the first secondary metabolite directly detected from anaerobic gut fungi, and it may serve the anaerobic gut fungi as an oxygen scavenger in the rumen of the host animal after forage intake. Hobson reported up to 0.6% oxygen transiently present in the gaseous phase of the rumen[5,173]. The gene cluster responsible for the production of baumin in *S. baumii* is not known at this time, which limited our ability to assign the gene cluster in anaerobic gut fungi. However, antiSMASH predicted only one PKS gene cluster (accession OCB83923.1) with more than two domains from the *S. baumii* genome. The core biosynthetic gene was a hybrid NRPS-Type I PKS with PKS architecture of KS-AT-DH-KR-ACP (ketosynthase-acyltransferase-dehydratase-ketoreductase-acyl carrier protein). The domain architecture of this PKS is similar to PKS family 4 of the anaerobic fungi, although some members of this family lack the dehydratase domain (PKS genes from *A. robustus* scaffold 127, *C. churrovis* scaffold 129, and *N. californiae* scaffold 428). However, protein sequence similarity is only ~30% between members of PKS family 4 from anaerobic fungi and the *S. baumii* PKS. Experimental validation of the gene cluster responsible for the synthesis of baumin is still necessary.

A

B

C

D. Baumin m/z 523.123
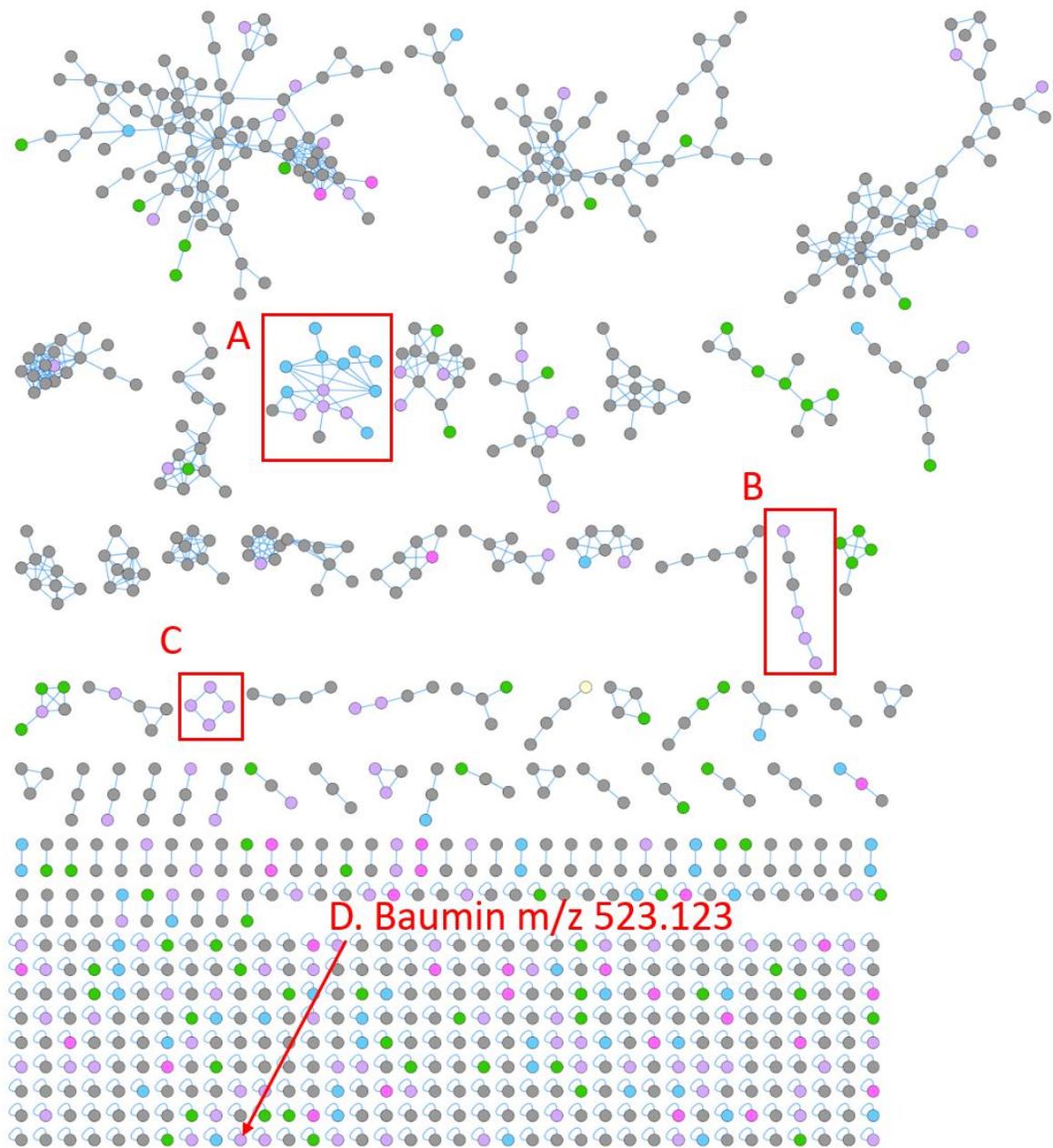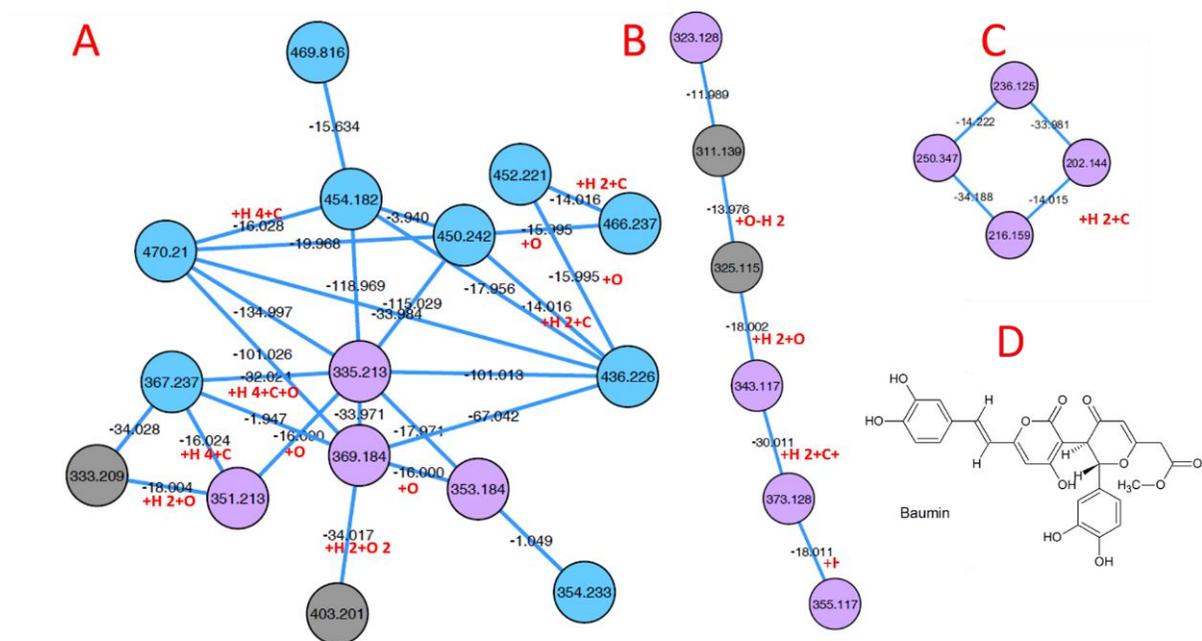
Figure 2.7: The molecular network generated from nonpolar untargeted metabolomics of *A. robustus* and *N. californiae* illustrates both known and novel natural products. Red rectangles enclose putative novel natural product clusters (A, B, and C) and baumin (D). Clusters A, B, and C are magnified below the network and the chemical structure of baumin in shown in D. Node colors are as follows: blue=feature detected in *N. californiae* supernatant only, pink=*A. robustus* supernatant only, lilac=*A. robustus* and *N. californiae,* green=control only (autoclaved and incubated grass in liquid growth medium), gray=fungal supernatant and control. Self-looping nodes were truncated below baumin.

2.3   Conclusions

Integrated 'omics' analysis of anaerobic gut fungi revealed the untapped potential of these non-model organisms as secondary metabolite producers. Species of anaerobic gut fungi from four distinct genera (*Anaeromyces*, *Caecomyces*, *Neocallimastix*, and *Piromyces*) possess the biosynthetic enzymes for polyketides, nonribosomal peptides, bacteriocins, and other natural product classes. The number of detected backbone genes per fungus is on the same order of magnitude as the prolific aspergilli. Upon inspection, some of the biosynthetic genes of anaerobic fungi were similar to bacteria, suggesting the possibility of horizontal gene transfer between fungi and bacteria in the rumen microbiome. HGT was further supported by the fact that in phylogenetic trees of NRPS condensations domains as well as

bacteriocins the fungal genes nested within or were sister to bacterial genes. Although many of the biosynthetic genes of anaerobic fungi were similar to bacteria, their regulation may still be typical of fungal secondary metabolism. Homologs of velvet regulatory proteins, which are known to link fungal development and secondary metabolism in filamentous fungi, were identified in the predicted proteins of anaerobic fungi. PKS genes identified within the fungal genomes were highly conserved between strains, indicating that polyketides may serve important biological functions for anaerobic fungi. Even during standard laboratory growth, transcriptomics and proteomics has demonstrated that much of their secondary metabolism is active. LC-MS/MS detected numerous secondary metabolites, including a compound putatively identified as the styrylpyrone baumin. Further experiments will be necessary to decipher the functions of the secondary metabolites of anaerobic fungi, but, among many possibilities, they may serve as regulators of the fungal life cycle or defense or compounds against bacterial competitors. In addition to their native function, natural products from anaerobic gut fungi are a promising source of novel antimicrobial peptides, antibiotics and therapeutics.

2.4   Materials and Methods

2.4.1   Routine cultivation of anaerobic gut fungi

A. robustus, *C. churrovis*, and *N. californiae* were isolated via reed canary grass enrichment from the feces of sheep or goat at the Santa Barbara Zoo, as described previously[6,7,29]. *P. finnis* was isolated from the feces of a horse at Verrill Farm Stables in Concord, MA, USA [6,7,29]. The fungal strains were routinely transferred every 3-4 days into fresh reduced liquid medium with 0.1 g of 4-mm milled reed canary grass as growth

substrate. *P. finnis* was cultivated in Medium C[174]. *A. robustus, C. churrovis*, and *N. californiae* were cultivated in a reduced formulation of Medium C containing 0.25 g of yeast extract (Thermo Fisher Scientific), 0.5 g Bacto™ Casitone, and 7.5% clarified rumen fluid.

2.4.2 *Mining fungal genomes for biosynthetic gene clusters using antiSMASH*

FASTA genome files for *A. robustus*[6], *C. churrovis* (Brown et al., manuscript in preparation), *N. californiae*[6], and *P. finnis*[6] (available from the MycoCosm[23] portal) were submitted separately to the antiSMASH 3.0 server[7,107,167]. Default parameters were used, and the ClusterFinder analysis option selected. Since *C. churrovis* was sequenced after the launch of antiSMASH 4.0, it was analyzed by the legacy command line version of antiSMASH 3.0.

2.4.3 *Protein BLAST analysis of core biosynthetic genes against NCBI non-redundant databases*

Core biosynthetic genes predicted by antiSMASH[106] were queried against the NCBI non-redundant protein databases using version 2.7.1 of the command line implementation of BLAST+[158], excluding Neocallimastigomycota from the results. Thresholds for hits were as follows: minimum E-value 10^{-8}, minimum qcovhsp 25%, and minimum identity 30%. The top hit for each query was considered to be the hit with the highest bitscore within these thresholds. The biosynthetic core gene in the reading frame with the most predicted domains was queried, except in cases where ClusterFinder[144] predicted multiple genes. In this case, all biosynthetic genes were searched. If the top hit for all genes was from the same taxonomic phylum, the result was counted once (e.g. cluster 3 on scaffold 152 of *N.*

californiae). Otherwise, each phylum was counted. AntiSMASH genes containing fewer than three domains were not included in this analysis.

2.4.4 *Horizontal gene transfer analysis of PKS domains, NRPS condensation domains, and bacteriocins*

PKS ketosynthase and NRPS condensation domains were selected based on PFAM[175] annotations of protein sequences from the corresponding genomes. Selected sequences were used to search for homologs using BLAST+[158] against NCBI's non-redundant database (downloaded July 2019) with an E-value threshold of $1x10^{-5}$. Additionally BLAST analysis was performed against fungal proteins from MycoCosm database[23], excluding sequences belonging to Neocallimastigomycota clade. For each ketosynthase domain we constructed a phylogenetic tree based on selected homologs. We selected up to 10 best hits from 1) prokaryota, 2) non-fungal eukaryotes and 3) non-Neocallimastigomycota fungal proteins, such that the maximum number of sequences used for each tree was 31. Sequences were aligned using MAFFT [176] with subsequent removal of non-reliable aligned positions using trimAl[177]. Phylogenetic trees were constructed using FastTree[178] and RAxML [179]. We considered as potential non-fungal HGT events the cases when a given Neocallimastigomycota domain was nested within a prokaryotae or non-fungal eukaryota clade and all branches of the tree had at least a 70% of bootstrap support values. Analysis of the bacteriocins was performed similarly, but the E-value threshold was relaxed to 0.1 and MMETSP[162] was queried in addition to MycoCosm and NCBI's non-redundant databases.

2.4.5 Identification of velvet homologs

The velvet domain protein family has been previously established in filamentous fungi[91]. The velvet domain family proteins were assigned to a Pfam[180], PF11754, as well as InterPro[181] family IPR021740. Filtered model proteins from of *A. robustus*, *C. churrovis*, *P. finnis*, and *N. californiae* belonging to PF11754 or IPR021740 were searched using the MycoCosm portal[23]. Two proteins were identified with the velvet motif in *C. churrovis*. Protein Id 623244 was a homolog of *vosA* from *Aspergillus nidulans* (accession ABI51618). Proteins were not identified in *A. robustus*, *P. finnis*, and *N. californiae* by Pfam domain search. To further expand the search, we used *vosA* from *A. nidulans* to query the other genomes by a protein BLAST+ search against all protein models in the remaining three strains of anaerobic gut fungi. Through this approach, we identified putative velvet proteins from *N. californiae* and *P. finnis*. The presence of a velvet domain was confirmed in all putative velvet proteins by CD-Search[182].

2.4.6 Classification of type I PKS genes into families and assessment of PKS gene cluster transcription

PKS genes predicted by antiSMASH [106] and SMURF [109] for *A. robustus*, *C. churrovis*, *N. californiae*, and *P. finnis* were grouped into families by OrthoFinder[165] with default parameters. OrthoFinder defines an orthogroup as a group of genes that are descended from a common gene in the last common ancestor. The advantages of OrthoFinder include the ability to remove sequence length bias from sequence similarity scores as well as the ability to define orthogroup similarity limits[165]. Further details about the OrthoFinder algorithm can be found in the listed reference[165]. The OrthoFinder algorithm can be downloaded from http://www.stevekellylab.com/software/orthofinder. Transcriptionally active genes in the

PKS gene clusters were determined using the metric of at least 5x coverage by RNA-seq reads over more than 95% of the length of the gene during standard laboratory cultivation described previously[6,29].

2.4.7 Curation of the biosynthetic genes using RNA-seq models

All biosynthetic genes were manually curated to ensure that they are fully supported by RNA-seq data. In the case of incomplete models, we used the BRAKER1 pipeline [183], which combines usage of RNA-seq read alignments with GeneMark-ET and AUGUSTUS gene finding to extend the gene models to completeness. RNA-seq data was previously acquired from fungal cultures grown on grasses and soluble sugars[6,29]. Clusters were delineated starting with the core biosynthetic gene and determining all genes in the 5' and 3' direction with at least 5x coverage across at least 95% of the gene length by RNA-seq reads and within 10 kbp consecutive intergenic distance of their neighbor.

2.4.8 Quantification of biosynthetic gene transcription and N6-methylation

Transcriptome assemblies for *A. robustus*, *C. churrovis*, *N. californiae*, and *P. finnis* were described previously[6,29]. To determine the number of transcribed biosynthetic genes, local BLAST libraries were prepared from the antiSMASH amino acid predictions of biosynthetic genes (e.g. PKS, NRPS). A protein BLAST was performed locally using Blast2GO[184] with the respective transcriptome as the subject sequences [184]. The biosynthetic genes were considered transcribed if hits were returned with E-value cut-off 0.001, similarity greater than 95%, and Hsp/Query greater than 95%. Only antiSMASH genes containing three or more catalytic domains (e.g. adenylation) with at least 100 amino acids were queried. Dense methylated adenine clusters were quantified as described previously[185].

Predicted proteins of *A. robustus* and *P. finnis* BGCs were queried for MACs within 500 bp of the transcription start site. N6-methyldeoxyadenine positions were collected either from previously published data [167] (*P. finnis* and *A. robustus*) or using the Sequel I PacBio sequencing platform (*C. churrovis*) followed by analysis using the smrtlink v8.0.0.80529 pb_basemods workflow. Following detection, 6mA modified sites were filtered and Methylated Adenine Clusters (MACs) were identified as described previously [167].in Mondo et al., 2017. Dense methylated adenine clusters were quantified as described previously [185]. Promoters were considered methylated if MACs were present within 500 bp of the transcription start site.

2.4.9   Sample preparation for proteomics

Anaerobic serum bottles containing 40 mL of Medium C [174] and 5 g/L cellobiose (Fisher Scientific) were preheated to 39 °C and then inoculated with 1.0 mL each of *A. robustus*, *C. churrovis*, *N. californiae* or *P. finnis* from routine passaging. After 3 days of growth (4 days for *P. finnis*), 1.0 mL of each culture was used to inoculate each serum bottle for proteomics. Each serum bottle contained 80 mL of medium C with 5 g/L cellobiose (Fischer Scientific) as the carbon source. For each fungal strain, six 80 mL cultures were prepared. After three days of growth, *C. churrovis* cultures were harvested and after six days of growth, cultures of *A. robustus*, *N. californiae*, and *P. finnis* were harvested. The cultures were transferred into 50 mL Falcon® tubes and contributed at 3200 *g* and °C for 10 min using a swinging bucket rotor (Eppendorf™ A-4-81). The supernatant was removed and each pellet was washed with 5.0 mL of pH 7.4 phosphate buffered saline (PBS) solution and centrifuged again to remove the PBS. Samples were frozen at -80 °C until the time of extraction.

For proteomic extraction, all chemicals were obtained from Sigma Aldrich (St. Louis, MO) unless otherwise noted. Fungal cell pellets were extracted utilizing a method similar to MPLEx[186] where the pellets were suspended in 5 mL ice-cold water, 6.75 mL methanol and homogenized with a disposable probe homogenizer (Omni International, Kennesaw, GA). Ice-cold chloroform was then added to the homogenate so that the chloroform:methanol:water ratio was 8:4:3 and the samples were vigorously vortexed for 1 minute. The samples were placed on ice for 5 minutes, followed by another 1-minute vortex step. The samples were then centrifuged at 5,000 g, 4 °C for 10 minutes. The top (polar metabolite) layer of the tri-phasic separation and the lower (non-polar metabolite) layer were not used in this study. The protein interphase pellet was washed wit 1 mL of ice-cold methanol by vortexing and centrifuging as above. The supernatant was removed and disposed, and the pellets allowed to dry slightly. The pellets were reconstituted in an 8M urea, 100 mM NH_4HCO_3 buffer, and a Bicinchoninic Acid (BCA) protein assay (Thermo Fisher Pierce, Waltham, MA) was performed to quantify the protein content in the pellet. 1 mg of protein was utilized for digestion, normalized to the same volume for all samples. Dithiothreitol was added to a 5 mM concentration in each sample, and the samples were incubated at 37 °C for 1 hour with shaking at 800 rpm on a thermomixer (Eppendorf, Hauppauge, NY). The samples were then diluted 8-fold with 50 mM NH_4HCO_3 and calcium chloride was added to a concentration of 1 mM. Trypsin was added in a 1:50 (w:w) trypsin:protein ratio and samples were incubated for 3 hours at 37 °C. Samples were centrifuged at 5000 g at room temperature for 10 minutes to pellet particulate material and the clarified samples were then subjected to C18 solid phase extraction (SPE) cleanup as described previously [7] with the exception that a Strata C18-E 50 mg column (Phenomenex,

49

Torrance, CA) was used. The samples were then diluted to a concentration of 0.1 ug/uL for LC-MS analysis.

2.4.10 Proteomics mass spectrometry and data analysis

Separation was performed prior to mass spectrometry by a Waters nano-Acquity M-Class dual pumping UPLC system (Milford, MA) using a 5 µL injection at 3 µL/min with reverse-flow elution onto the analytical column at 300 nL/min. The gradient profile of mobile phases of (A) 0.1% formic acid in water, and (B) 0.1% formic acid in acetonitrile was the following (min, %B): 0, 1; 2, 8; 20, 12; 75, 30; 97, 45; 100, 95; 110, 95; 115, 1; 150, 1. The trapping column was 150 µm ID and 4 cm in length, and the analytical column was 75 µm ID and 70 cm in length. The columns were packed in-house using 360 µm OD fused silica (Polymicro Technologies Inc., Phoenix, AZ) with 2-mm sol-gel frits for media retention and contained Jupiter C18 media (Phenomenex, Torrence, CA). Particle sizes for the trapping and analytical columns were 5 µm and 3 µm, respectively.

Proteomics data was collected on a Q-Exactive Plus mass spectrometer (Thermo Scientific, San Jose, CA) with a homemade nano-electrospray ionization interface. The electrospray emitters were prepared from 150 µm OD x 20 µm ID chemically etched fused silica[187]. The spray voltage was 2.2 kV, and the ion transfer tube temperature was 250 °C. Data were collected for 120 min after a 20 min delay from time of sample injection and trapping. FT-MS spectra were acquired with a resolution of 35k (AGC target 3e6) from 400-2000 m/z. The top 12 FT-HCD-MS/MS spectra were acquired in data dependent mode, excluding singly charged ions, with a resolution of 17.5 k (AGC target 1e5) and an isolation window of 2.0 m/z using a normalized collision energy of 30 and exclusion time of 30 s.

Proteomics data analysis was performed as previously described [7], with the exception that peptide fragments were mapped to the transcriptomes [6,29] translated into all open reading frames as well as to the antiSMASH-predicted biosynthetic enzymes.

2.4.11 Preparation of fungal supernatant samples for metabolomics

Routinely passaged *A. robustus*, *C. churrovis*, *N. californiae*, or *P. finnis* was inoculated into 60 mL serum bottles (VWR International) containing 40 mL of anaerobic, autoclaved Medium C[174] with 0.4 g of 4-mm reed canary grass. After three days of growth at 39 °C, these seed culture were used to inoculate four replicate Hungate tubes per fungal strain containing 9 mL of anaerobic, autoclaved Medium C and 0.1 g of 4-mm milled reed canary grass. For each serum bottle or Hungate tube, 1.0 mL of fungal inoculum was used. Four Hungate tubes containing anaerobic, autoclaved Medium C containing reed canary grass were incubated at 39 °C for use as a control. Cultures and controls were harvested after six days of incubation, centrifuged at 3220 *g* for 10 min with the swinging bucket rotor (Eppendorf F-34-6-38) at 4 °C, and the fungal supernatant was frozen at -80 °C for exo-metabolomics analysis.

2.4.12 Extraction and LC-MS/MS of nonpolar metabolites

Ethyl acetate extraction of nonpolar metabolites from fungal supernatant was performed as follows: 2 mL ethyl acetate was added to 1.5 to 2 mL fungal supernatant, vortexed and sonicated for 10 min in a water bath (room temperature), centrifuged (5 min at 5000 rpm), then top ethyl acetate layer removed to another tube. For 2 mL supernatant samples, this process was repeated with another 2 mL ethyl acetate added to the sample, then top layer

removed and combined with the previous extract. Extracts were dried in a SpeedVac (SPD111V, Thermo Scientific) and stored at -20 °C.

In preparation for LC-MS analysis, 150 to 300 µL LC-MS grade methanol containing 1 µg/mL internal standard (2-Amino-3-bromo-5-methylbenzoic acid, Sigma) was added to dried extracts, followed by a brief vortex and sonication in a water bath for 10 minutes, then centrifugation for 5 minutes at 5000 rpm. 150 µL of resuspended extract was then centrifuge-filtered (2.5 min at 2500 rpm) using a 0.22 µm filter (UFC40GV0S, Millipore) and transferred to a glass autosampler vial. Reverse phase chromatography was performed by injecting 2 µl of sample into a C18 chromatography column (Agilent ZORBAX Eclipse Plus C18, 2.1x50 mm 1.8 µm) warmed to 60°C with a flow rate of 0.4 mL/min equilibrated with 100% buffer A (100% LC-MS water with 0.1% formic acid) for 1 minute, followed by a linear gradient to 100% buffer B (100% acetonitrile with 0.1% formic acid) at 7 minutes, and then held at 100% B for 1.5 minutes. MS and MS/MS data were collected in both positive and negative ion mode using a Thermo Q Exactive HF mass spectrometer (ThermoFisher Scientific, San Jose, CA), with full MS spectra acquired ranging from 80-1200 m/z at 60,000 resolution, and fragmentation data acquired using an average of stepped collision energies of 10, 20 and 40 eV at 17,500 resolution, and 20, 50 and 60 eV for a single replicate. Mass spectrometer source settings included a sheath gas flow rate of 55 (au), auxiliary gas flow of 20 (au), sweep gas flow of 2 (au), spray voltage of 3 kV and capillary temperature of 400 degrees C. Sample injection order was randomized and an injection blank of methanol only run between each sample. Raw data is available for download at https://genome.jgi.doe.gov/portal/.

2.4.13 Metabolomics data analysis and identification of baumin from A. robustus

Peak-finding was performed with MZmine[188]. Filtering the features that were at least four-fold higher in intensity than the control resulted in 116 peaks. Metabolite Atlas workflow tools were used to remove features that were visually recognized as artifacts of the peak-finding process (isotopes, background ions, unnecessary adducts, etc.) and to refine retention times for feature integration across files[189]. These features were further filtered to those with sample/control ratios of at least 4.

Each MS2 fragmentation spectrum was searched with Pactolus, an in-house implementation of the MIDAS[190] scoring algorithm. The MIDAS scoring algorithm traverses the possible fragmentation paths and scores an identification based on *m/z* values matching structures along a fragmentation path. The Pactolus implementation precomputes and stores the fragmentation paths for each molecule, whereas the MIDAS algorithm computes the fragmentation paths on the fly. The Pactolus code can be found at https://github.com/biorack/pactolus. Suggested identifications for each MSMS spectrum were mapped back to the filtered list of features above. For the feature *m/z* 523.1232 at 4.7 min, the compound baumin was identified as a top hit as calculated by Pactolus. To investigate the likelihood of this identification, an isotope simulation using XCalibur 2.2SP1 (Thermo Scientific, San Jose) was used to determine potential molecular formulas of the feature. The formula $C_{27}H_{22}O_{11}$, which corresponds to baumin, was calculated as a protonated adduct with an error of -0.288 ppm. Further support for this identification is given by the MS/MS spectra, with detected *m/z* supporting likely fragments of the baumin structure. Peak-finding was performed as above for *C. churrovis*, *N. californiae*, and *P. finnis* to ascertain whether baumin was also produced by those strains.

The MS-GF+ (1.3.0) Workflow was used in the ProteoSAFe web server, available through the Center for Computational Mass Spectrometry, to run structure and class prediction by SIRIUS 4.0[171] and CANOPUS[172], respectively. The job is publically available at the following URL:

https://proteomics2.ucsd.edu/ProteoSAFe/status.jsp?task=6e6cf05848424ca1bdba0f5b48ccafc6

We checked *A. robustus* for its genetic capability to produce flavonoids. No gene models were assigned to the Flavonoid Biosynthesis KEGG[191] map, which is a component of MAP01060 on the MycoCosm portal[23] The following genes related to flavonoid biosynthesis in higher-order fungi, selected from Supplementary File 1 of the reference by Mohanta[192], were queried using BLAST+[158] against filtered model proteins of *A. robustus* in the MycoCosm portal[23]: RAQ52167.1 (chalcone synthase, *Aspergillus flavus*), PYH41479.1 (chalcone isomerase, *Aspergillus saccharolyticus*), PIG83686.1 (dihydroflavonal-4-reductase, partial, *Aspergillus arachidicola*), CEL03464.1 (Isoflavone reductase family protein, *Aspergillus calidoustus*), RAH85834.1 (quercetin 2,3-dioxygenase anaerobically complexed with the substrate kaempferol, *Aspergillus japonicus* CBS 114.51), RAQ53266.1 (leucoanthocyanidin dioxygenase, *Aspergillus flavus*), GAO86351.1 (myricetin O-methyltransferase, *Aspergillus udagawae*), CCE28660.1 (related to naringenin, 2-oxoglutarate 3-dioxygenase, *Claviceps purpurea*), 20.1, RAQ48556.1 (quercetin 2, *Aspergillus flavus*). There were no hits for any of these sequences (E-value threshold 10^{-5}).

2.4.14 Visualization of nonpolar untargeted metabolomics data via molecular networking

Molecular networks were constructed from the tandem mass spectrometry data described above. Two separate networks were constructed: one for *A. robustus* and *N. californiae* and a medium C control, and the other for *C. churrovis, P. finnis* and a medium C control. Molecular networks were created using the online workflow at GNPS[168]. All MS/MS peaks within +/- 17 Da of the precursor m/z were removed. The MS/MS spectra were further filtered by selecting only the top six peaks in the +/- 50 Da window throughout the spectrum. Consensus spectra were created by clustering the data using MS-Cluster [193] with a MS/MS fragment ion tolerance of 0.5 Da and a parent mass tolerance of 2.0 Da. Consensus spectra with fewer than two spectra were eliminated. A network was then created with the following criteria used for the edges: 1) cosine score above 0.7, and 2) greater than six matched peaks. Furthermore, edges between nodes were only considered in the network if and only if each of the consensus spectra represented by the nodes were part of other node's top 10 most similar nodes. The spectra in the network were subsequently queried against the spectral libraries of GNPS. The library spectra were filtered to be consistent with the input data. Only pairings between network spectra and library spectra with scores above 0.7 and at least six matched peaks were kept.

The mass spectrometry data were deposited on public repository MassIVE (MSV000085907).

The GNPS jobs for *A. robustus* and *N. californiae* can be accessed at:

http://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=af74716b912a435eb53c1307a1dad092

The GNPS job for *C. churrovis* and *P. finnis* can be accessed at:

https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=6fb6a9367cf34b669b8bc00862541af9

3 Co-cultivation of the anaerobic fungus *Anaeromyces robustus* with *Methanobacterium bryantii* enhances transcription of carbohydrate active enzymes

Adapted from *Journal of Industrial Microbiology & Biotechnology*, Vol 46, Candice L. Swift, Jennifer L. Brown, Susanna Seppälä, Michelle A. O'Malley, Co-cultivation of the anaerobic fungus *Anaeromyces robustus* with *Methanobacterium bryantii* enhances transcription of carbohydrate active enzymes, 1427-1433, Copyright 2019, with permission from Springer.

3.1 Introduction

Anaerobic fungi isolated from the guts of herbivores produce an abundance of biomass-degrading enzymes capable of breaking down a wide range of plant matter into fermentable sugars[6]. These microbes are responsible for degrading up to 50% of the biomass ingested by herbivores, through a combination of mechanical breakdown by rhizoidal growth and a rich array of secreted enzymes[194–198]. Gut fungi produce a synergistic combination of enzymes responsible for breaking down and modifying plant poly- and oligosaccharides, classified as Carbohydrate Active Enzymes (CAZymes)[6,199]. Cellulase, beta-glucosidase, and hemicellulase activities in enzymatic secretions from gut fungi are comparable to those observed in engineered preparations of *Aspergillus* sp. and Trichoderma sp., making anaerobic gut fungi attractive candidates for bioprocessing applications[6].

While the biomass-degrading capabilities of anaerobic fungi alone are impressive, studies have shown that the effectiveness of cellulose or plant biomass breakdown can be

56

enhanced through co-culture with methanogenic archaea[119,120,200–203]. For example, co-cultivation of the fungi *Piromonas communis* and *Neocallimastix frontalis* with *Methanobrevibacter smithii* on lucerne stem resulted in enhanced removal of xylose and glucose by *P. communis* and enhanced removal of glucose by *N. frontalis*[203]. A similar study using these same pairings as well as a *Sphaeromonas communis* paired with *M. smithii*, revealed that the extent and rate of xylan utilization increased in co-culture[202]. In addition to these studies, the literature widely reports that co-cultivation of anaerobic gut fungi with methanogens fosters a synergistic relationship where methanogens remove hydrogen and convert it to methane[117,119,120,122,201,204]. This partnership drives increases in the production of acetate, CO_2, and CH_4 and decreases formate, lactate, malate, succinate, and ethanol, resulting in a shift in production from the less valuable electron sink products of the anaerobic fungi to methane, a source of renewable energy[117,119,120,122,201,204,205].

Reported advantages encourage the utilization of fungus-methanogen co-cultures and consortia in bioprocessing applications, yet it remains unknown whether the reported impacts of co-culturing are due to increased fungal growth, elevated transcription of enzymes in co-culture, or other factors. While previous work measured alterations in biomass degradation[116–120], sugar utilization[204], metabolite profiles[116–120,204], and enzyme activity[119] when anaerobic fungi are co-cultured with methanogens, the mechanistic underpinnings for these differences remain unknown. Notably, there is growing evidence that unlike other fungi, the anaerobic fungi organize part of their secreted enzymes in cellulosome-like structures that contain multiple different enzyme functionalities[6,7]. Fungal cellulosomes assemble by sequence-divergent non-catalytic dockerin domains, but whereas similar structures are known from bacteria, details of fungal cellulosome structure, function, and regulation are still not well understood[7,206,207]. Nevertheless, a previous study discovered

that the majority of the dockerin-containing proteins in three anaerobic gut fungal strains were secreted CAZymes, using a combination of genomic and proteomic techniques[7]. Based on the observed increase in fungal biomass breakdown for co-cultures of fungi with methanogenic archaea, it is possible that fungal cellulosome transcription increases when the fungi are co-cultivated with a methanogen.

Here, we present the first examination of global transcriptomic response and CAZyme regulation in anaerobic fungi cultivated with associated methanogens. Co-cultivation of the anaerobic fungus *Anaeromyces robustus* with methanogen *Methanobacterium bryantii* suggests that elevated transcription plays a key role in increased biomass degradation that is often observed for such co-culture partnerships. The specific gene sets responsible for shifts in CAZyme production and altered metabolism are identified, which can be further exploited for use in the bioproduction of value-added products obtained from biomass breakdown. This finding also invites new insights and follow-up studies to probe the interwoven metabolism and interactions between microbes in the rumen microbiome.

3.2   Results and Discussion

Anaerobic gut fungi possess a wealth of carbohydrate active enzymes (CAZymes) [6,7,29,196] that could be harnessed for lignocellulosic deconstruction and bio-based manufacturing. In the guts of large herbivores, these fungi are associated with methanogenic archaea[208,209], yet it remains unclear how association with methanogens impacts fungal behavior and degradation activity. Based on previous observations, we hypothesized that in the presence of methanogens, anaerobic gut fungi upregulate their CAZyme production to amplify biomass-degradation activity[202,203]. To test this hypothesis, we cultivated a synthetic consortium of the model anaerobic fungus *Anaeromyces robustus* with the methanogen

Methanobacterium bryantii (DSM No.-863, DSMZ) and performed RNA-sequencing on the fungus to determine whether CAZymes are transcriptionally activated in co-culture. Details of the co-culture preparation are provided in Figure 3.1.
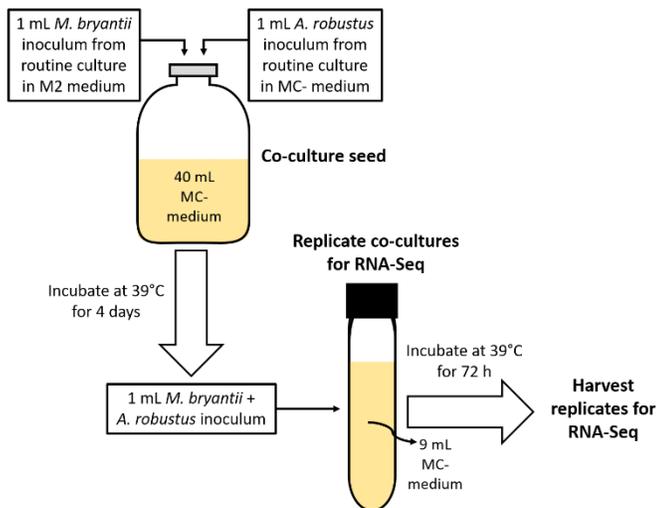
Figure 3.1: Schematic of procedure for co-cultivation of *A. robustus* with *M. bryantii* in preparation for RNA-Seq.

We found that 421 genes were positively regulated and 364 genes were negatively regulated when *A. robustus* was grown in co-culture with *M. bryantii* compared to *A. robustus* grown in isolation (p-adjusted less than 0.05 and greater than the absolute of one log2fold change). Previous work found that the *A. robustus* transcriptome (grown on glucose, reed canary grass, Avicel™, cellobiose, and filter paper) contains at least 17,127 transcripts[6]. Thus, approximately 5% of total *A. robustus* genes were differentially regulated in co-culture. Specifically, we found that 105 CAZymes (12% of the total predicted CAZymes of *A. robustus*) were upregulated (Figure 3.2) and 29 were downregulated. Upregulated genes predicted proteins with a wide array of activities on various polysaccharides including both cellulose and hemicellulose (Figure 3.3). 19 glycoside

hydrolase, 2 glycocosyltransferase, 6 carbohydrate esterase, and 2 polysaccharide lyase families were represented in the upregulated genes.
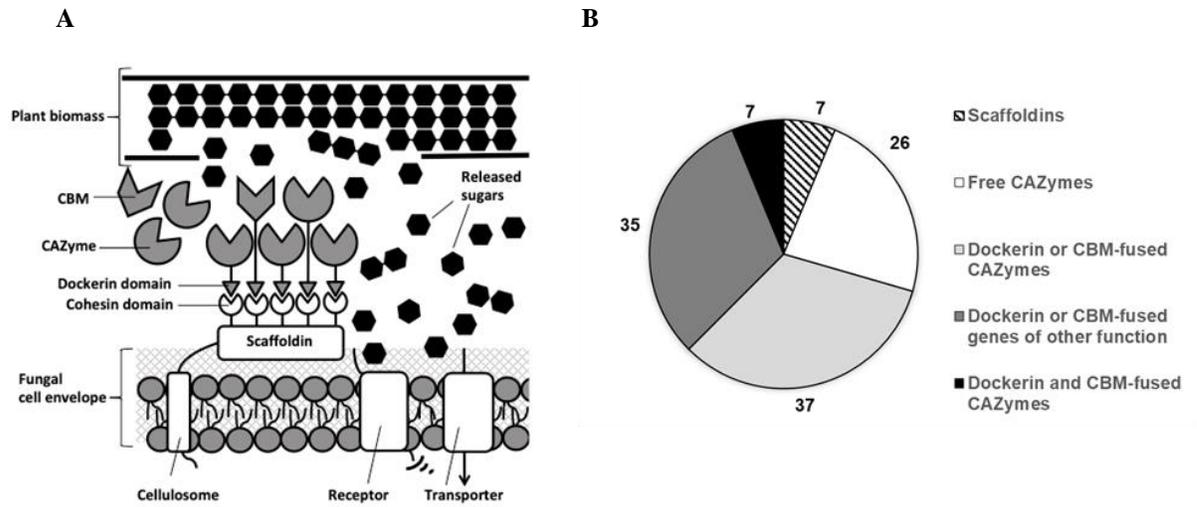
Figure 3.2: Co-cultivation of the anaerobic fungus *A. robustus* with methanogen *M. bryantii* upregulates genes encoding fungal biomass-degrading machinery. (a) Artistic depiction of free carbohydrate active enzymes (CAZymes) and components of fungal cellulosomes degrading plant biomass, (b) Number of upregulated genes associated with biomass degradation in fungal-methanogen co-culture relative to fungal monoculture. Transcripts that encode for components of cellulosomes (scaffoldins, dockerin-fused) are indicated, as well as soluble (free) CAZymes and those associated with carbohydrate binding modules (CBMs).
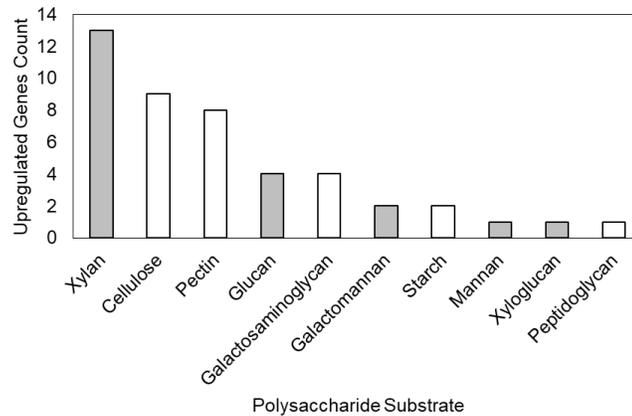
Figure 3.3: Co-cultivation of *M. bryantii* with *A. robustus* upregulates CAZyme genes associated with degradation of diverse polysaccharides. Gray bars highlight polysaccharide components of hemicellulose. Only genes in CAZyme families with activity unique to one type of polysaccharide substrate were counted in this analysis.

Anaerobic fungi feed by secreting biomass-degrading enzymes to the environment, but unlike other fungi, the anaerobic fungi organize part of their secreted enzymes in cellulosomes. The integral building blocks of fungal cellulosome were recently identified in and include CAZymes with carbohydrate binding modules (CBMs) and fungal dockerin domains: the dockerin domains interact with cohesin domains on large scaffoldin proteins that hold the enzymes together, likely increasing substrate channeling between the synergistically acting enzymes (Fig. 1)[7]. Notably, the upregulated genes encode free (dockerin and CBM-lacking) CAZymes as well as dockerin- or CBM-fused CAZymes and scaffoldins (Fig. 1). 75% of the upregulated CAZymes were dockerin-fused, CBM-fused, or both. Additionally, 7 scaffoldins were upregulated. These results strongly suggest that transcription of fungal cellulosomes increases when *A. robustus* is cultivated in the presence of *M. bryantii*. The maximum log2fold change of all upregulated CAZymes in co-culture was 4.1 for a predicted carbohydrate esterase family 1 protein, and the median and average log2fold changes were 1.5 and 1.7, respectively. The upregulated enzymes play central roles in gut fungal carbohydrate metabolism, as evident in Table 3.1. These results support our hypothesis that co-cultivation of *A. robustus* with *M. bryantii* enhances the biomass-degrading capability of *A. robustus* by upregulating CAZyme expression. Combined with previous evidence in the literature[120,200–203,210], these results suggest more generally that other synthetic pairings of anaerobic gut fungi with methanogens improve the effectiveness of biomass breakdown by gut fungi via the same mechanism.

Table 3.1: Top 20 positively regulated genes in *A. robustus* when co-cultured with *M. bryantii* for Carbohydrate Metabolism and Metabolism of Complex Carbohydrates KEGG pathway classes

Protein ID	Log2fold change	KEGG pathway
327759	3.26	N-Glycan degradation
233233	2.39	Pentose and glucuronate interconversions
105274	2.08	N-Glycan degradation

283457	2.06	Starch and sucrose metabolism
273262	2.05	N-Glycan degradation
293755	2.04	N-Glycans biosynthesis
326455	1.73	Aminosugars metabolism
251333	1.72	N-Glycan degradation
292627	1.67	Starch and sucrose metabolism
230297	1.62	Starch and sucrose metabolism
271403	1.56	Starch and sucrose metabolism
28972	1.53	Starch and sucrose metabolism
328568	1.52	N-Glycan degradation
327760	1.48	N-Glycan degradation
328416	1.45	Starch and sucrose metabolism
327642	1.43	Galactose metabolism
295030	1.31	N-Glycans biosynthesis
249043	1.30	Starch and sucrose metabolism
271905	1.29	N-Glycan degradation
286401	1.21	Pentose and glucuronate interconversions

Although the majority of the upregulated putative CAZymes belong to known families (e.g. glycosyl hydrolases), 33% are dockerin- and/or CBM-fused but have other or unknown functions (Figure 3.2). Surprisingly, we found that four genes encoded dockerin- and/or CBM-fused putative proteases and one gene encoded a dockerin-fused histone acetyltransferase. The remaining proteins may represent uncharacterized biomass degrading enzymes. A previous transcriptomic study[211] of catabolite repression in gut fungi identified five *A. robustus* genes encoding proteins of unknown function that were co-regulated with CAZymes, but these five genes were not found to be upregulated in co-culture. These results highlight the value of co-cultivation of anaerobic gut fungi with methanogens as a means to identify enzymes with potentially novel biomass-degrading activities for future characterization.

We also investigated the role of the remaining positively regulated genes beyond CAZymes. We found that 63% of all positively regulated genes were annotated with a euKaryotic Orthologous Group, or KOG [212], class, excluding general function prediction or

unknown function classes. Excluding CAZymes, the top five represented KOG classes in the positively regulated genes in co-culture were: (1) Signal transduction mechanisms, (2) Replication, recombination and repair, (3) Transcription, (4) Carbohydrate transport and metabolism, and (5) Post-translational modification, protein turnover, chaperones. These KOG classes suggest that *A. robustus* senses co-cultivation with *M. bryantii* and responds by transcriptional and post-transcriptional regulation of carbohydrate metabolism.

The fourth most highly represented KOG class, carbohydrate transport and metabolism, implies that *A. robustus* grown in co-culture with *M. bryantii* increases the expression of sugar transporters and the rate of sugar utilization. This implication is consistent with a previous study by Joblin and colleagues[203], in which synthetic consortia of anaerobic gut fungi with *Methanobrevibacter smithii* removed sugar from lucerne stem faster than the fungal monocultures. The most common transporter classes for small molecules in microbes are the Major Facilitator Superfamily (MFS) and the ATP Binding Cassette (ABC) transporters[213]. Our analyses suggest that both classes of transporters are upregulated in the co-culture. However, although ABC transporters are involved in sugar uptake in prokaryotes, they are typically associated with extrusion processes in eukaryotes and may be involved in signaling processes[213,214]. On the other hand, we also identify two MFS transporters that show strong homology to sugar transporters in other microorganisms. MFS transporters have been implicated in the regulation of cellulase genes in *Trichoderma reesei*[215]. The upregulation of these putative sugar transporter genes may be in response to the increased availability of sugar from enhanced plant biomass breakdown. Furthermore, a closer inspection of the InterProScan annotations suggests that at least two G-protein coupled receptors (GPCRs) are upregulated in the co-culture. These fungal GPCRs appear to belong to Class 3, but they have an unusual architecture with a very large amino-terminal that is

different from the amino-terminal domain that is typically associated with metazoan Class 3 GPCRs [216]. Although experimental evidence is lacking, it is tempting to speculate that these GPCRs are involved in sensing and regulating sugars and biomass-degrading genes.

As demonstrated by the differential regulation of 785 genes at mid-log, *A. robustus* responds rapidly to co-cultivation with *M. bryantii*. However, the difference in the concentration of fermentation products at mid-log between co-culture and monoculture supernatant was not statistically significant (p-adjusted < 0.05), with the exception of lactate, which was not detectable in the monoculture (Figure 3.4). The transitory accumulation of lactate in co-culture relative to monoculture at mid-log has been reported previously[204]. These results suggest that the metabolite profile of the co-culture is not fully distinct from the monoculture until stationary phase. This work demonstrates the value of transcriptomics in addition to metabolite measurements to understand co-culture dynamics.

3.3   *Conclusions*

Here, we demonstrated that the anaerobic gut fungus *Anaeromyces robustus* upregulated diverse families of CAZymes when co-cultivated with *Methanobacterium bryantii.* In addition, we found two putative sugar transporters and two Class 3 GPCRs were also upregulated, suggesting that co-culture with methanogens drives carbohydrate sensing and import. Anaerobic gut fungi are attractive organisms for biomass breakdown due to their number and wealth of CAZymes, which is superior to all other sequenced fungi to date. The upregulation of CAZymes in a synthetic co-culture with a methanogen is a proof-of-principle that future efforts to engineer anaerobic gut fungi as platforms for biomass breakdown would benefit from including a methanogenic partner in order to leverage the synergy between these two organisms.

3.4 Materials and Methods

3.4.1 Routine cultivation of A. robustus and M. bryantii

The anaerobic fungal strain *A. robustus* was isolated via reed canary grass enrichment from fecal pellets collected from a sheep at the Santa Barbara Zoo, as described previously[6,7]. The fungal isolate was grown anaerobically at 39°C in Hungate tubes filled with 9 mL of autoclaved modified Medium C [174] ("MC-"), containing 1.25 g yeast extract, 5 g Bacto™ Casitone, and 7.5 vol% clarified rumen fluid, with 0.1 g of milled reed canary grass as growth substrate, and supplemented with vitamin solution post-autoclaving[210]. Fungal zoospores were transferred every 3-4 days into fresh modified Medium C containing 0.1 g of milled reed canary grass. *A. robustus* was treated monthly with 0.1 mL of 10,000 U/mL Penicillin Streptomycin (Gibco™) per 10 mL culture. *M. bryantii* (DSM No.-863, DSMZ) was cultivated in a variant of the M2 media recipe [210] containing 2 g/L sodium acetate, 4 g/L sodium formate, 2 g/L yeast extract, and 4 g/L Bacto™ Casitone. Liquid medium was dispensed by 9.0 mL aliquots into Hungate tubes anaerobically with an 80% H_2 and 20% CO_2 gas mix in the headspace. After autoclaving, 0.1 mL of sterile-filtered 100x vitamin solution[210] supplemented with 0.9 µg/mL Cyanocobalamin and 40 mg/mL CoM was added to each tube.. 1.0 mL of growing *M. bryantii* culture was then inoculated into each tube. Each month 1.0 mL of culture was transferred into fresh M2 medium. Routine growth was measured by accumulated pressure production[217] for anaerobic fungi, and by methane production for the methanogen.

3.4.2 Cocultivation of anaerobic gut fungi with M. bryantii

Serum bottles containing 40 mL of modified Medium C ("MC-") with 0.4 mL 100x vitamin solution and 0.4 g reed canary grass were inoculated with the following treatments from routine cultures of *A. robustus* (4 days of growth) and *M. bryantii* (6 days of growth): 1.0 mL of *A. robustus* or a combination of 1.0 mL of *A. robustus* and 1.0 mL of *M. bryantii*. A schematic of the preparation of co-cultures is depicted in Figure 3.1. Growth of the fungus was ensured by measurement of accumulated pressure, as described previously [217]. Growth of *M. bryantii* in isolation does not accumulate pressure (Figure 3.4) because the molar balance of methanogenesis produces a deficit of two moles of gas compared to the reactants[218,219]. Therefore, headspace samples of the co-culture were taken to ensure hydrogen consumption and methane production by *M. bryantii* as measured by gas chromatography (GC). After 4 days of growth at 39°C, 1.0 mL for each treatment was separately used to inoculate Hungate tubes (at least two replicates per treatment) containing 9 mL of modified Medium C ("MC-") supplemented with 0.1 mL of 100x vitamin solution and 0.1 g reed canary grass as growth substrate. Growth of the fungus was quantified by measurement of accumulated pressure, and methane production by *M. bryantii* in co-culture was measured by a TRACE™ 1300 Gas Chromatograph (Thermo Scientific). Both fungal monocultures and fungal-methanogen co-cultures were harvested at mid-log, as determined by accumulated pressure (Figure 3.5) and transferred into 15 mL Falcon tubes. Cells and plant biomass were pelleted by centrifugation at 3220 *g* using a swinging bucket rotor (Eppendorf™ A-4-81). Supernatant was removed by pipette and 1.0 mL RNAlater (Sigma-Aldrich®) was added to each sample. The samples were vortexed and frozen at -80°C until RNA extraction.
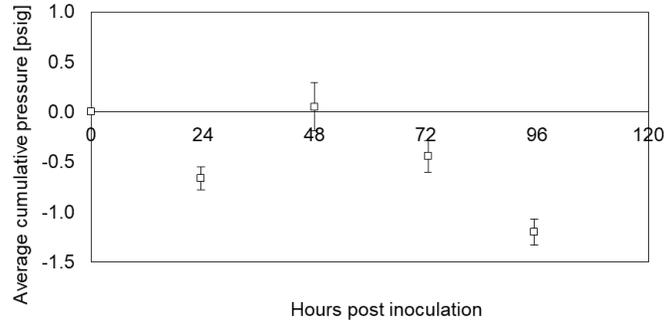
Figure 3.4: *M. bryantii* does not accumulate pressure in monoculture. Each data point is an average of three replicates.

● *A. robustus*   ○ *A. robustus & M. bryantii*
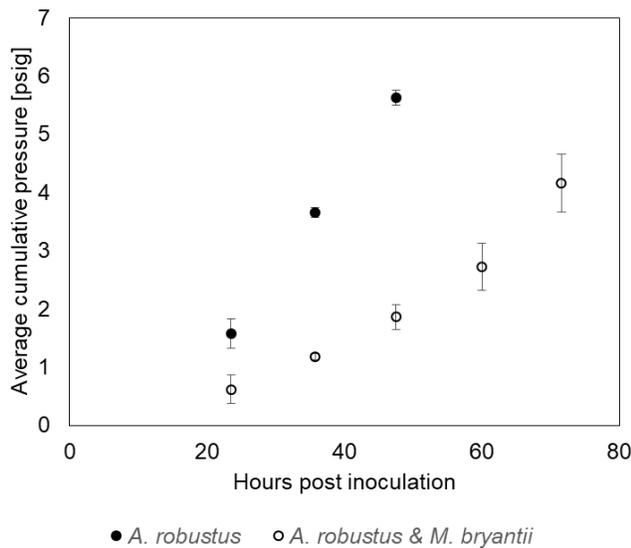
Figure 3.5: Growth of *A. robustus* monoculture and *A. robustus* with *M. bryantii* co-culture as measured by accumulated fermentation gas pressure. Each data point is an average of five replicates.

3.4.3   Extraction of RNA from A. robustus and A. robustus/M. bryantii co-cultures

Samples were thawed on ice and centrifuged at 3220 *g* for 10 min. RNAlater was removed by pipette and each pellet was transferred into a 2-ml screw-cap tube containing 1.0 mL of 0.5 mm zirconia/silica beads (Biospec). Tubes and beads were autoclaved before use. The cells were lysed using the Biospec Mini-Beadbeater-16 for 1 min. The tubes were centrifuged at 13,000 *g* for 3 min using a microcentrifuge (Eppendorf™ 5424). The supernatant was collected, using gel-loading tips (Fisher Scientific) to maximize yield. Subsequently, total RNA was extracted using the RNeasy Mini kit (QIAGEN). Instructions were followed for "Purification of Total RNA from Plant Cells and

Tissues and Filamentous Fungi" including the optional on-column DNAse digest. RNA quality was assessed by Agilent TapeStation. The RNA Integrity Number (RINe) of all samples prior to sequencing was above 7.0.

3.4.4 Library preparation, sequencing, and data analysis

Fungal mRNA libraries for co-culture and monoculture samples were prepared using the Illumina® Truseq® Stranded mRNA kit, which selects for eukaryotic polyadenylated mRNA. The libraries were sequenced by 75 bp single-end reads on an Illumina® NextSeq500 with a high output kit (more than 400 million reads). HISAT[220] was used to align reads to the *Anaeromyces robustus* genome, which is available for download at the Joint Genome Institute MycoCosm portal[7,23]. Only single, distinct primary alignments were allowed for reads. FeatureCounts[221] was used to extract counts per gene for each treatment. Genes were tested for differential regulation between co-culture (two replicates) and fungal monoculture (three replicates) using the DESeq2 package in R [222]. Functional classifications for differentially regulated genes (absolute log2fold change greater than 1, *p*-adjusted less than 0.05) were assigned from the KEGG[191], GO[223], KOG[212], and CAZyme annotations available in the MycoCosm portal[23]. CAZyme polysaccharide activities were determined from the Carbohydrate Active Enzymes database (http://www.cazy.org/)[224]. CAZymes were determined to be dockerin- or Carbohydrate Binding Module (CBM)-fused based on sequence information detailed in previous work [7].Using InterProScan annotations[181], the dataset was filtered for putative transporters and G-protein coupled receptors (GPCRs). TMHMM[225] was used to verify that the down-selected sequences contain predicted transmembrane segments and BLAST[226] was used to identify homologous sequences.

3.4.5   Detection of fatty acids by High Pressure Liquid Chromatography (HPLC)

An Agilent1260 Infinity HPLC (Agilent) was used to measure volatile fatty acids (VFAs) in co-cultures and monocultures. Samples were acidified to 5 mM sulfuric acid and let incubate at room temperature for 5 min. Following the incubation, samples were centrifuged at 21,000 g for 5 min and the supernatant was passed through a 0.22 μm PVDF filter. The samples were then injected into the HPLC by the autosampler. 5 mM sulfuric acid was used as the carrier through an HPLC Organic Acid Catalysis column (Aminex® 87H Ion Exclusion Column). Acetate, lactate, and formate were measured by the Variable Wave Detector (VWD). Ethanol was measured by the Refractive Index detector. Standards were prepared for each VFA except ethanol with concentrations of 0.1 and 1.7 g/L. Ethanol standards were 0.1 and 8 g/L.

4 Co-culture of anaerobic fungi with rumen bacteria establishes an antagonistic relationship

4.1 Introduction

Microbial antagonism can take many forms: antibiosis (the production by an organism of a compound that inhibits or kills another organism), competition for nutrients and space, parasitism, and others[227]. Although often discussed in the context of biological control agents that protest postharvest crops[227–229], microbial antagonism has also been recognized to have a profound impact on microbial communities, especially host-associated communities[230]. For example, microbial antagonism can increase microbial diversity in some situations[231,232], protect against invasion by pathogens[233], and facilitate genome evolution through the acquisition of genetic material from killed cells[234]. Mathematical modeling suggests that communities dominated by antagonistic relationships are more stable and resilient to perturbations than those dominated by cooperative relationships[235].

Microbial relationships, especially antagonistic ones, within the rumen microbiome are complex and not well-characterized. In particular, knowledge of rumen fungi (class Neocallimastigomycetes) and their interactions with other microbial community members is lacking. Rumen fungi, also called anaerobic gut fungi, thrive in the digestive tracts of large herbivores as part of a biomass-degrading consortium with bacteria, methanogenic archaea, and protozoa[5,236]. Bacteria outnumber fungi in the rumen by at least four orders of magnitude[5,236]. Co-cultivation of fungi with bacteria suggests that the nature of the interaction between rumen fungi and bacteria depends on the specific fungal-bacterial pairing. Antagonistic relationships, in which the cellulolytic activity of the fungus was

inhibited, were observed between *R. flavefaciens* and *Neocallimastix frontalis* MCH3 or *Piromyces communis* FL[121], *Piromyces communis* and *Selenomonas ruminantium*[122], as well as *R. flavefaciens* and *Orpinomyces joyonii* or *N. frontaslis*[123]. Some previous studies of rumen fungi co-cultivated with the cellulolytic rumen bacteria *Fibrobacter succinogenes* have shown no effect on biomass degradation[121,123,124], implying neither mutualism nor antagonism between these organisms. However, Joblin and colleagues found that *F. succinogenes* inhibited the degradation of ryegrass stems by *N. frontalis* in co-culture with *Methanobrevibacter smithii*, whereas the presence of *F. succinogenes* enhanced degradation by co-cultures of *Caecomyces* spp. with *M. smithii*[125]. In a separate study by Roger and colleagues[123], the presence of *F. succinogenes* had no impact on the degradation of wheat straw or maize stem by *N. frontalis* or *Orpinomyces* (*Neocallimastix*) *Joyonii*.

Transcriptomics allows us to probe more deeply the interactions between rumen fungi and bacteria. Co-culture transcriptomics has proven to be a valuable tool by which to investigate the nature of microbial interactions, as demonstrated in recent publications[237–240]. For example, RNA-sequencing of the anaerobic fungus *Anaeromyces robustus* in co-culture with methanogen *Methanobacterium bryantii* revealed that the fungus upregulated 105 genes encoding carbohydrate-active enzymes (CAZymes), representing 12% of total predicted CAZymes[238]. However, RNA-sequencing of both fungi and bacteria in co-culture remains rare due to the technical challenge of depleting ribosomal RNA from both microbes. Here, we co-cultivated pairings of rumen fungi with *Fibrobacter* sp. UWB7 (hereafter *F.* sp. UWB7), a close relative of *F. succinogenes*[2], and performed the first dual transcriptomic characterization of a rumen bacteria and fungus in co-culture. By this approach, we tested the hypothesis that the relationship between the *F.* sp. UWB7 and anaerobic gut fungi is antagonistic. Furthermore, we performed untargeted nonpolar metabolomics to determine

whether the introduction of *F*. sp. UWB7 triggers the production of possible defense compounds by the fungus. Specifically, we cultured *Anaeromyces robustus* with *F*. sp. UWB7 on crystalline cellulose (Avicel® PH-101, Sigma) or switchgrass as well as *Caecomyces churrovis* with *F*. sp. UWB7 on switchgrass and compared to the respective fungal and bacterial monocultures.

4.2   Results and Discussion

Secondary metabolites, although not strictly necessary for the growth or survival of an organism under all growth conditions[9], are often secreted during antagonistic relationships between microorganisms[37,115,143]. Previous work mining the high quality genomes of anaerobic fungi revealed that anaerobic fungi are capable of synthesizing secondary metabolites (Chapter two). *A. robustus* and *C. churrovis* encoded 43 and 32 biosynthetic enzymes for various classes of secondary metabolites, including nonribosomal peptide synthetases (NRPSs) and polyketide synthases (PKSs). We hypothesized that some of the secondary metabolites produced by *A. robustus* and *C. churrovis* are compounds used for defense or competition against rumen bacteria.

4.2.1   Co-cultivation with rumen bacteria induces stress in anaerobic fungi and activates components of fungal secondary metabolism

Although not stable for many generations of batch passaging, anaerobic fungi can grow with *F*. sp. UWB7 for a sufficient duration to capture RNA from both organisms. *A. robustus* or *C. churrovis* were grown in isolation for 24 hours before the introduction of *F*. sp. UWB7 (Figures 4.1 and 4.2). Fungi and bacteria were grown together until the fungus reached mid-log phase, at which point the co-cultures and corresponding fungal

monocultures were harvested for RNA extraction and sequencing. In response to the presence of *F*. sp. UWB7 (Figure 4.3), cultures of *Anaeromyces robustus* or *Caecomyces churrovis* both upregulated genes encoding stress response proteins (chaperones), indicating that the presence of the bacteria invoked a fungal stress response (Figure 4.4). The fungal stress was observed during growth on switchgrass or Avicel® PH-101. However, roughly ten times more genes were differentially regulated comparing fungal co-culture to monoculture during growth on Avicel® relative to growth on switchgrass (Table 4.1). This difference may reflect the preference of *F*. sp. UWB7, a cellulose-degrading specialist[241], for crystalline cellulose over complex plant matter, resulting in both more robust bacterial growth and a stronger bacterial signal to the fungi.
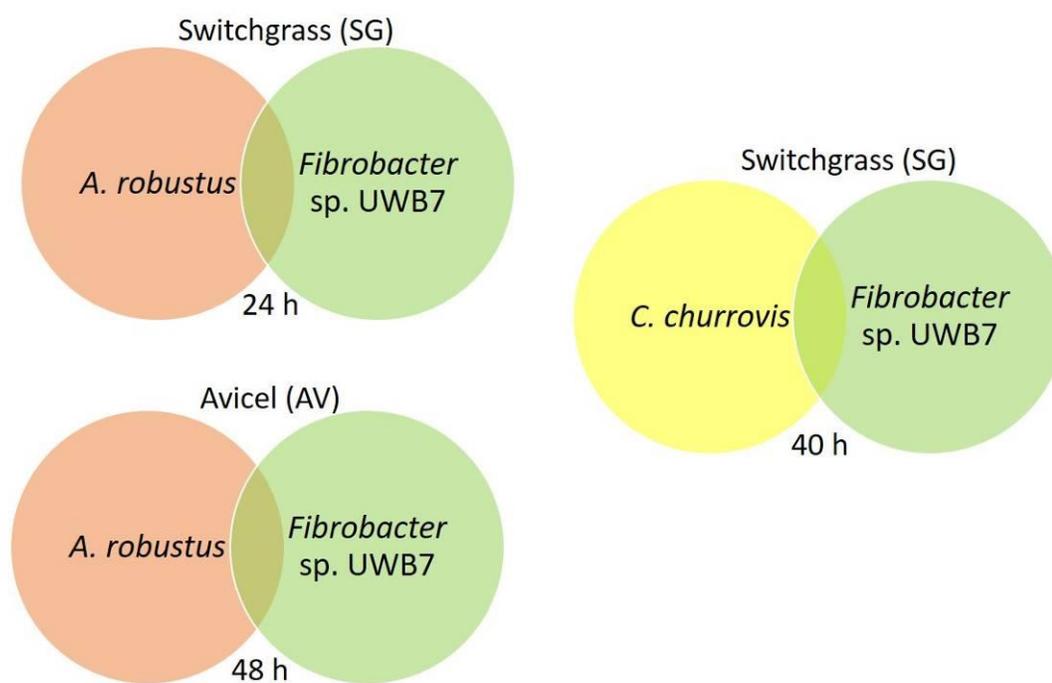
Figure 4.1: Schematic of the co-cultivation pairings of anaerobic fungi with the rumen bacteria *F*. sp. UWB7 on two different carbon substrates. Fungal cultures were grown for 24 h prior to the introduction of *F*. sp. UWB7 and subsequently co-cultured for the duration listed below each Venn diagram.
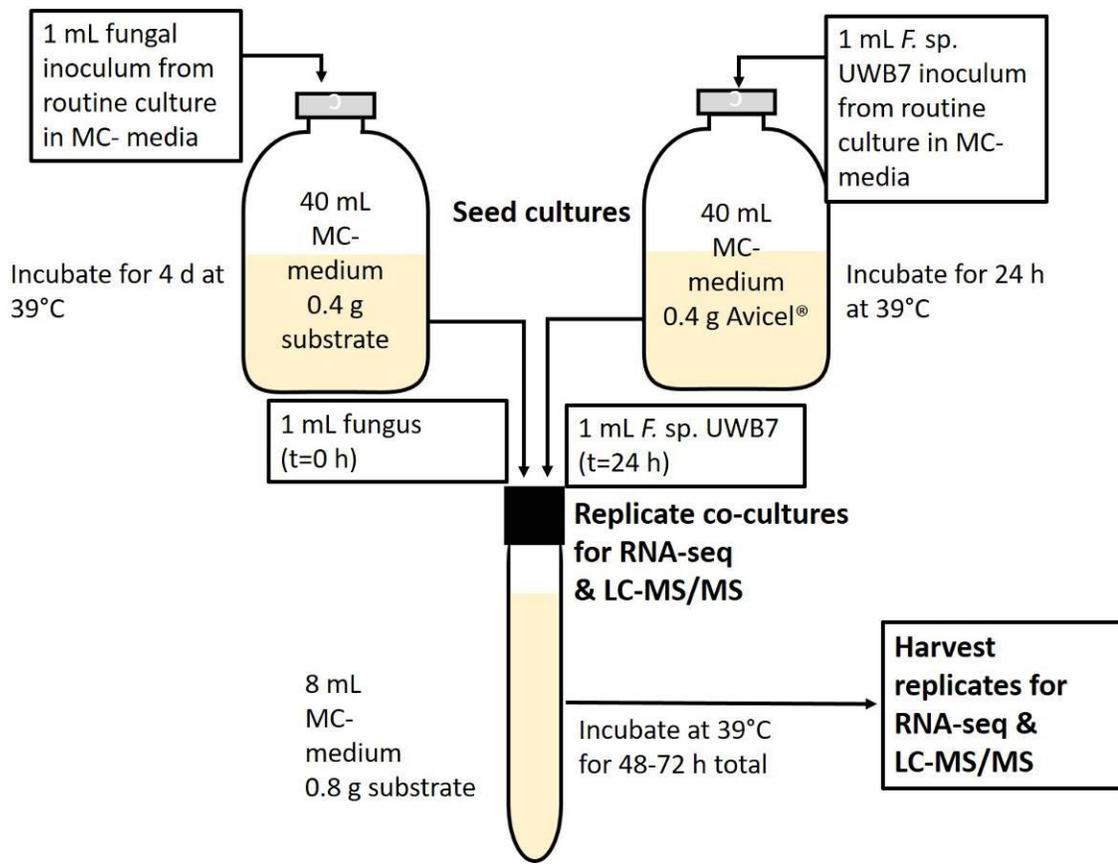
Figure 4.2: Detailed experimental schematic of the preparation of cultures for transcriptomics and metabolomics.
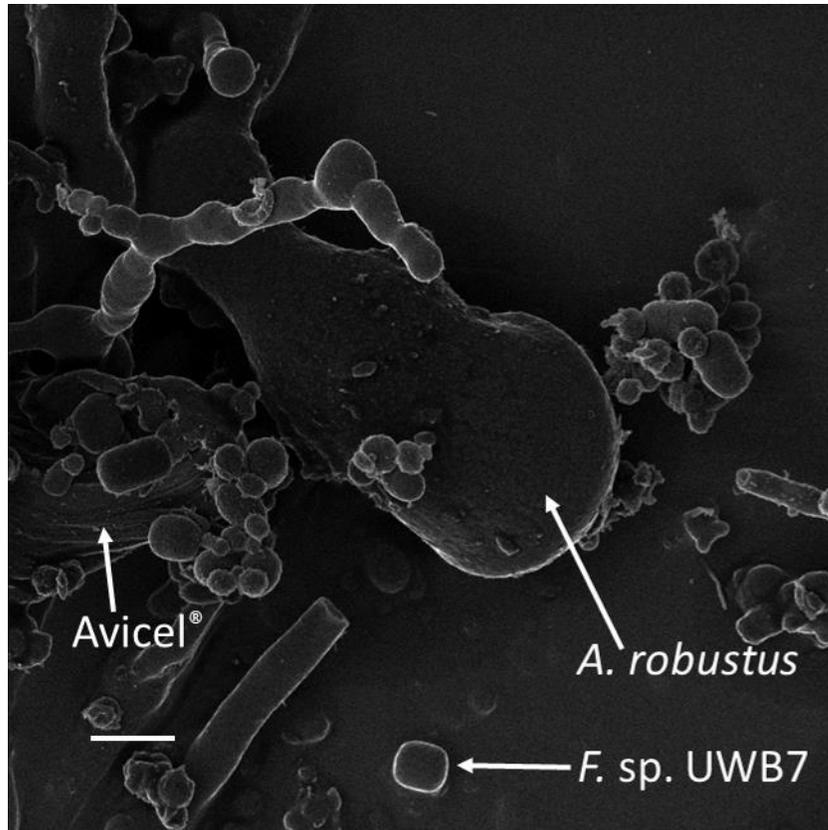
Figure 4.3: Helium ion micrograph of *A. robustus* grown in co-culture with *F*. sp. UWB7 on Avicel®. The presence of *A. robustus* is indicated by a sporangium, and the presence of *F*. sp. UWB7 indicated by single cells. Field of view 20.00 µm. Scale bar 2.00 µm.
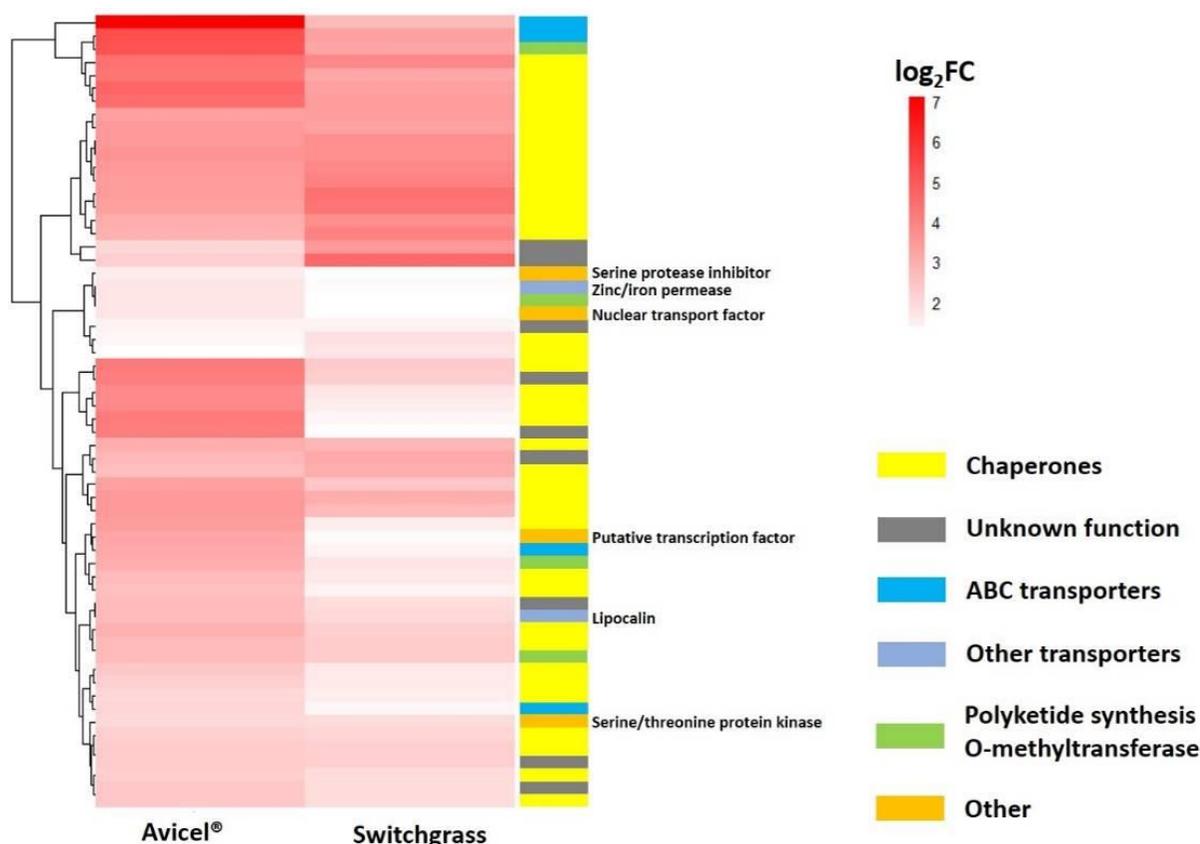
Figure 4.4: *A. robustus* **upregulates transporters, chaperones, and O-methyltransferases in response to co-cultivation with *F*. sp. UWB7 on different carbon substrates**. Heatmap representing the $\log_2$ fold change of *A. robustus* transcript abundance in co-culture with *F*. sp. UWB7 relative to respective fungal monoculture at mid-log on two different carbon substrates (Avicel® or switchgrass). Only fungal transcripts upregulated at least two-fold with *p*-adjusted < 0.05 in co-culture versus monoculture are shown. Putative functions are designated as assigned in MycoCosm[23] (KOG[212] or InterPro[181]).

Table 4.1: Number of differentially regulated fungal genes in co-culture with *F*. sp. UWB7 compared to fungal monoculture (absolute $\log_2$fold change<1 and adjusted *p*-value<0.05). AV=Avicel® and SG=switchgrass.

Fungus, substrate	Total differentially regulated genes	Upregulated in co-culture	Downregulated in co-culture
A. robustus, Avicel®	2937	1022	1915
A. robustus, switchgrass	237	135	102
C. churrovis (SG)	1151	579	333

Chapter two has demonstrated that the genomes of anaerobic gut fungi encode a variety of biosynthetic enzymes of natural products. In addition to a general stress response, *A.*

robustus upregulated six non-ribosomal peptide synthetases (NRPSs, and one polyketide synthase (PKS)-like enzyme (Table 4.2) at least two-fold (*p*-adjusted <0.05) in co-culture with *F*. sp. UWB7 compared to *A. robustus* monoculture (both cultures grown on Avicel®). Neighboring genes were co-regulated with a pair of NRPS genes (Figure 4.5). One NRPS gene cluster was downregulated in the Avicel® co-culture (Figure 4.5), which suggests that the nonribosomal peptide not be a defense compound. Instead, the nonribosomal peptide may serve another, at present unknown, function. In higher-order fungi, secondary metabolites are linked to different developmental stages of fungi[84–86] and some of the secondary metabolites from anaerobic fungi may also serve such a purpose. In addition, five genes encoding putative polyketide O-methyltransferases were upregulated with a $\log_2$fold change of 2.7 or greater (Table 4.3). Surprisingly, a predicted protein (271870) containing a condensation domain), which would normally form a modular domain on an NRPS[242], was upregulated two-fold in co-culture relative to monoculture. Taken together, these data suggest that *A. robustus* regulates polyketide and non-ribosomal peptide synthesis in response to microbial challenge by *F*. sp. UWB7.

Table 4.2: Differentially regulated (adjusted *p*-value < 0.05) biosynthetic genes for secondary metabolites. Genes marked with an asterisk in the proteinId column indicate the gene is co-regulated with neighboring genes (Figure 4.2).

MycoCosm proteinId	SM type	Log2fold change	Scaffold
A. robustus (Avicel®)			
193122	NRPS	2.9	480
294553	NRPS	2.9	182
271076	NRPS	2	279
231391*	NRPS	1.8	77
266215	PKS-like	1.6	49
218823*	NRPS	1.2	77
330657	NRPS	1.1	540
328517*	PKS	(-) 1.0	207
C. churrovis (switchgrass)			
17094	PKS	(-)	3.6
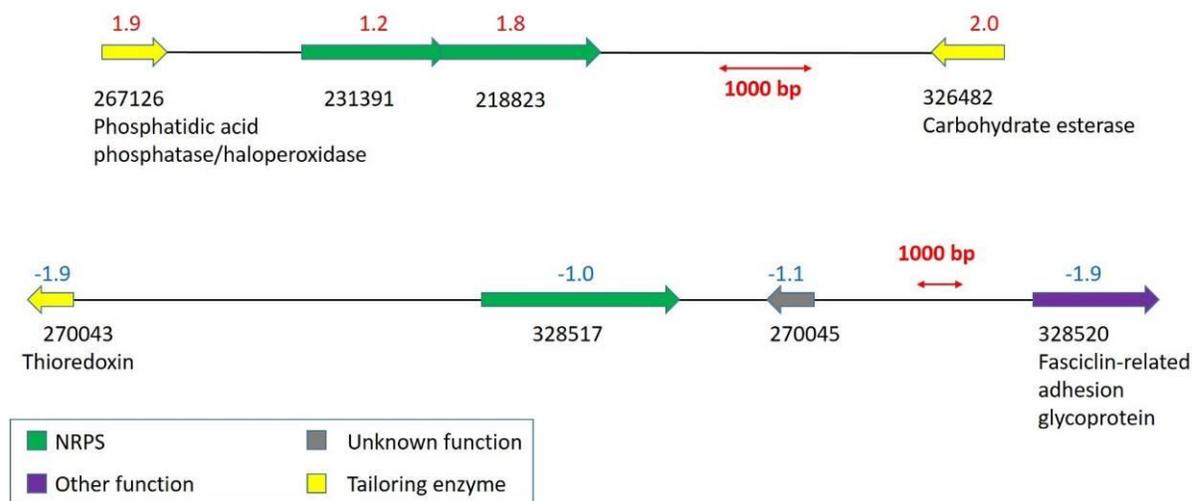
Figure 4.5: Genes that were coordinately regulated in co-culture of *A. robustus* with *F.* sp. UWB7 versus fungal monocultures. All cultures were grown on Avicel®. Log$_2$fold change is shown above each gene and the MycoCosm proteinId is shown below each gene.

Table 4.3: Anaerobic gut fungi *A. robustus* and *C. churrovis* upregulate genes involved in polyketide and non-ribosomal peptide synthesis as well as SAM-dependent methyltransferases potentially involved in the regulation of secondary metabolism when co-cultured with *F.* sp. UWB7. Adjusted *p*-value less than 0.05 for all comparisons.

MycoCosm protein Id	InterPro Id	Log2fold change	Product name or InterPro annotation
A. robustus (**Avicel®**)			
298422	IPR016874	5.1	Polyketide synthesis O-methyltransferase
325735	IPR003455,IPR016874	4.1	Polyketide synthesis O-methyltransferase
325736	IPR003455	3.2	O-methyltransferase domain protein
325734	IPR003455, IPR16874	3.0	Polyketide synthesis O-methyltransferase
266064	IPR003455, IPR16874	2.7	Polyketide synthesis O-methyltransferase
297314	IPR013217	1.4	S-adenosyl-L-methionine-dependent methyltransferase
271870	IPR001242	1.2	Condensation domain
C. churrovis (**switchgrass**)			
525731	IPR007213, IPR016874, IPR029063	4.1	S-adenosyl-L-methionine-dependent methyltransferase
619287	IPR007213, IPR029063	3.8	Leucine carboxyl methyltransferase-domain containing protein
581095	IPR007213, IPR029063	3.5	S-adenosyl-L-methionine-dependent methyltransferase
270385	IPR007213, IPR016874, IPR029063	3.4	S-adenosyl-L-methionine-dependent methyltransferase
175508	IPR007213, IPR016874, IPR029063	3.0	Putative tetracenomycin polyketide synthesis O-methyltransferase
462789	IPR007213, IPR016874, IPR029063	2.5	S-adenosyl-L-methionine-dependent methyltransferase*

4.2.2 *Anaerobic fungi regulate their secondary metabolism via epigenetic modifications in the presence of rumen bacteria*

Both *A. robustus* and *C. churrovis* upregulated genes encoding S-adenosyl methionine (SAM)-dependent methyltransferases when co-cultured with *F.* sp. UWB7 (Table 4.3). One of these protein may perform a similar function LaeA or Lae1, which act in complex with other proteins as global regulators of secondary metabolism in higher order fungi[87,88,243]. LaeA is reported to modulate gene expression of biosynthetic gene clusters (BGCs) via chromatin remodeling[97], and studies have found that epigenetic modifications such as histone acetylation or methylation can regulate expression of biosynthetic gene clusters in fungi[97,244,245]. We propose that anaerobic gut fungi remodel chromatin via histone modifications in order to modulate their secondary metabolism, similar to what has been suggested for higher-order fungi[97,98]. When co-cultured with *F.* sp. UWB7, both *C. churrovis* and *A. robustus* upregulated genes with putative functions in histone methylation or acetylation (Table 4.4), which are both modifications known to be involved in determining heterochromatin or euchromatin locations[93]. It is possible that one of the highly upregulated methyltransferases in co-culture acts as a global regulator of secondary metabolism in anaerobic gut fungi, similar to LaeA and Lae1. However, the distant evolutionary relationship between Neocallimastigomycetes and higher-order fungi as well as the current lack of genetic tractability of rumen fungi makes previous approaches used to pinpoint LaeA homologs unreliable or impossible at this time.

To test whether there were differences in the amount of histone modifications H3K4me3 and H3K27me3 between fungal monocultures and fungal co-cultures with *F* sp. UWB7, we performed Western blots on monoculture and co-culture cell lysates using

antibodies raised to *S. cerevisiae* H3K4me3 and H3K27me3 (Supplementary Figures 6.2.1 and 6.2.2). The exposure time for the H3 loading control increased by a factor of ten between monoculture and co-culture, indicating decreased fungal biomass in co-culture. However, the exposure times were nearly equivalent between H3K27me3 blots of monocultures and co-cultures, indicating that despite the decreased fungal biomass in co-culture, there was an increased proportion of H3K27me3 marks. H3K27me3 is known to be a repressor of transcription, whereas H3K4me3 is an activating mark[246]. Consistent with the downregulation of genes due to the enhancement of H3K27me3 marks in co-culture, more genes were downregulated than upregulated when comparing *A. robustus* co-cultures to monocultures during growth on Avicel® (Table 4.1). These results support that epigenetic modifications influence gene regulation when *A. robustus* is exposed to *F.* sp. UWB7. Furthermore, two genes encoding heterochromatin-associated protein HP1 were greater than two-fold downregulated in co-culture (protein Ids 290815, 266437) and another gene (280338)encoding a homolog of the WSTF-ISWI chromatin remodeling complex, which has been implicated in the replication of heterochromatin[247], was eight-fold downregulated in co-culture. In aspergilli, heterochromatin protein-1 and H3K9me3 marks have been associated with the repression of secondary metabolite gene clusters[97,98]. Taken together, these findings suggest that the secondary metabolism of anaerobic fungi is regulated via epigenetic marks and chromatin remodeling, consistent with higher-order fungi.

Table 4.4: Genes upregulated in co-culture with *F.* sp. UWB7 relative to fungal monoculture with putative functions in chromatin remodeling. Adjusted *p*-value less than 0.05.

MycoCosm proteinId	InterPro Id	Log2fold change	InterPro annotation
A. robustus (Avicel®)			
328372	IPR013216	5.4	Methyltransferase type 11
325551	IPR013216	2.0	Methyltransferase type 11
266199	IPR000182, IPR016181	2.0	Acyl-CoA N-acyltransferase
47572	IPR013216	1.9	Methyltransferase type 11
C. churrovis (switchgrass)			
619316	IPR013216, IPR025714, IPR029063	3.7	Methyltransferase type 11
451944	IPR013216, IPR029063	2,3	Methyltransferase type 11
627120	IPR000181, IPR000182	2.2	Acyl-CoA N-acyltransferase
479563	IPR013216, IPR029063	2.0	Methyltransferase type 11
522030	IPR000182	1.4	Acyl-CoA N-acyltransferase
78835	IPR00181, IPR000182	1.1	Acyl-CoA N-acyltransferase

To further understand the fungal regulatory response to co-cultivation with *F.* sp. UWB7, we analyzed the eukaryotic Orthologous Groups (KOGs)[212] of the differentially regulated transcripts (Supplementary Figures 6.2.1-6.2.3). In the case of *A. robustus* cultured with *F.* sp. UWB7 on Avicel®, the percentage of significantly downregulated fungal genes in co-culture was greater than upregulated for the majority of the KOG classes (absolute $\log_2$fold change ≥ 1, *p*-adjusted <0.05). In other words, co-cultivation repressed most fungal cellular and metabolic processes, with the exceptions of the KOG classes of post-translational modifications and secondary metabolism. Table 4.5 lists the upregulated genes in co-culture within the KOG class of secondary metabolism. In addition to the NRPS genes described above, 12 multidrug exporters belonging to the ABC superfamily were upregulated. The percentage of differentially regulated genes in all of the KOG classes was greater than 10%, with a median of 23%. In contrast, cultivation on switchgrass resulted in far fewer differentially regulated genes in each KOG class between co-culture and

monocultureA. *robustus* significantly regulated a median value of 1% of genes in each KOG

class, and *C. churrovis* regulated 5%. These findings further support that the cultivation of of

F. sp. UWB7 with *A. robustus* on Avicel® resulted in a stronger bacterial signal to the

fungus. More broadly, it is clear that the choice of substrate, in addition to the specific

organisms, has a profound impact on the gene regulation of microorganisms in co-cultures.

Table 4.5: Upregulated genes *A. robustus* (co-culture with *F.* sp. UWB7 relative to *A. robustus* monoculture) assigned to the KOG class "Secondary metabolites biosynthesis, transport and catabolism" of. All cultures were grown with Avicel® as the substrate. Only genes with $\log_2$fold change greater than one and adjusted *p*-value less than 0.05 are shown. Genes on the same scaffold (bold) are co-localized, immediate neighbors. The multidrug/pheromone exporter class is KOG0055 and the nonribosomal peptide synthetase KOG is KOG1178.

MycoCosm proteinId	Scaffold	log2FC	KOG name
266148	47	-7.1	Multidrug/pheromone exporter, ABC superfamily
275781	27	-5.2	Multidrug/pheromone exporter, ABC superfamily
327152	110	-3.7	Multidrug/pheromone exporter, ABC superfamily
330958	721	-3.2	Multidrug/pheromone exporter, ABC superfamily
329547	322	-3.1	Multidrug/pheromone exporter, ABC superfamily
305700	**422**	-3	Multidrug/pheromone exporter, ABC superfamily
297543	**434**	-2.1	Multidrug resistance-associated protein/mitoxantrone resistance protein, ABC superfamily
330191	**422**	-1.6	Multidrug/pheromone exporter, ABC superfamily
265866	40	-1.6	Multidrug/pheromone exporter, ABC superfamily
330237	**434**	-1.4	Multidrug resistance-associated protein/mitoxantrone resistance protein, ABC superfamily
330852	**626**	-1.4	Multidrug resistance-associated protein/mitoxantrone resistance protein, ABC superfamily
226999	**626**	-0.96	Multidrug resistance-associated protein/mitoxantrone resistance protein, ABC superfamily
193122	480	-2.9	Non-ribosomal peptide synthetase/alpha-aminoadipate reductase and related enzymes
294553	182	-2.9	Non-ribosomal peptide synthetase/alpha-aminoadipate reductase and related enzymes
271076	279	-2	Non-ribosomal peptide synthetase/alpha-aminoadipate reductase and related enzymes
231391	**77**	-1.8	Non-ribosomal peptide synthetase/alpha-aminoadipate reductase and related enzymes
218823	**77**	-1.2	Non-ribosomal peptide synthetase/alpha-aminoadipate reductase and related enzymes
330657	540	-1.1	Non-ribosomal peptide synthetase/alpha-aminoadipate reductase and related enzymes

4.2.3 Rumen bacteria upregulate genes encoding components of drug efflux pumps in response to the presence of anaerobic fungi

To further probe the relationship between *F.* sp. UWB7 and anaerobic gut fungi, we sequenced the corresponding prokaryotic messenger RNA in co-culture and monoculture, using both eukaryotic and prokaryotic ribosomal RNA depletion (see Methods). When co-cultivated with *C. churrovis* using switchgrass as the carbon source, *F.* sp. UWB7 upregulated 143 genes and downregulated 261 genes (4 and 8% of predicted genes in IMG/M[248]). Putative transporters comprised 12% of the upregulated genes. Table 4.6 summarizes upregulated genes encoding transporters (log2fold change≥1.0, adjusted *p*-value<0.05).Notably, a predicted TolC family protein (Locus tag Ga0136279_1901) was two-fold upregulated in co-culture relative to *F.* sp. UWB7 monoculture (*p*-adjusted 1.5×10^{-15}). TolC proteins are components of efflux pumps in Gram-negative bacteria and these pumps can transport a wide array of molecules, including antibiotics[249]. In addition, two genes encoding the adaptor subunits of RND efflux pumps (Ga0136279_1902 and Ga0136279_0657) were two- and three-fold upregulated in co-culture (*p*-adjusted$\leq 10^{-6}$), suggesting that they may be part of a regulon. Upon inspection of the gene neighborhoods, the TolC family protein Ga0136279_1901 and adapter subunit Ga0136279_1902 were neighboring genes. The co-regulation of genes encoding components of multi-drug efflux pumps has been previously reported: in Enterobacteriaceae, the genes encoding the multi-drug efflux pump AcrAB-TolC (*acrA*, *acrB*, and *tolC*) form a regulon[250], although *tolC* is not co-localized with *acrA* and *acrB*. The two-fold upregulated gene Ga0136279_2553, annotated as an ABC transporter substrate binding protein also bordered a TolC family protein (Ga0136279_2554), although the gene encoding the putative TolC protein was not

differentially regulated in co-culture. For effective efflux directly into the external environment, both the outer membrane channel, such as TolC, and the periplasmic adaptor are necessary[251]. Therefore, it is significant that both the TolC and adaptor protein homologs are upregulated when *F.* sp. UWB7 is co-cultivated with *C. churrovis*.

Besides drug efflux pumps, *F.* sp. UWB7 also upregulated at least two-fold six genes encoding chaperones, supporting the induction of a bacterial stress response by co-cultivation with anaerobic fungi. Notably, *F.* sp. UWB7 also upregulated 32-fold a putative HicB antitoxin (Ga0136279_0693), which could be part of a toxin-antitoxin system[252]. In addition, *F.* sp. UWB7 two-fold upregulated a gene encoding a putative abortive phage resistance protein (Ga0136279_0760), which are typically part of an RNA toxin-antitoxin system[253].

Table 4.6: Anaerobic gut fungi *A. robustus* and *C. churrovis* upregulate genes involved in polyketide and non-ribosomal peptide synthesis as well as SAM-dependent methyltransferases potentially involved in the regulation of secondary metabolism when co-cultured with *F.* sp. UWB7. Adjusted *p*-value less than 0.05 for all comparisons.

Locus Tag	Log2fold change	Product name
Ga0136279_2636	2.9	ABC transporter ATP-binding protein
Ga0136279_2635	2.2	Putative ABC transport system permease protein
Ga0136279_2405	2.2	type II and III secretion system protein
Ga0136279_1256	1.8	Outer membrane protein beta-barrel domain-containing protein
Ga0136279_0657	1.7	Urea ABC transporter substrate-binding protein
Ga0136279_2085	1.6	MULTISPECIES: efflux RND transporter periplasmic adaptor subunit
Ga0136279_1465	1.6	Zinc ABC transporter substrate-binding protein
Ga0136279_2620	1.5	MULTISPECIES: ammonium transporter
Ga0136279_2553	1.5	Transporter
Ga0136279_2080	1.4	MULTISPECIES: ABC transporter substrate-binding protein
Ga0136279_2080	1.4	Iron complex outermembrane recepter protein*
Ga0136279_1904	1.3	General secretion pathway protein E
Ga0136279_1405	1.2	TonB family C-terminal domain-containing protein*
Ga0136279_1902	1.1	MULTISPECIES: efflux RND transporter periplasmic adaptor subunit
Ga0136279_1405	1.0	TRAP transporter large permease subunit
Ga0136279_0818	1.0	Calcium/sodium antiporter
Ga0136279_1391	1.0	MULTISPECIES: TolC family protein

*Locus tags Ga0136279_2080 and Ga0136279_2080 are not transporters, but are part of the TonB receptor complex involved in iron transport[254–256].

4.2.4 The metabolic profile in co-culture is distinct from bacterial and fungal monocultures

To further test the hypothesis that co-cultivation of anaerobic gut fungi with *F.* sp. UWB7 triggers the production of fungal defense compounds, we performed untargeted nonpolar metabolomics analysis on fungal-bacterial co-cultures and the respective fungal or bacterial monocultures. We constructed a principal component analysis (PCA) plot using MetaboAnalyst[257] (Figure 4.6). In the two-dimensional score plot, bacterial monocultures and fungal-bacterial co-cultures grown on switchgrass show a high degree of overlap. However, the cultures grown on Avicel® (*A. robustus* monoculture, *F.* sp. UWB7 monoculture, and co-culture) were distinct from each other and this is even more readily apparent in the three-dimensional score plot (Supplementary Figure 8.2.6). Taken together, these data suggest that, in contrast to cultures grown on switchgrass, the metabolic profiles observed in the co-cultures of *A. robustus* with *F.* sp. UWB7 on Avicel® are distinct from those observed in the respective monocultures.
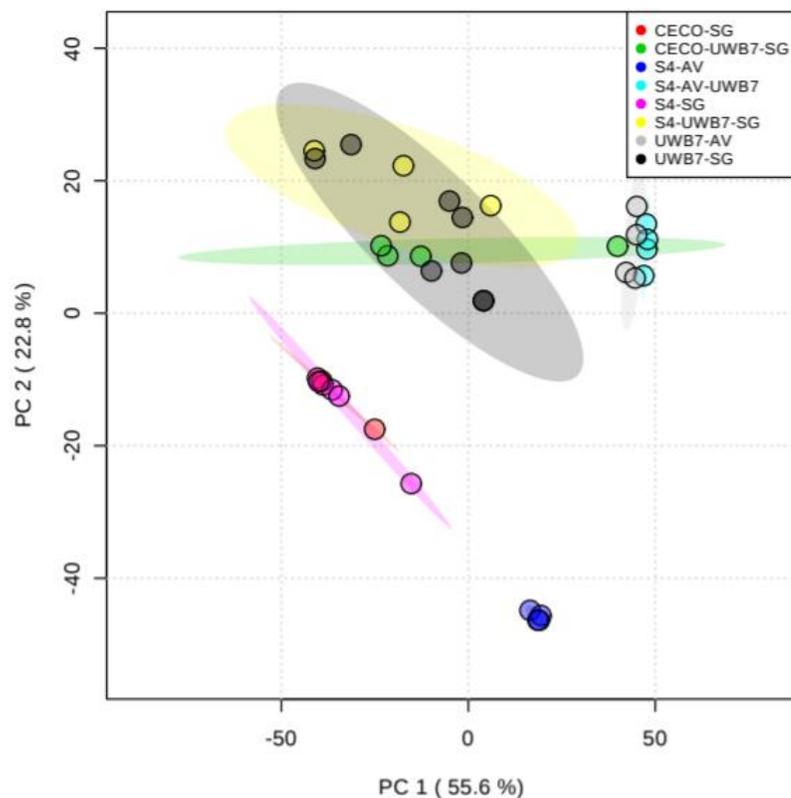
Figure 4.6: The metabolic profile of *A. robustus* co-cultured with *F.* sp. UWB7 on Avicel®️ is distinct from the respective fungal and bacterial monocultures. Two-dimensional principal component analysis (PCA) plots of the untargeted nonpolar metabolomics data for co-cultures and monocultures of A. robustus, C. churrovis, and F. sp. UWB7. S4=A. robustus, Ceco=C. churrovis, UWB7=F. sp. UWB7, AV=Avicel®️, SG=switchgrass. Plots were rendered by MetaboAnalyst[258].

We further investigated the distribution of nonpolar metabolites between monocultures and co-cultures by constructing molecular networks using Global Natural Products Social Molecular Networking (GNPS)[168] and visualizing the networks in Cytoscape[259] with three-way coloring[260] (Figure 4.7). This approach highlighted a group of unknown metabolites unique to *F.* sp. UWB7 that were not observed in fungal monocultures. These metabolites were not enriched by co-cultivation with anaerobic fungi and therefore may represent constitutively produced bacterial metabolites. The closest known node in the cluster matched 1-palmitoyl-sn-glycero-3-phosphoethanolamine (*m/z* 454.2791), a lysophospholipid (LPL), which comprises a small concentration of of Gram-

negative bacterial cell membranes[261]. Structure and molecular class predictions from SIRIUS[171] and CANOPUS[172] suggested that the unknown nodes may represent glycerophosphoethanolamines. Overall this cluster of nodes likely represents components of the bacterial membrane that are released into the supernatant after cell death and lysis.

Fatty acids & derivatives

Glycerophosphoethanolamines

Lysophospholipids

Peptides

Peptides
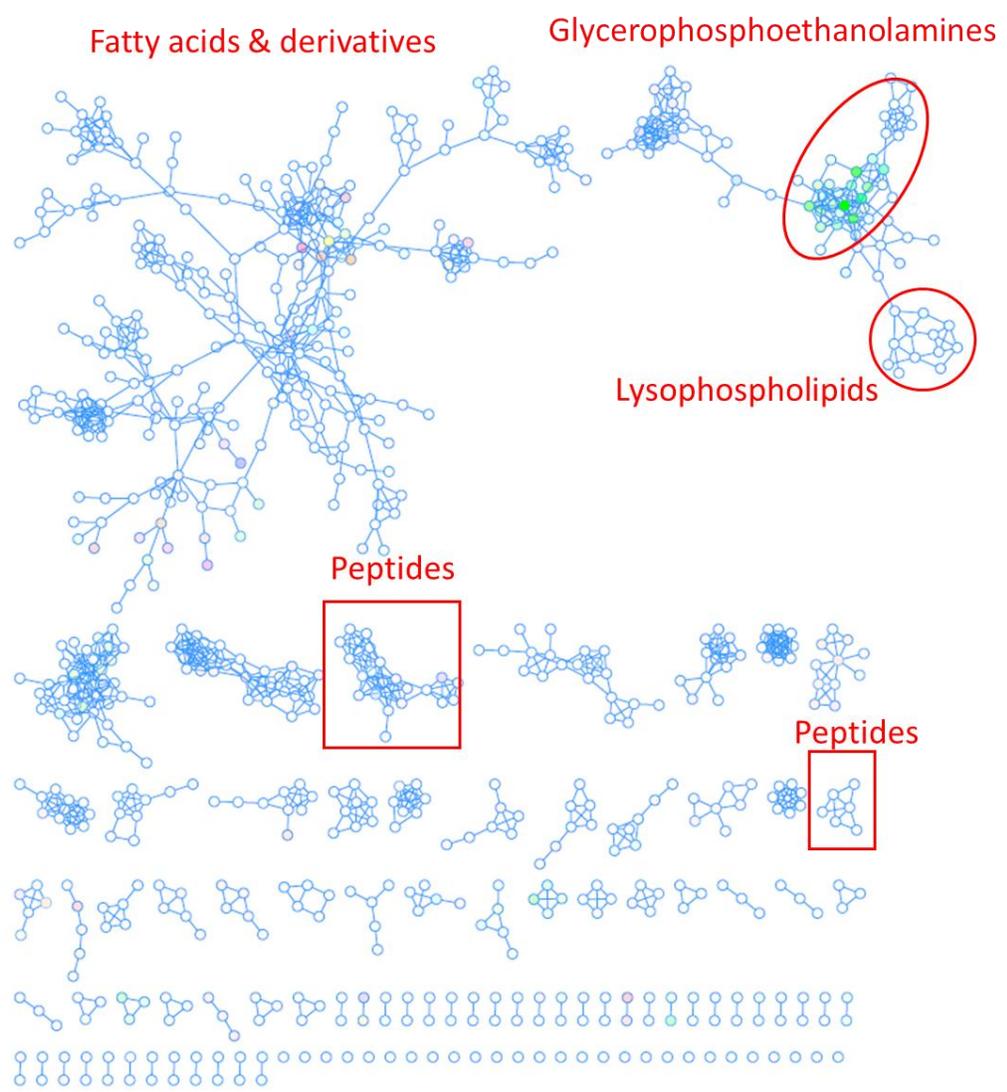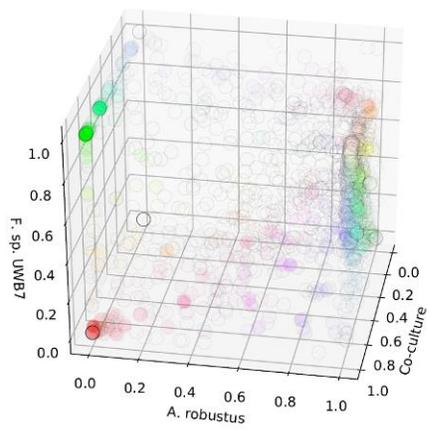
Figure 4.7: Cultivation of anaerobic fungi with *F*. sp. UWB7 reveals diverse shared metabolites as well as a group of unknown bacterial metabolites. Combined feature-based

molecular network created by GNPS[168] from positive and negative ion mode LC-MS/MS data from the nonpolar metabolites of *A. robustus*, *C. churrovis*, and *F.* sp. UWB7 co-cultures and monocultures grown on Avicel® or switchgrass substrates. Self-looping nodes were truncated. Three-way coloring[260] was used to visualize features in the *A. robustus* monocultures, *A. robustus-F.* sp. UWB7 co-cultures, and *F.* sp. UWB7 monocultures (all grown on Avicel®). Transparency of the nodes was set to emphasize nodes with high intensity in the co-culture of *A. robustus-F.* sp. UWB7.

By comparing the peak height ratios of features between culture conditions, we searched for metabolites that were enriched by co-cultivation with *F.* sp. UWB7. We screened these metabolites for those that were consistently observed across all four biological co-culture replicates. We applied an enrichment threshold of four-fold, reasoning that a four-fold enhancement would most likely not be caused by the simple addition of bacterial and fungal biomass in co-culture, but rather to increased production by one or both of the organisms. Four features matched these stringent criteria (Figure 4.8). All four features are single nodes in the molecular network depicted in Figure 4.7 and are thus not likely to be part of a family of structurally related compounds. Notably, the feature *m/z* 244.2272 was not observed in *F.* sp. UWB7 monoculture, but was 12-fold enriched in co-culture relative to *A. robustus* monoculture on Avicel®, which suggests that this is a unique fungal metabolite with enhanced production in response to the presence of *F.* sp. UWB7.

89

| Substrate | | | Avicel® | | Switchgrass | | | |
| | | | A. robustus | | A. robustus | | C. churrovis | |
Spectra	m/z	RT	Co-culture/UWB7 monoculture	Co-culture/fungal monoculture	Co-culture/UWB7 monoculture	Co-culture/fungal monoculture	Co-culture/UWB7 monoculture	Co-culture/fungal monoculture
NA	196.0403588	2.39	12.89	19.97				
A	244.227173	5.80	16.88	10.51	12.35	8.95		
NA	380.2768314	4.98	*	12.13				
B	408.3083649	7.43			4.52	13.20	4.98	4.96
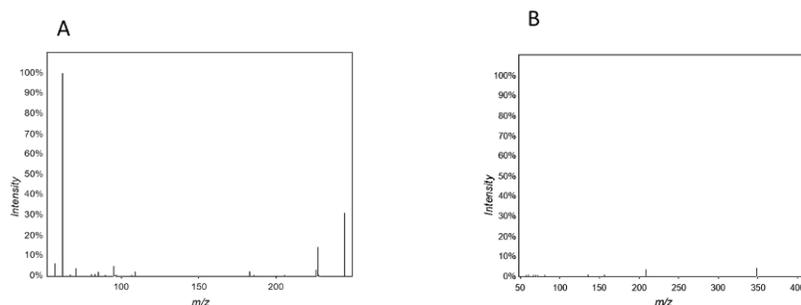
Figure 4.8: Metabolites enriched in fungal-bacterial co-cultures relative to monocultures. Untargeted nonpolar metabolomics features enriched at least four-fold in co-culture of *A. robustus* with *F.* sp. UWB7 on Avicel® compared to both bacterial and fungal monocultures (one-tailed student's *t*-statistic<0.05). No MS/MS was acquired for features marked "NA." All features were detected in positive ion mode. *Feature was not detectable in UWB7 monoculture.

4.3 Conclusions

We have demonstrated using dual transcriptomics that, despite previous somereports that *Fibrobacter succinogenes* had no effect on rumen fungi[121,123], as assessed by extent of biomass degradation, co-cultivation of the close relative *F.* sp. UWB7 with *A. robustus* or *C. churrovis* resulted in drastic changes to both bacterial and fungal transcriptomes, including upregulation of bacterial drug-efflux pumps and fungal chaperones, polyketide O-methyltransferases, PKSs, and NRPSs. Furthermore, fungal genes encoding putative histone modifying enzymes were upregulated in co-culture. Histone 3 lysine 27 trimethylation marks increased and heterochromatin-associated protein-1 was downregulated in co-culture. Together, these results suggest that, similar to higher-order fungi, anaerobic fungi regulate their secondary metabolism via chromatin remodeling. Taken together, these data support that anaerobic gut fungi activate their secondary metabolism via epigenetic and transcriptional regulation when challenged by rumen bacteria. The metabolic outcome of

these transcriptional changes may be the production of a fungal defense compound produced by a PKS or NRPS. Consistent with this hypothesis, untargeted nonpolar metabolomics supports that at least one fungal metabolite is enriched by co-cultivation with *F.* sp. UWB7. Anaerobic fungi and the antagonistic relationships of the rumen microbiome may prove to be a valuable source of antibiotics.

4.4   *Materials and Methods*

4.4.1   *Isolation and cultivation of anaerobic gut fungi*

Anaeromyces robustus was isolated via reed canary grass enrichment from the fecal pellet of a Churro sheep at the Santa Barbara Zoo, as described previously[6,7]. *Caecomyces churrovis* was isolated similarly [29]. Both fungi were cultivated anaerobically in Hungate tubes at 39 °C with reed canary grass as the carbon source in a modified formulation ("MC-") of complex medium C[262], containing 0.25 g/L yeast extract (before boiling), 0.5 g/L Bacto[TM] Casitone (before boiling) and 7.5 vol% clarified rumen fluid. The media was supplemented with vitamins after autoclaving as described by Teunissen and colleagues[210]. Cultures were passaged every 3-4 days into fresh media via a 1.0 mL sterile syringe.

4.4.2   *Cultivation of Fibrobacter sp. UWB7*

The strain *Fibrobacter* sp. UWB7 was a generous gift from Garret Suen at the University of Wisconsin-Madison. Details of the isolation of this strain are described in Neumann and colleagues[241]. *F.* sp. UWB7 was cultivated at 39 °C anaerobically in Hungate tubes containing 9.0 mL of MC- media supplemented with vitamin solution, prepared as described above, and 0.1 g of Avicel® PH-101 (Sigma Aldrich, St. Louis, MO). Each Hungate tube was inoculated with 1.0 mL of cryostock or live *F.* sp. UWB7 culture.

4.4.3 Overview of the co-cultivation conditions of anaerobic gut fungi with Fibrobacter sp. UWB7

An overview of the co-cultivation pairings, carbon substrates, and co-cultivation incubation times is depicted in Figures 4.1 and 4.2. Briefly, *A. robustus* was co-cultivated with *F.* sp. UWB7 on both Avicel® PH-101 (Sigma Aldrich, St. Louis, MO) and milled switchgrass (gift from U.S. Department of Agriculture), whereas *C. churrovis* was co-cultivated with *F.* sp. UWB7 on switchgrass only, due to the slow growth of *C. churrovis* on Avicel®[29]. Since *A. robustus* and *C. churrovis* are expected to grow more slowly than *F.* UWB7, as evidenced by the order-of-magnitude larger specific growth rate of *F. succinogenes* compared to *C. churrovis* on soluble sugars[29,263], both strains of anaerobic gut fungi were allowed to grow for 24 hours prior to inoculation with *F.* sp. UWB7. Co-cultures were subsequently allowed to grow for an additional 24-72 hours prior harvesting for RNA extraction. The length of each co-cultivation Incubation for the co-cultivation pairings was set by time necessary for the fungus to reach mid-log growth phase, as assessed by cumulative pressure[217] of fungal monocultures as well as visual assessment.

4.4.4 Co-cultivation of anaerobic fungi with Fibrobacter sp. UWB7

Anaerobic liquid growth medium "MC-" was prepared following the recipe for complex medium C[262], with yeast extract, Bacto™ Casitone, and clarified rumen fluid reduced to 0.25 g/L (concentration before boiling), 0.5 g/L (before boiling), and 7.5 vol%, respectively. A 100-mL capacity serum bottle was filled with 80 mL of MC- liquid medium and 0.8 g of switchgrass or Avicel® PH-101 (Sigma Aldrich, St. Louis, MO). The serum bottle and its contents were flushed with a CO_2, autoclaved, and supplemented with 0.8 mL of 100x vitamin solution[210]. The serum bottle was preheated to 39 °C and then a seed culture of *A.*

robustus was started by inoculating 1.0 mL of the routinely passaged *A. robustus* described above into the liquid medium using a 1 mL sterile syringe. The seed culture was immediately vented following inoculation and then incubated at 39 °C for four days. This seed culture was used to inoculate cultures to be harvested for subsequent RNA-seq. Cultures were prepared in replicates of four by inoculating 1.0 mL of *A. robustus* into 8.0 mL (co-culture) or 9.0 mL (monoculture) of MC- containing 0.1 g switchgrass or Avicel® PH-101 (Sigma Aldrich, St. Louis, MO). Prior to inoculation, the medium and substrate were autoclaved, supplemented with 0.1 mL of 100x vitamin solution[210] post-autoclaving, and preheated to 39 °C. The fungal culture were grown at 39 °C for 24 h and then co-cultures were started by inoculating 1.0 mL of *F.* sp. UWB7 into each of four replicates. The seed culture of *F.* sp. UWB7 was grown for 24 h at 39 °C from 1.0 mL of inoculum in a serum bottle containing 80 mL of MC-, 0.8 g of Avicel® PH-101 (Sigma Aldrich, St. Louis, MO), 0.8 mL of 100x vitamin solution[210], and a CO_2 headspace. *A. robustus* monocultures and *A. robustus-F.* sp. UWB7 co-cultures were incubated for an additional 24 h (switchgrass substrate) or 48 h (Avicel®). The contents of each Hungate tube were transferred to 15 mL Falcon tubes (Fisher Scientific, Waltham, MA) and centrifuged at 3200 *g* and 4 °C using a swinging bucket rotor (Eppendorf[TM] A-4-81) for 10 min. The supernatant was saved at -80 °C for subsequent LC-MS/MS and 1.0 mL of RNAlater® (Sigma Aldrich, St. Louis, MO) was added to the pellet to preserve the RNA. The pellet from each culture was frozen at -80 °C until lysis.

The co-cultivation of *C. churrovis* with *F.* sp. UWB7 was performed identically to the co-cultivation of *A. robustus* with *F.* sp. UWB7, except that length of the co-cultivation incubation was 40 h.

4.4.5 RNA extraction and QC

Samples were thawed on ice and then centrifuged at 3220 *g* at 4°C using a swinging bucket rotor (Eppendorf™ A-4-81) for 10 min. RNAlater® was decanted. The pellets were transferred into 2-mL screw-cap tubes containing 1.0 mL of autoclaved 0.5 mm zirconia/silica beads (Biospec) and 600 μL buffer RLT (QIAGEN, Hilden, Germany) with 1 vol% 2-mercaptoethanol (Sigma Aldrich, St. Louis, MO). The tube was briefly vortexed and then the cells were lysed using the Biospec Mini-Beadbeater-16 for 45 s. The tubes were placed on ice for 30 s. Following lysis, the tubes were centrifuged for 3 min at 13,000 *g* and 20 °C using a microcentrifuge (Eppendorf™ 5424). The supernatant was removed using gel-loading tips (Fisher Scientific, Waltham, MA) to maximize yields and centrifuged again to remove residual cell debris for 2 min at 20,000 *g*. The supernatant from each tube was then transferred into 2 mL round-bottom sample tubes (QIAGEN catalog number 990381). RNA was extracted using a QIAcube following the RNeasy Mini protocol for animal cells with QIAshredder homogenizenization and optional on-column DNase digest.

Quantity and quality of RNA was assessed by a QuBit fluorometer and TapeStation (Agilent), respectively. All RNA integrity number equivalents (RIN^e) were above 6.0, assessed by either eukaryotic or prokaryotic ribosomal markers for co-cultures.

4.4.6 RNA library preparation and sequencing

In fungal monocultures, fungal mRNA was selectively enriched by capturing polyadenylated RNA using poly-T beads. For bacterial monocultures, ribosomal RNA was depleted using the Illumina Ribo-Zero rRNA Removal Kit (Yeast) spiked into the Illumina Ribo-Zero Gold rRNA Removal Kit (Epidemiology). To obtain both bacterial and fungal libraries from the co-cultures, each sample was divided and 200 ng was used as the input

into each alternative pipeline: 1) polyA selection for the fungal library, or 2) ribosomal depletion by Illumina Ribo-Zero rRNA Removal Kit (Yeast) spiked into the Illumina Ribo-Zero Gold rRNA Removal Kit (Epidemiology) for the library enriched in bacterial mRNA. Stranded RNASeq libraries were created by the Joint Genome Institute and quantified by qPCR. Libraries were sequenced by paired-end 150 bp reads using the NovaSeq (Illumina, San Diego, CA).

4.4.7   RNA-seq data analysis

Raw reads were evaluated for artifact sequences using BBDuk[264]. Detected artifacts identified using kmer matching (kmer=25) were trimmed from the 3' end of reads. Reads were further filtered by removing RNA spike-in reads, PhiX reads, and reads containing any N's. Reads were trimmed for quality using the phred trimming method (set at Q6). Following trimming, short reads of less than 50 bases were removed. Filtered reads were aligned to the reference genome for the respective organism (fungal genomes available on the MycoCosm portal[23] and *F.* sp. UWB7 genome GenBank assembly accession GCA_900142945.1) using HISAT2 version 2.1.0[220]. Raw gene counts were generated by featureCounts[221] using the gene annotation files available in MycoCosm[23] for *A. robustus* or *C. churrovis* and the IMG[248] for *F.* sp. UWB7. Differential expression analysis was performed using DESeq2 (version 1.18.1)[222], with a minimum absolute $\log_2$fold change of 1.0 and statistical significance threshold of *p*-adjusted less than 0.05.

4.4.8   Extraction and LC-MS/MS

2 mL of ethyl acetate was added to 1.5 mL of fungal supernatant, vortexed then sonicated for 10 min in a water bath (room temperature), centrifuged (5 min at 5000 rpm),

and then the top ethyl acetate layer was removed to another tube. Extracts were dried in a SpeedVac (SPD111V, Thermo Scientific) and stored at -80 °C.

In preparation for LC-MS analysis, 150 µL LC-MS grade methanol containing 1 µg/mL internal standard (2-Amino-3-bromo-5-methylbenzoic acid, Sigma) was added to dried extracts, followed by a brief vortex and sonication in a water bath for 10 minutes. 150 µL of resuspended extract was then centrifuge-filtered (2.5 min at 2500 rpm) using a 0.22 µm filter (UFC40GV0S, Millipore) and transferred to a glass autosampler vial. Reverse phase chromatography was performed by injecting 2 µl of sample into a C18 chromatography column (Agilent ZORBAX Eclipse Plus C18, 2.1x50 mm 1.8 µm) warmed to 60°C with a flow rate of 0.4 mL/min equilibrated with 100% buffer A (100% LC-MS water with 0.1% formic acid) for 1 minute, followed by a linear gradient to 100% buffer B (100% acetonitrile w/ 0.1% formic acid) at 7 minutes, and then held at 100% B for 1.5 minutes. MS and MS/MS data were collected in both positive and negative ion mode using a Thermo Q Exactive HF mass spectrometer (ThermoFisher Scientific, San Jose, CA), with full MS spectra acquired ranging from 80-1200 *m/z* at 60,000 resolution, and fragmentation data acquired using an average of stepped collision energies of 10, 20 and 40 eV at 17,500 resolution. Sample injection order was randomized and an injection blank of methanol only run between each sample. Raw data is available for download at https://genome.jgi.doe.gov/portal/.

4.4.9   Metabolomics data analysis: creation of molecular network via Global Natural Products Social Molecular Networking

The Feature-Based Molecular Networking (FBMN) workflow[265] on GNPS[168] (https://gnps.ucsd.edu) was used to construct a molecular network. First, peak-finding was

performed with MZmine (version 2.39)[188]. An MZmine workflow was used to generate a list of features (*m/z*, rt values obtained from extracted ion chromatograms containing chromatographic peaks within a narrow *m/z* range) and filtered to remove isotopes, adducts and features without MS/MS. ADAP chromatogram builder and deconvolution modules were used[266].

The mass spectrometry data were deposited on the MassIVE public repository (MSV000086033).

The molecular networking GNPS job can be publicly accessed at:

https://gnps.ucsd.edu/ProteoSAFe/status.jsp?task=aeeb3b1a8fac4b67b54b6f1171a3053f

For each feature, the most intense fragmentation spectrum, was uploaded to GNPS. All MS/MS fragments were removed within +/- 17 Da of the precursor m/z. Window-filtering was achieved by selecting only the top 6 fragment ions in the +/- 50 Da window throughout the spectrum. Parameters were set as follows: precursor ion mass tolerance of 0.05 Da and MS/MS fragment ion tolerance of 0.05 Da. The edges of the molecular network were specified to have a cosine score greater than 0.70 and more than 6 matching peaks. The edges were further filtered such that an edge was permitted if and only if the joined nodes were present in the other respective node's top 10 most similar nodes. Lowest scoring edges were removed from molecular families such that no family contained more than 100 nodes. All spectra within the molecular network were queried against GNPS spectral libraries[168]. Each library spectrum was filtered following the same procedure as that applied to the input data. The minimum criteria for a match between a network spectra and a library spectra were that the score be greater than 0.7 and that at least 6 peaks match. MS/MS spectra were annotated by DEREPLICATOR[267].

It should be noted that a spectra match to a database spectra is not a definitive identification of the feature: it could be an isomer with a similar fragmentation, an ion with a close but not exact m/z but similar fragmentation pattern, or an in-source degradation product of another larger molecule (the degradation product may look similar to the database match).

GNPS positive and negative mode networks were merged using a custom Python script to group nodes having a retention time difference less than 0.15 minutes and an *m/z* difference less than 20 parts-per-million assuming the negative mode species ionized as an [M-proton]- and the positive mode species ionized as an [M+proton]+.

Finally, the molecular network was visualized using Cytoscape[259] and three-way coloring. Given three numerical values to compare, the corresponding hue for each value can be calculated according to Baran and colleagues[260] using a custom Python script. The transparency of each node is determined by the value of each normalized to the minimum and maximum of the set of values. In this case, the three values to compare were the GNPS-normalized peak areas of each feature (averaged across four biological replicates), for three different treatments: (1) *A. robustus* monoculture (Avicel® substrate), (2) *F.* sp. UWB7 monoculture (Avicel® substrate), and (3) *A. robustus-F.* sp. UWB7 co-culture (Avicel® substrate). The method of per sample normalization selected in the GNPS job was "row sum normalization (per file sum to 1,000,000)," and the mean was chosen as the aggregation method per group (treatment). For the molecular network depicted in Figure 4.7, the transparency of each node was normalized with respect to the minimum and maximum GNPS-normalized peak areas of the co-culture condition.

Structure and class prediction of the unknown bacterial metabolites (glycophosphoethanolamines, Figure 4.7) was performed by SIRIUS 4.0[171] and

CANOPUS[172] by the MS-GF+ (1.3.0) Workflow in the ProteoSAFe web server from the Center for Computation Mass Spectrometry. The job may be viewed and cloned from the following URL:

https://proteomics2.ucsd.edu/ProteoSAFe/status.jsp?task=e49148e8624c4e4cb0c3fbe09918ab6c

4.4.10 Principal Component Analysis (PCA) of metabolomics data using MetaboAnalyst 4.0

MetaboAnalyst 4.0[268] was used construct principal component analysis (PCA) plots from the peak heights feature table of all samples, generated using MZmine2[188,269]. The "prcomp" function in R[270], which requires the package "chemometrics," was used internally within the MetaboAnalyst to perform the PCA.

4.4.11 Growth of fungal cultures for Western blotting analysis of fungal epigenetic modifications

The seed culture of *A. robustus* was prepared as follows: 1.0 mL of *A. robustus* from routine cultivation was transferred into a 60 mL glass serum bottle preheated to 39 °C containing 40 mL of MC- media with 0.4 g of Avicel® PH-101 (Sigma Aldrich, St. Louis, MO) and supplemented with 0.4 mL of 100x vitamin solution[210] after autoclaving. The seed culture was grown for four days. From the seed culture, 1.0 mL was inoculated into each of six 80 mL cultures of MC- with 0.8 g Avicel® PH-101 (Sigma Aldrich, St. Louis, MO) supplemented with 0.8 mL of 100x vitamin solution[210]. Three of these cultures were incubated at 39 °C for 24 h and then each bottle was inoculated with 1.0 mL of *F*. sp. UWB7 seed culture. The remaining three bottles were also incubated at 39 °C, but they not inoculated with *F*. sp. UWB7. The *F*. sp. UWB7 seed culture was prepared by inoculating

1.0 mL of active culture into a 60 mL serum bottle containing 40 mL of MC- supplemented with vitamin solution[210] and 0.4 g of Avicel® PH-101 (Sigma Aldrich, St. Louis, MO). The active culture was *F.* sp. UWB7 revived from cryostock one week prior and passaged one time. The co-cultures and monocultures were grown for a total of 72 h following the fungal inoculation. The cultures were then transferred into 50 mL Falcon tubes and centrifuged using a fixed angle rotor (Eppendorf F-34-6-38) at 4 °C and 3000 *g* for 10 min. The cell pellets were stored at -80 °C until lysis.

4.4.12 Extraction of fungal cultures for Western blotting analysis of fungal epigenetic modifications

The frozen cell pellets prepared above were resuspended in 3 ml of 2M NaOH (Fisher Scientific, Waltham, MA, USA) with 10% V/V beta mercaptoethanol (Sigma Aldrich, St. Louis, MO, USA). The solution was gently mixed and incubated on ice for 5 minutes to promote hydrolysis of the fungal cell wall. The solution was then centrifuged at 14,000 *g* for 30 seconds at 4 °C. The resulting pellet was resuspended in 3 ml of high salt extraction buffer containing 40 mM HEPES pH 7.5 (Fisher Scientific, Waltham, MA, USA), 350 mM NaCl (Fisher Scientific, Waltham, MA, USA), 0.1% W/V Tween 20 (BioRad, Hercules, CA, USA), and 10% V/V glycerol (Fisher Scientific, Waltham, MA, USA). The solution was immediately centrifuged at 14,000 *g* for 30 seconds at 4 °C. The cell pellets were resuspended in 3 ml of 2X Sodium Dodecyl Sulfat e (SDS) sample buffer, 0.1M Tris-HCl pH 6.8 (Fisher Scientific, Waltham MA, USA), 4% W/V sodium dodecyl sulfate (Fisher Scientific, Waltham, MA, USA), 0.2% W/V bromophenol blue (Sigma Aldrich, St. Louis, MO, USA), and 20% V/V glycerol (Fisher Scientific, Waltham, MA, USA), and 10% V/V beta mercaptoethanol (Sigma Aldrich, St. Louis, MO, USA). Cell pellets incubated at 100

°C for 10 minutes, prior to being centrifuged at 14,000 *g* for 30 seconds at 4 °C. The supernatants were stored at -20 °C until further use.

4.4.13 *Western blotting of fungal epigenetic modifications*

30 microliters of previously frozen cell lysates prepared in SDS sample buffer were gently mixed prior to loading on a 15% polyacrylamide gel. *Candida glabarata* whole cell lysate, extracted as mentioned in extraction of fungal cultures for Western blotting, was used as histone H3 nuclear loading control as well as a positive control for H3K4me3 Western blots. *Piromyces* sp. UH3-1 whole cell lysate, was used as a loading control for H3K27me3 Western blots. Gel electrophoresis occurred for 65 minutes at 150 volts under constant voltage at room temperature. The gel and Immobilion PVDF membrane (Fisher Scientific, Waltham, MA, USA) were briefly washed with 100% methanol (Fisher Scientific, Waltham, MA, USA) prior to being washed with 1X Towbin buffer containing 25 mM Tris pH 8.3 (Thomas Scientific, Swedesboro, NJ, USA), 192 mM glycine (Sigma Aldrich, St. Louis, MO, USA), and 10% V/V methanol. The gel and membrane were overlayed on top of 9 pieces of 3MM chromatography paper (Fisher Scientific, Waltham, MA, USA) that were soaked in 1X Towbin buffer. After overlaying the gel and membrane on top of the 9 sheets of chromatography paper, an additional six sheets of chromatography paper already saturated with 1X Towbin buffer were overlayed on top of the gel. Both membrane and chromatography paper were previously cut to dimensions of 5.5 cm X 8.5 cm in order to match the dimensions of the resolving gel. After rolling out the transfer sandwich to remove air bubbles, the proteins were transferred under semi-dry conditions using a Hoefer Hsi Semi-phor TE70 semi-dry transfer unit (Holliston, MA, USA) for 90 minutes at 42 mAmps under constant amperage at room temperature. Membranes were blocked overnight with 3%

W/V milk (Great Value, Bentonville, AR, USA) with 0.15% W/V sodium azide (Sigma Aldrich, St. Louis, MO, USA) dissolved in 1X Tris Buffered Saline pH 7.5 (Fisher Scientific, Waltham, MA, USA) at 4 °C. The following day, membranes were washed at room temperature for 30 minutes with 1X TBS buffer pH 7.5, exchanging the buffer every 10 minutes. Primary antibodies were diluted in 10ml of 1X TBS buffer and incubated with the membranes for approximately 3 hours at room temperature on a rocker at slow speed. Rabbit anti-histone H3 antibody (#ab1791, Abcam, Cambridge, MA, USA) diluted 1:10,000 in 1X TBS buffer was used as a nuclear loading control. Rabbit anti H3K4me3 antibody (#39016, Activemotif, Carlsbad, CA, USA) was diluted 1:50,000 in 1X TBS buffer. Rabbit anti H3K27me3 antibody (07-449, Upstate-Millipore, Lake Placid, NY, USA) was diluted 1:5000 in 1X TBS pH 7.5. After 3 hours the blots were washed with 1X TBS pH 7.5 for 30 minutes, exchanging the buffer every 10 minutes. Horse radish peroxidase conjugated goat anti-rabbit secondary antibody (#111-035-003, Jackson ImmunoReasearch, West Grove, PA, USA) was diluted 1:10,000 in 1X TBS pH 7.5 and 10 ml of this solution was added to each blot, which incubated on a rocker for 3 hours at room temperature. The blots were washed with 1X TBST buffer (10 mM Tris pH 7.5, 100 mM NaCl, 0.1% W/V Tween 20) for 30 minutes, exchanging the buffer every 10 minutes. 300 microliters of Crescendo horse radish peroxidase reagent (Fisher Scientific, Waltham, MA, USA) was added to each blot, and the blots were imaged in a BioRad Chemidoc imager under the default chemiluminesence settings and auto-adjusted exposure time. For the H3 blots, the positive loading control was masked during auto-adjusted exposure to avoid overwhelming the sample signals.

4.4.14 Helium ion microscopy

The *A. robustus-F.* sp. UWB7 co-culture was prepared as described above. The cell pellet, including the Avicel® growth substrate, was harvested and suspended in phosphate buffered saline (PBS, pH 7.5) in a 15 mL Falcon tube, to which glutaraldehyde (Sigma-Aldrich, St. Louis, MO) was added to a final concentration of 2 vol%. The tubes were incubated at room temperature for 1 hour on a rotator. The tubes were then centrifuged at 700 g for 10 minutes at 4 °C and the buffer was removed. The pellet was resuspended in 10 mL of 25 vol% ethanol and incubated for another hour. This process of suspension, incubation, and centrifugation was repeated for a stepwise ethanol dehydration series with 30%, 50% and 70% ethanol steps. Twice more the pellet was washed with 10 mL of 100% ethanol and incubated for 15 min and finally resuspended in 5 mL of 100% ethanol. The cells were then dried via critical point drying with an Autosamdri-815 (Tousimis, Rockville, MD) and carbon dioxide as a transitional fluid, sputter-coated with conductive carbon, and imaged by an Orion helium ion microscope (Carl Zeiss Microscopy, Peabody, MA).

4.4.15 Data availability

All sequencing reads has been deposited in the Sequencing Read Archive (SRA) and are associated with NCBI BioProject PRJNA666900.

5 Anaerobic fungal genomes encode stress response genes similar to early eukaryotic parasites

Secondary metabolism is one way that fungi respond to specific challenges encountered in their environment. Two other mechanisms that fungi use to cope with stress are the unfolded protein response (UPR) and heat shock response (HSR). Unlike secondary metabolism, UPR and HSR are shared with all eukaryotes, although the former has distinct adaptions in different kingdoms and the latter in different species.

5.1 Introduction

The ability to cope with different kinds of environmental stress is a universal trait of life. However, depending on the type of stress, the response can be highly conserved between organisms, such as the heat shock response (HSR), or the response can vary significantly between the kingdoms of life, such as in the case of the unfolded protein response (UPR). Heat shock causes both cytosolic and ER stress, triggering the HSR and potentially the UPR, in which cells respond to misfolded proteins by attenuating protein production. In biotechnology, cellular stress responses can decrease product titer or even cause cell death. Fungi have adapted to cope with many kinds of stressors in their native environments. Heat shock proteins (hsps) are used to respond to cytosolic stress triggered by various stimuli, including heat[271–273], osmotic pressure[274], and low pH[275,276]. The expression of hsps in the HSR as well as the UPR are part of the comprehensive response of an organism, called the environmental stress response, to various types of environmental stresses [277,278]. The environmental stress response is well-established for Ascomycota, especially for

Saccharomyces cerevisiae[279,280]. Both the HSR and UPR are often triggered during expression of proteins, especially those that are a product of heterologous expression efforts[281–283].

Although the UPR is induced by stress to the ER rather than the cytosol, some features of the UPR and HSR are shared, such as the upregulation of genes with functions in polypeptide translocation, vesicular transport from the ER, and ER-Associated Degradation (ERAD) in order increase protein-folding capacity and to limit protein production[284]. Although the UPR is found in all eukaryotes, the cellular machinery involved in the response differs between the kingdoms of life[285]. The unfolded protein response of fungi is well-established, especially for yeasts and filamentous fungi applied in biotechnology[282,286–288]. Fungi are known to possess the transmembrane endoribonuclease Ire1, which senses misfolded proteins, oligomerizes and autophosphorylates, activating an endoribonuclease domains that removes a constitutively expressed intron from mRNA encoding the transcription factor Hac1[289,290]. This transcription factor triggers the upregulation of genes encoding chaperones to promote protein folding and those with functions in vesicular transport and protein turnover[287,291,292]. In contrast, members of Metazoa use both the Ire1-initiated pathway as well as two others: ATF6 and PERK[285,293] PKR-like Endoplasmic Reticulum Kinase (PERK) is also a transmembrane sensing protein, but instead of cleaving an intron, it phosphorylates the eukaryotic initiation factor 2 alpha subunit (eIF2α), which prohibits GTP exchange and subsequent delivery of Met-tRNA to initiate translation[294–296]. This event results in global inhibition of protein translation, except for certain proteins[293].

PERK is not the only kinase of eIF2α that regulates translation in eukaryotes. Different kinases of eIF2α evolved to sense diverse types of environmental stress: ER stress (PERK); nutrient limitation (GCN2); viral infection (PKR/PKZ); and heme deprivation, heat

shock, or oxidative stress (HRI)[297]. The domain architecture for each kinase is distinct[297]. The only eIF2α kinase with a transmembrane domain is PERK. The archetypical PERK also contains an N-terminal signal peptide, IRE1-like stress sensing and dimerization domain, and a catalytic kinase domain with an internal insert[297]. GCN2 is ubiquitously present in fungi, whereas HRI is not found in all fungi. To date, no instance of a PERK-like kinase has been reported in fungi[297].

Understanding how to circumvent cellular stress responses facilitates fungal bioprocessing, especially heterologous protein production, as illustrated by previous studies of filamentous fungi[298] and yeasts[283,299,300]. With appropriate genetic engineering strategies, knowledge of the UPR can be leveraged for enhanced production of secreted proteins can to improve product titers, without the induction of deleterious effects such as ER-Associated Degradation (ERAD)[298]. One successful example was the sevenfold boost in production of laccase by *Aspergillus niger* var. *awamon* as a result of the overexpression of active *hacA* transcription factor[301]. Similarly, constitutive overexpression of *HAC1* in *Saccharomyces cerevisiae* resulted in a 70% increase in the secretion of foreign protein α-amylase from *Bacillus amyloliquefaciens*[300]. However, this approach is not successful for all heterologous production schemes. For example, constitutive overexpression of *HAC1* in *Saccharomyces crevisiae* did not increase in the secretion of the foreign protein endoglucanase from *Trichoderma reesei*[300].

However, our knowledge of the fungal stress response has largely been limited to Dikarya (higher-order fungi). Neocallimastigomycetes, a clade of the phylum Chytridiomycota, are early-diverging fungi native to the digestive tracts of large herbivores. They specialize in the degradation of plant biomass and possess the largest array of carbohydrate active enzymes (CAZymes) of any sequenced fungi to date and thus hold great

promise for applications in biotechnology[6,302,303]. Recent improvements in anaerobic fungal cultivation[207,304] and the sequencing of high-quality genomes[7] and transcriptomes[7] have advanced our understanding of these non-model organisms, but the HSR and UPR have not been characterized. As some of the most ancestral fungi on the fungal evolutionary tree[305], Neocallimastigomycetes provide unique evolutionary insights when compared to other early-branching eukaryotes.

Here, we delineate the major components of the UPR, including the key stress-sensing enzymes, chaperones, and some target genes associated with the secretory pathway. We further validate the identified components of the HSR and UPR within Neocallimastigomycetes by subjecting two representative strains from this class (*Anaeromyces robustus* and *Neocallimastix californiae*) to heat shock and subsequently tracking their transcriptomic response. Lastly, we report on and quantify an unusual prevalence of small heat shock proteins within the genomes of Neocallimastigomycetes relative to other fungi, which may have implications in the stability of their proteomes.

5.2   Results and Discussion

5.2.1   Neocallimastigomycetes share components of the metazoan unfolded protein response

The stress response of Ascomycota is well-established in regards to various environmental stimuli[277–280], recombinant protein expression[282,283,299], and especially the components of the UPR[287] and HSR[306] in *Saccharomyces cerevisiae*. However, knowledge of the stress response of early-diverging fungi in the fungal tree of life, such as

Neocallimastigomycetes, and the extent of its conservation with other fungi or kingdoms of life is lacking.

We delineated components of the UPR and target genes of the secretory pathway in four strains of Neocallimastigomycetes (*Anaeromyces robustus*, *Caecomyces churrovis*, *Neocallimastix californiae*, and *Piromyces finnis*) by homology to model organisms (Table 5.1). Homologs of Ire1 were identified in all four genomes, as well as the critical chaperones KAR2, calnexin, and calreticulin, and the catalyst of disulfide bond formation ERO1 (ER oxidoreductin 1). In most cases, the genes were highly conserved between Neocallimastigomycetes and *S. cerevisiae*, with percent identities and coverages greater than 40 and 80%, respectively (Table 5.1). *S. cerevisiae* only possesses a single gene with similarity to the calnexin of mammalian cells, rather than separate calnexin and calreticulin proteins[307]. In contrast, homologs of *S. cerevisiae* calnexin and *D. melanogaster* calreticulin were identified in all four Neocallimastigomycetes (Table 5.1). Surprisingly, protein sequences similar to the metazoan PERK transmembrane protein were also identified (Figure 5.1, Table 5.1).

Closer inspection of the PERK-like sequences in Neocallimastigomycetes showed sequence similarity only in the catalytic domain (Figure 5.2) and not the luminal dimerization domain. The luminal domain of bovine PERK has been shown to bind directly with misfolded proteins, which is thought to induce dimerization and autophosphorylation[308]. Although the presence of an additional UPR pathway in fungi is unprecedented, kinases of eIF2α that respond to ER stress have been identified in *Toxoplasma gondii*[309], an early eukaryotic parasite belonging to the Apicomplexa phylum. *Toxoplasma gondii* has two eIF2α kinases: TgIF2K-A responds to ER stress, and TgIF2K-B

is thought to respond to cytosolic stress[309]. Both TgIF2K-A and TgIF2K-B sequences are divergent from metazoan PERK except in the kinase domain[309].

The domain architecture of PERK candidates from Neocallimastigomycetes (Table 5.1) is illustrated in Figure 5.3. *N. californiae* 503500 and *P. finnis* 367815 both contained N-terminal PAS domains, but *A. robustus* 13827 and *C. churrovis* 543476 lack these domains. *A. robustus* 13827 is located at the start of a scaffold, which indicates that the gene may be truncated. Thus, the full gene may contain a PAS domain. PAS domains are known sensing modules of signal transduction proteins, such as kinases[310,311]. Therefore, these regions may perform sensing of misfolded proteins, similar to the IRE1-like regulatory region of metazoan PERK. SignalP-5.0[312] predicted no signal peptides in the PERK-like kinases from Table 5.1, but TMHMM2.0[225] identified a transmembrane helix in all sequences. However, the transmembrane helices were C-terminal relative to the kinase domain, which is different than the transmembrane kinases identified in Apicomplexa, in which case the transmembrane helices were N-terminal relative to the kinase domain[309] (Figure 5.3). Sequence alignment by Clustal Omega[313] and visualization with Jalview[314] of PERK-like kinases from Neocallimastigomycetes to representative members of Apicomplexa, including TgIF2K-A from *T. gondii*, only demonstrated consensus in the kinase domain regions. However, sequence alignment of *P. finnis* 367815, *N. californiae* 503500, *C. churrovis* 543476, and *A. robustus* 13827 indicated a high degree of similarity, even in regions outside of the kinase domain (Figure 5.4). Given the sequence similarity in the kinase domain to metazoan PERK proteins, as well as the precedent of transmembrane kinases that respond to ER stress in apicomplexans, it is probable that the transmembrane kinases of Neocallimastigomycetes perform a similar function in sensing and responding to unfolded proteins.
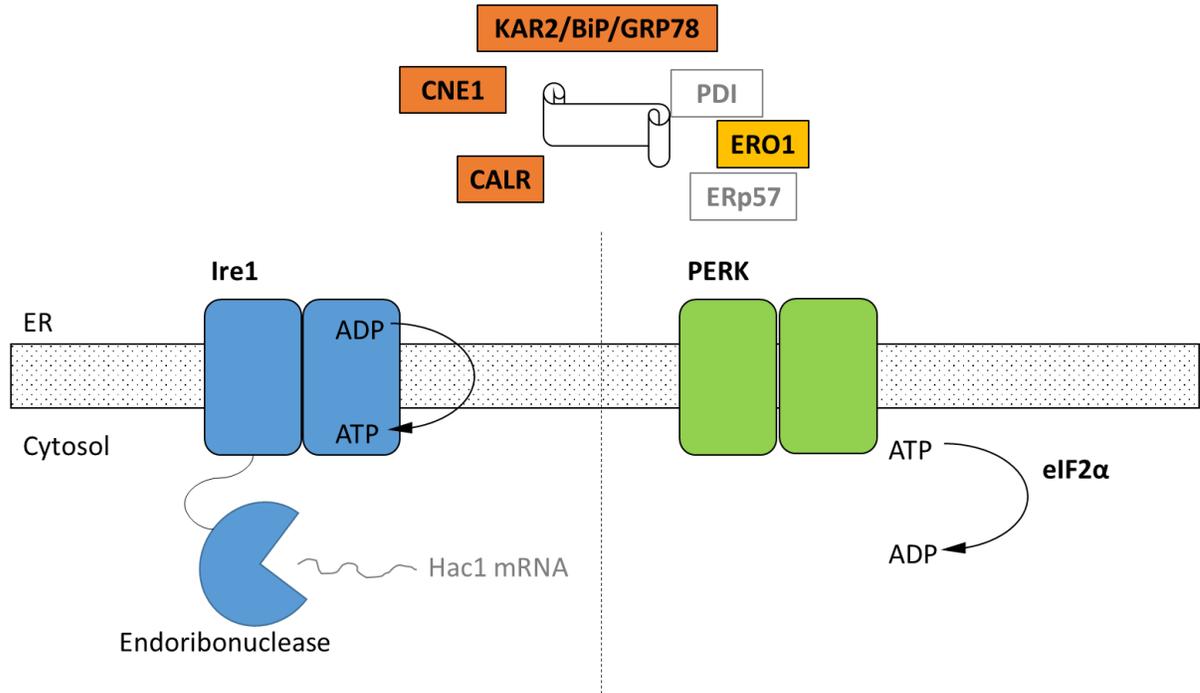
Figure 5.1: Schematic of the cellular components of the UPR identified within the genomes of four representative strains of Neocallimastigomycetes. Black text indicates homologs of UPR components that were identified in in the genomes of *A. robustus*, C. *churrovis*, *N. californiae*, and *P. finnis* (Table 5.1) whereas gray text signifies parts not identified. Ire1 and PERK are two transmembrane receptors of misfolded proteins. Ire1 activates the Hac1 transcription factor through an alternative splicing event. In the metazoan UPR, PERK oligomerizes and then phosphorylates eIF2α, resulting in global translational repression. KAR2/BiP/GRP78 functions in protein translocation and folding. Following translocation into the lumen, chaperones KAR2 and lectins calnexin (CNE1) and calreticulin (CALR) work to fold the polypeptide. Disulfide bond formation is achieved through Ero1, protein disulfide isomerase (PDI), and ERp57.

Table 5.1: Key machinery of the unfolded protein response in Neocallimastigomycetes and homology to model organisms. Protein Ids are listed from MycoCosm[23]. E-value for all results was less than 10^{-8}. Accession numbers used as query sequences are as follows: KAR2 (NP_012500), IRE1 (NP_011946.1), PERK (NP_649538), eIF2alpha (NP_001285329), and, ERO1 (NP_013576), CNE1 (NP_009343), and CALR (NP_001262430).

Fungus	Protein Id	% Identity	% Subject Coverage
KAR2, *S. cerevisiae*			
A. robustus	292878	69.2	89.0
C. churrovis	594704	52.4	45.4
N. californiae	377732	69.3	89.1
N. californiae	378759	69.3	89.1
P. finnis	579097	69.5	89.1
IRE1, *S. cerevisiae*			
A. robustus	260887	56.3	80.4
C. churrovis	496540	50.7	44.7
N. californiae	384816	50	42.9
N. californiae	412463	58	82.1
P. finnis	301873	54.0	78.2
PERK, *D. melanogaster*			
A. robustus	13827	49.1	25.6
C. churrovis	543476	48.2	24.4
N. californiae	503500	48	21.2
P. finnis	367815	49.8	22.2
eIF2alpha, *D. melanogaster*			
A. robustus	261532	49.6	83.4
C. churrovis	517625	49.2	83.4
N. californiae	500664	48.4	79.5
N. californiae	523066	50.4	84.9
P. finnis	581387	49.3	78.6
ERO1, *S. cerevisiae*			
A. robustus	230787	43.7	43.5
A. robustus	231651	46.5	41.2
C. churrovis	486941	41.1	47.8
C. churrovis	522184	44.6	41.0
N. californiae	517753	45.8	42.5
N. californiae	450298	44.0	41.9
N. californiae	428806	43.9	51.8
P. finnis	402783	45.8	41.7
P. finnis	323943	46.2	43.4
CNE1 (calnexin), *S. cerevisiae*			
A. robustus	221501	40.8	70.8
C. churrovis	422995	43.0	64.3
N. californiae	522025	41.0	70.9
N. californiae	508661	41.0	70.0
P. finnis	580920	43.8	52.5
CALR (calreticulin), *D. melanogaster*			
A. robustus	232658	52.9	70.4
C. churrovis	432815	53.6	65.6
N. californiae	674626	53.8	69.7
P. finnis	580787	53.3	74.3
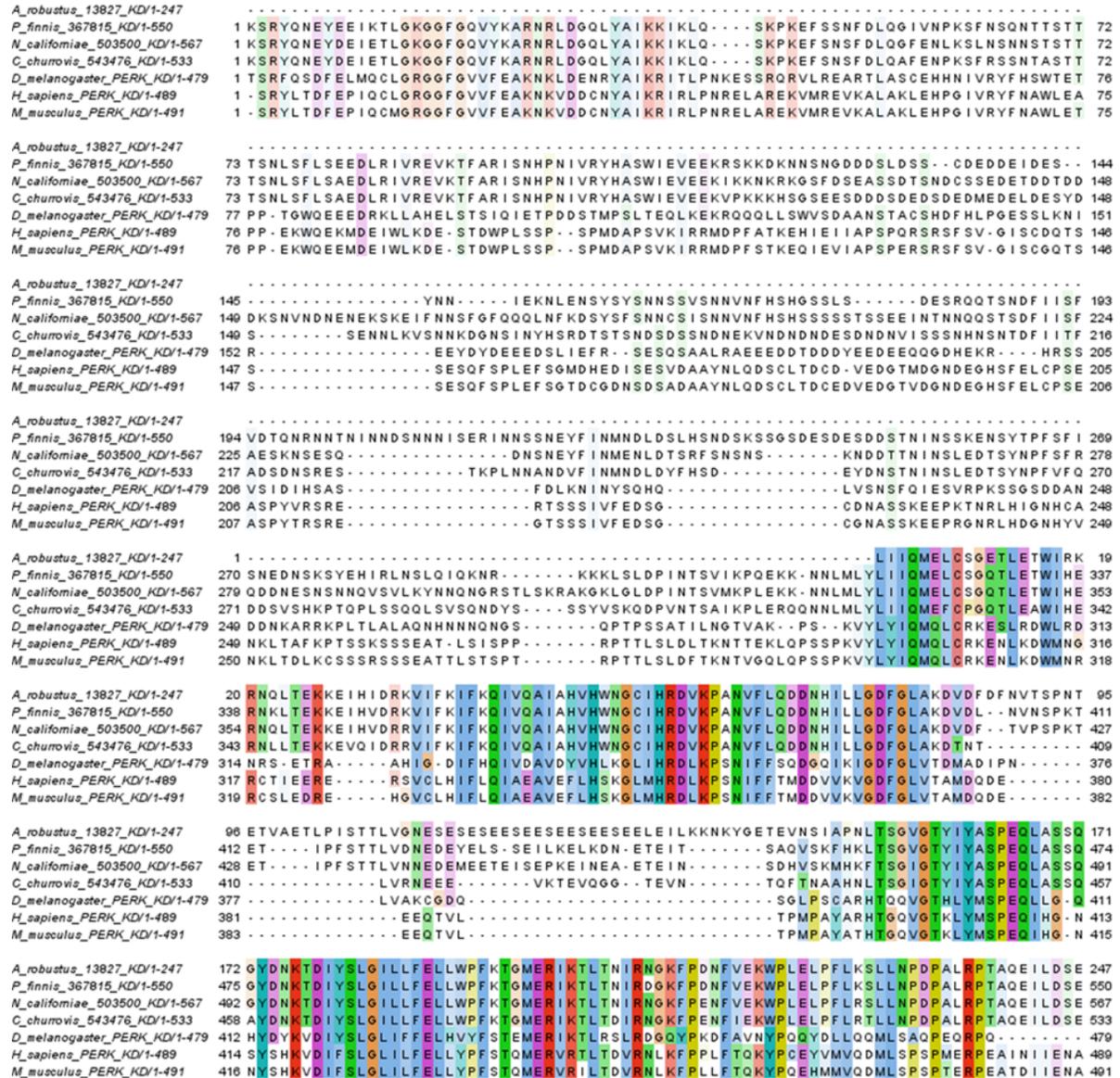
Figure 5.2: Conserved residues between kinase domains of PERK-like proteins from Neocallimastigomycetes and representative kinase domains from metazoan PERK proteins. Black bars indicate protein kinase catalytic domains (Conserved Domain Database[315] accession cl21453). Coloring follows Clustal X designations for amino acid properties: Blue=hydrophobic, red=positive, magenta=negative, green=polar, pink=cysteine, orange=glycine, yellow=proline, cyan=aromatic. A threshold of 30% conservation was used to set transparency.

112
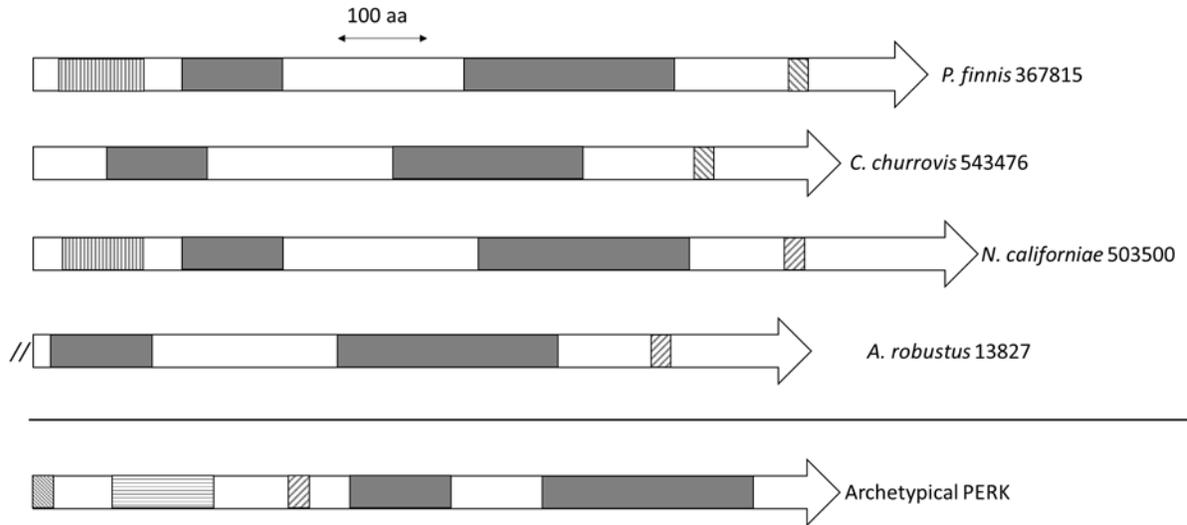
Figure 5.3: Domain architecture of PERK and PERK-like kinasesfrom Neocallimastigomycetes. Domains are indicated by rectangles, with the following shading or pattern fill: vertical lines=PAS domain, left diagonal lines=signal peptide, horizontal lines=IRE1-like regulatory region, right diagonal lines=transmembrane domain, gray=kinase domain. Double slashes indicate that this gene is located at the edge of a scaffold.

Figure 5.4: Conserved residues between PERK-like proteins from Neocallimastigomycetes. Black lines above amino acid residues indicate protein kinase catalytic domains (Conserved Domain Database[315] accession cl21453). Coloring follows Clustal X designations for amino acid properties: Blue=hydrophobic, red=positive, magenta=negative, green=polar, pink=cysteine, orange=glycine, yellow=proline, cyan=aromatic. A threshold of 30% conservation was used to set transparency.

Several PERK-like kinases have been implicated in the regulation of developmental life stages of parasites from the phylum Apicomplexa. TgIF2α kinases have been linked to the formation of a latent state, the bradyzoite cyst, in the life cycle of *Toxoplasma gondii*[309]. In the subfamily *Leishmania*, LinPERK regulates the differentiation of promastigotes to amastigotes[316,317]. TcK2 in *Trypanosoma cruzi* regulates differentiation from epimastigotes

(the proliferative form of the parasite) to metacyclic-trypomastigotes (the infective form of the parasite)[318]. TbeIF2αK2 in *Trypanosoma* has a unique role in responding to persistent ER stress by translocating to the nucleus to override transcription[319]. These examples support more specific roles for apicomplexan PERK-like proteins than a general ER-stress response. Neocallimastigomycetes, although not parasitic, also have a complex life cycle: a motile zoospore seeks and encysts on plant biomass, subsequently growing into a rhizomycelium, which develops sporangia (sac-like structures filled with zoospores) that erupt and release zoospores to begin the cycle again[5]. The regulation of this life cycle is not understood. It has also been hypothesized that some Neocallimastigomycetes form a dormant spore state, which has been observed for one species[320]. It is possible that the PERK-like proteins are similarly linked to developmental changes in the life cycle of Neocallimastigomycetes or even in the formation of a dormant spore state.

The genomes of Neocallimastigomycetes contain a significant number of genes with evidence of horizontal gene transfer (HGT) with rumen bacteria, including genes encoding carbohydrate active enzymes (CAZymes)[7] and non-ribosomal peptide synthetases (Swift et al, manuscript in preparation). We tested the hypothesis that the PERK-like transmembrane kinases may have arisen by HGT with ruminant hosts. Although half of all the top-scoring BLAST hits for the Neocallimastigomycete transmembrane kinases were metazoan proteins, gene phylogenies did not support the occurrence of HGT with ruminants. Similarly, we tested whether apicomplexan transmembrane kinases were the result of HGT with the parasites' hosts, but HGT was not supported except for the case of *Theileria annulata*, in which the gene product (accession XP_953607.1) was sister to metazoan genes from *Piliocolobus*. Therefore, the PERK-like transmembrane kinases in both apicomplexans and Neocallimastigomycetes were likely the result of vertical evolution, although it is still

possible that one or more HGT events occurred in the ancestors of both the apicomplexans and Neocallimastigomycetes.

5.2.2 Comparative transcriptomics of the responses of A. robustus and N. californiae to heat shock reveals affirms signature stress response genes

Neocallimastigomycetes are highly sensitive to temperature fluctuations, most likely due to the fact that they are native to the rumen, which is tightly temperature-controlled at 39 °C[5]. In the laboratory, rumen fungi grow optimally at 39-42 °C[321,322]. To capture the transcriptomic response to heat shock, without inducing cell death, we measured fungal growth curves at a range of heat shock temperatures and for varying durations to test an optimal temperature shift and duration that would limit, but not completely suppress, growth. We found that temperature shock at 48 °C for 15 minute met these criteria.

The global transcriptomic responses of *A. robustus* and *N. californiae* to a 15-minute duration heat shock at 48 °C were captured, with the dynamics of the response measured by time points at 15 minute intervals up to one hour after stress was induced. Differentially regulated genes were identified as those with an absolute $\log_2$ fold change greater than one compared to a control condition without heat shock harvested immediately prior to the heat shock of the test cultures (*p*-adjusted <0.05). The count of up- and downregulated genes at each time point indicates that the greatest response, as measured by the number of differentially regulated genes and the magnitude of the largest $\log_2$ fold change, occurred 45-60 minutes after heat shock (Figure 5.5).
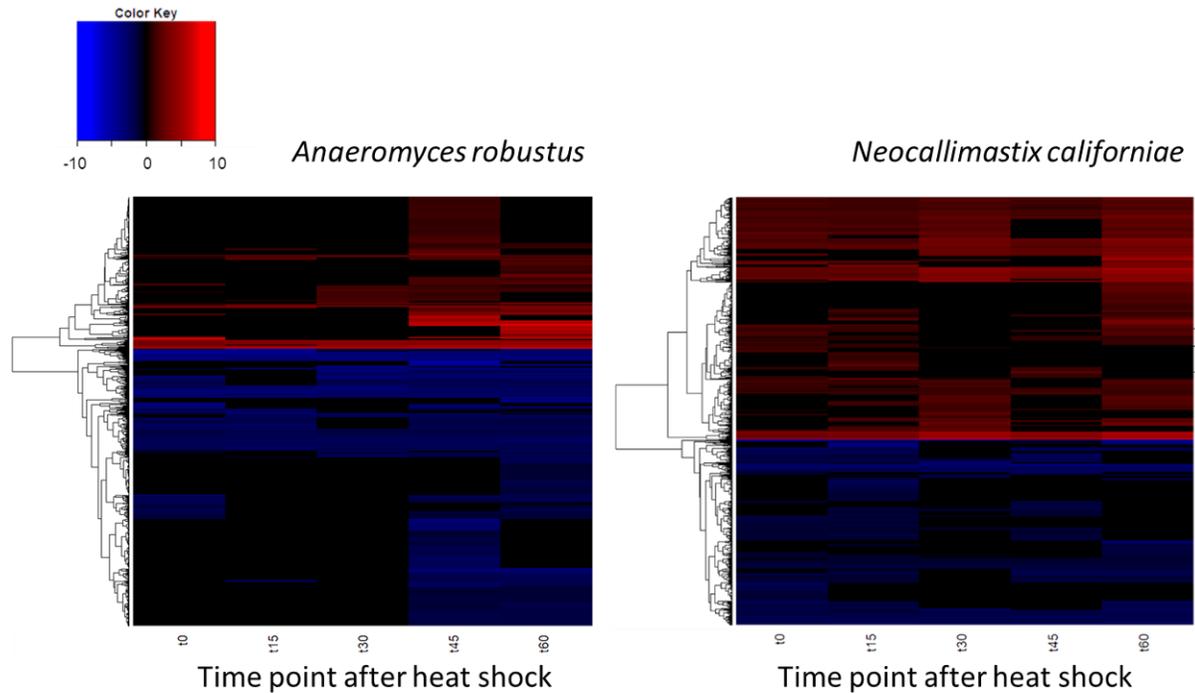
Figure 5.5: Dynamics of the transcriptomic responses to heat shock differ between *A. robustus* and *N. californiae*. Each row represents the $\log_2$ fold change of a transcript relative to the control without heat shock. Only transcripts with absolute $\log_2$FC greater than or equal to one and adjusted p-value less than or equal to 0.05 are shown. X-axis labels for each heat map are as follows: "t0" represents time zero, immediately after the 15 minute heat shock, "t15" represents 15 minutes after t0 time point, and so forth up to 60 minutes after the completion of the heat shock ("t60"). Plots were rendered using the 'heatmap.2' function in the 'ggplots' package of R software version 3.4.3[270].

In both *A. robustus* and *N. californiae* analysis of the eukaryotic Orthologous Groups (KOGs)[212] revealed that the proportion of differentially regulated genes was highest in the group Cellular Processes and Signaling (Figure 5.6). Posttranslational modification, protein turnover, chaperones was the class with the most differentially expressed genes one hour after heat shock compared to before heat shock. The majority of the upregulated genes within this KOG class were chaperones belonging to the HSP20 or HSP70 families. For *A. robustus* 7% of the total genes classified in this KOG class were at least two-fold upregulated and for *N. californiae* 16% were upregulated (*p*-adjusted<0.05). This finding supports that HSR was activated in *A. robustus* and *N. californiae* upon exposure to thermal

stress at 48 °C. Further supporting that a stress response was activated, *A. robustus* upregulated 12 of its 15 genes encoding glutamate dehydrogenase (E.C. 1.4.1.4) at least two-fold (*p*-adjusted<0.05) one hour after heat shock. In *Saccharomyces cerevisiae*, glutamate dehydrogenase has been linked to the phenotype of resistance to thermal and oxidative stress-induced apoptosis[323]. The upregulation of multiple copies of glutamate dehydrogenase implies that glutamate dehydrogenase may perform a similar function in *A. robustus*.
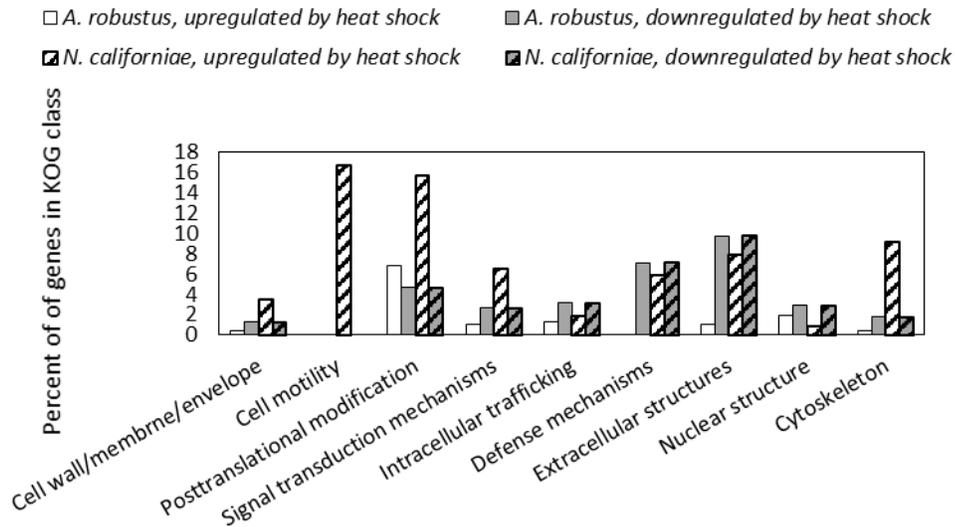
Figure 5.6: Differential regulation in response to heat shock of *A. robustus* and *N. californiae* genes in each class belonging to the eukaryotic Orthologous Group[212] (KOG) Cellular Processes and Signaling. Genes were counted as differentially regulated only for those with an absolute $\log_2 FC \geq 1$ (*p*-adjusted$\leq$0.05) one hour after completion of heat shock compared to the control without heat shock. Data labels for total regulated genes are the total number of gene models in the KOG class, excluding CAZymes. MycoCosm[23] KOG assignments were used for all gene annotations.

To verify whether ER stress was also induced, we searched for differentially regulated genes that indicated changes in protein folding and an increased burden of misfolded proteins. Previous work in model ascomycetes[282] indicated that genes with functions in protein traffic and secretion are upregulated in response to denaturants such as dithiothreitol (DTT) and tunicamycin, which are frequently used to elicit UPR[282,284]. However, HSR also targets genes related to the secretory pathway and can relieve ER

stress[284]. We found that both *N. californiae* and *A. robustus* significantly upregulated genes within the KOG class Intracellular trafficking, secretion, and vesicular transport (Tables 5.2 and 5.3). Retrograde transport from the Golgi to the ER is particular was upregulated, as evidenced by the upregulation of genes encoding putative COPI subunit proteins (Table 5.3). Consistent with the accumulation of misfolded proteins associated with ER stress, both *N. californiae* and *A. robustus* upregulated genes involved in the protein degradation, such as ubiquitination enzymes and proteasomes (Tables 5.2 and 5.3). Other key players of the secretory pathway that are implicated in the UPR include ER oxidoreductin (ERO1) and the hsp70 chaperone KAR2[324]. One of the *N. californiae* homologs of KAR2, MycoCosm[23] protein Id 377732, was upregulated 60 minutes after heat shock by a $\log_2$ fold change of 1.47 relative to the control without heat shock (*p*-adjusted<0.05).

Table 5.2: *N. californiae* **genes upregulated by heat shock indicative of UPR/HSR.** Log_2 fold change of the gene expression at 60 minutes after completion of heat shock relative to control with heat shock. Adjusted *p*-value less than or equal to 0.05. Annotations refer to the eukaryotic Orthologous Groups (KOGs)[212], InterPro[181], or Pfam[180] and are available from the MycoCosm portal[23].

ProteinId	Log2FC	KOG Annotation
Intracellular trafficking, secretion and vesicular transport		
448600	2.41	Lectin VIP36, involved in the transport of glycoproteins carrying high mannose-type glycans
666096	2.10	Annexin
702979	2.37	ER-Golgi vesicle-tethering protein p115
520343	4.57	Synaptic vesicle protein EHS-1 and related EH domain proteins
678035	2.33	GTP-binding protein SEC4, small G protein superfamily, and related Ras family GTP-binding proteins
450301	2.17	Endocytosis/signaling protein EHD1
462258	2.03	Endocytosis/signaling protein EHD1
459874	2.00	GTPase Rab11/YPT3, small G protein superfamily
321299	1.77	GTP-binding ADP-ribosylation factor Arf6 (dArf3)
698286	1.51	Protein involved in membrane traffic (YOP1/TB2/DP1/HVA22 family)
636991	1.48	ER-Golgi vesicle-tethering protein p115
696904	2.23	ER-Golgi vesicle-tethering protein p115
673754	2.09	Vesicle coat complex AP-3, beta subunit
416162	2.04	ER-Golgi vesicle-tethering protein p115
672010	1.80	ER-Golgi vesicle-tethering protein p115
667258	1.74	Annexin
Protein turnover		
678702	3.54	Subtilisin-related protease/Vacuolar protease B
702757	2.33	Ubiquitin-conjugating enzyme (PF00179)
702758	1.93	Ubiquitin-like proteins
700100	1.92	Ubiquitin-like proteins
454694	1.81	Ubiquitin and ubiquitin-like proteins
701610	1.73	Subtilisin-related protease/Vacuolar protease B
381558	1.48	Ubiquitin and ubiquitin-like proteins
461227	1.10	Ubiquitin-conjugating enzyme (PF00179)
507800	1.01	Subtilisin-related protease/Vacuolar protease B

Table 5.3. *A. robustus* genes upregulated by heat shock indicative of UPR/HSR. Log_2 fold change of the gene expression at 45 minutes after completion of heat shock relative to control with heat shock. Adjusted *p*-value less than or equal to 0.05. Annotations refer to the eukaryotic Orthologous Groups (KOGs)[212], InterPro[181], or Pfam[180]and are available from the MycoCosm portal[23]. Genes specifically discussed in the main text are emphasized in italics.

ProteinId	Log2FC	KOG Annotation
Intracellular trafficking, secretion and vesicular transport		
291613	5.52	Endocytosis/signaling protein EHD1
297960	2.06	Transport protein particle (TRAPP) complex subunit
294812	1.70	Nuclear transport receptor Karyopherin-beta2/Transportin (importin beta superfamily)
299217	1.68	Karyopherin (importin) alpha
249042	1.53	GTPase Rab6/YPT6/Ryh1, small G protein superfamily
328807	1.48	Sorting nexin SNX11
239866	1.42	Clathrin adaptor complex, small subunit
325124	1.37	GTP-binding ADP-ribosylation factor-like protein yARL3
261239	1.24	Golgi protein
325616	1.22	Prolactin regulatory element-binding protein/Protein transport protein SEC12p
289472	1.22	Nuclear transport receptor Karyopherin-beta2/Transportin (importin beta superfamily)
287286	1.18	Karyopherin (importin) beta 1
293071	1.17	*Vesicle coat complex COPI, beta' subunit*
271324	1.17	Vacuolar sorting protein VPS24
280100	1.14	GTPase Rab1/YPT1, small G protein superfamily, and related GTP-binding proteins
2608481	1.14	Septin family protein (P-loop GTPase)
271187	1.13	Septin family protein (P-loop GTPase)
292264	1.11	Annexin
211639	1.10	Synaptobrevin/VAMP-like protein SEC22
241413	1.04	SNARE protein YKT6, synaptobrevin/VAMP syperfamily
294106	1.04	**Vesicle coat complex COPI, gamma subunit**
276638	1.03	Vesicle coat complex AP-1/AP-2/AP-4, beta subunit
Protein turnover		
288193	1.76	26S proteasome regulatory complex, ATPase RPT3
225527	1.55	26S proteasome regulatory complex, ATPase RPT3
222683	1.30	Ubiquitin-specific protease UBP14
324815	1.42	20S proteasome, regulatory subunit beta type PSMB1/PRE7
292040	1.40	Ubiquitin and ubiquitin-like proteins
292798	1.33	Ubiquitin-like protein
197276	1.30	Ubiquitin activating enzyme UBA1
222683	1.30	Ubiquitin-specific protease UBP14
236994	1.24	Ubiquitin-conjugating enzyme (PF00179)
275087	1.21	20S proteasome, regulatory subunit beta type PSMB4/PRE4
285507	1.21	26S proteasome regulatory complex, subunit RPN9/PSMD13
264651	1.18	20S proteasome, regulatory subunit beta type PSMB7/PSMB10/PUP1
200113	1.16	26S proteasome regulatory complex, ATPase RPT5
112561	1.12	26S proteasome regulatory complex, subunit RPN7/PSMD6

ProteinId	Log2FC	KOG Annotation
269892	1.11	Ubiquitin-conjugating enzyme (PF00179)
228527	1.11	Ubiquitin-conjugating enzyme E2
267208	1.09	Tripeptidyl peptidase II
329084	1.08	20S proteasome, regulatory subunit alpha type PSMA2/PRE8
242802	1.02	20S proteasome, regulatory subunit alpha type PSMA7/PRE6
292679	1.01	Ubiquitin-conjugating enzyme (IPR015368)
262301	1.00	20S proteasome, regulatory subunit beta type PSMB5/PSMB8/PRE2

A. robustus differentially regulated a total of 66 genes assigned to the KOG class Signal transduction mechanisms at one or more time points relative to the control without heat shock. Nine of the upregulated genes were annotated as Serine/threonine protein kinases (KOG1187). Since the protein kinase Hog1 in *S. cerevisiae* initiates the environmental stress response[277,325], we wondered whether any of these protein kinases were homologs. In *A. robustus*, protein Id 197439 was upregulated two-fold (*p*-adjusted≤0.05) at 30, 45, and 60 minutes after heat shock versus the no heat shock control. Protein BLAST[226] alignment of *A. robustus* protein Id 197439 from MycoCosm[23] to *S. cerevisiae* Hog1 (accession NP_013214) resulted in 48% identity and 95% coverage between the sequences (E-value 4e-108). These findings suggest that this gene may be evolutionarily related to the Hog1 of higher order fungi. However, there was no corresponding homolog in *N. californiae* that was upregulated in response to heat shock, although protein Ids 424919, 409946, and 523809 were identified as top hits when *A. robustus* 197439 was searched against all filtered model proteins of *N. californiae* using BLAST+[158].

Within the group Information Storage and Processing, *A. robustus* upregulated 6% of genes in the class Chromatin structure and dynamics. N6-adenine methylation has been shown to mark transcriptionally active genes in early-diverging fungi[185]. In *A. robustus* the gene encoding protein Id 282395, containing an N6-adenine-specific DNA methylase

domain (IPR002052) was upregulated by more than 32-fold 60 minutes after heat shock compared to the control without heat shock (*p*-adjusted<0.05). This finding suggests that *A. robustus* may epigenetically regulate gene expression in response to heat shock.

5.2.3   Neocallimastigomycetes harbor a disproportionate number of small chaperones among fungi

The majority of the upregulated genes encoding heat shock proteins in the response of *A. robustus* and *N. californiae* to heat shock were predicted to be small (~20 kDa. Although many small hsps are sequence divergent, especially between different organisms[326], the upregulation of these putative small hsps in response to heat shock corroborates the sequence-based prediction. Upon inspection of the genomes of Neocallimastigomycetes, we found that the number of small hsps per genome is on the same scale as some plants (greater than 20 small hsps)[327]. Since the average genome sizes of Neocallimastigomycetes are at least an order of magnitude smaller than plant genomes[7,327,328], this indicates that Neocallimastigomycetes genomes are relatively enriched in small heat shock proteins. Similarly, hsps are also enriched on a gene count basis, since the number of gene models for *A. robustus*, *C. churrovis*, and *P. finnis* is less than the model plant *Arabidopsis thaliana*[329]. This strong preference for small chaperones is not observed in most cother fungi, as evidenced in Figure 5.7. c Wallemiomycetes are a notable exception in the higher-order fungi that also encode a greater ratio of small to large hsps. Since *Wallemia* are halophiles[330,331], this may indicate that small hsps confer a selective advantage in this environment. By sampling all published genera available from the MycoCosm portal[23] for each clade on the fungal evolutionary tree, we observed that Neocallimastigomycetes have the highest percentage (0.58%) of hsps (size 70 kDa or 20 kDa) out of all predicted genes, as

well as six times the number of predicted hsps belonging to the 20 kDa class compared to the number of genes within the 70 kDa class. *A. robustus* and *N. californiae* upregulated more HSP20 chaperones relative to any other hsp class (Figure 5.8), in line with the overrepresentation of small chaperones in their genomes. The proportion of hsps in each class out of all upregulated hsps was similar between *A. robustus* and *N. californiae*, although the percent utilization of each class differed. For example, *A. robustus* upregulated 46% of the total HSP20 genes and *N. californiae* upregulated 90% of all HSP20 genes. The remainder of the HSP20 genes are likely upregulated by other environmental stressors, which may include pH stress, osmotic stress, or others.
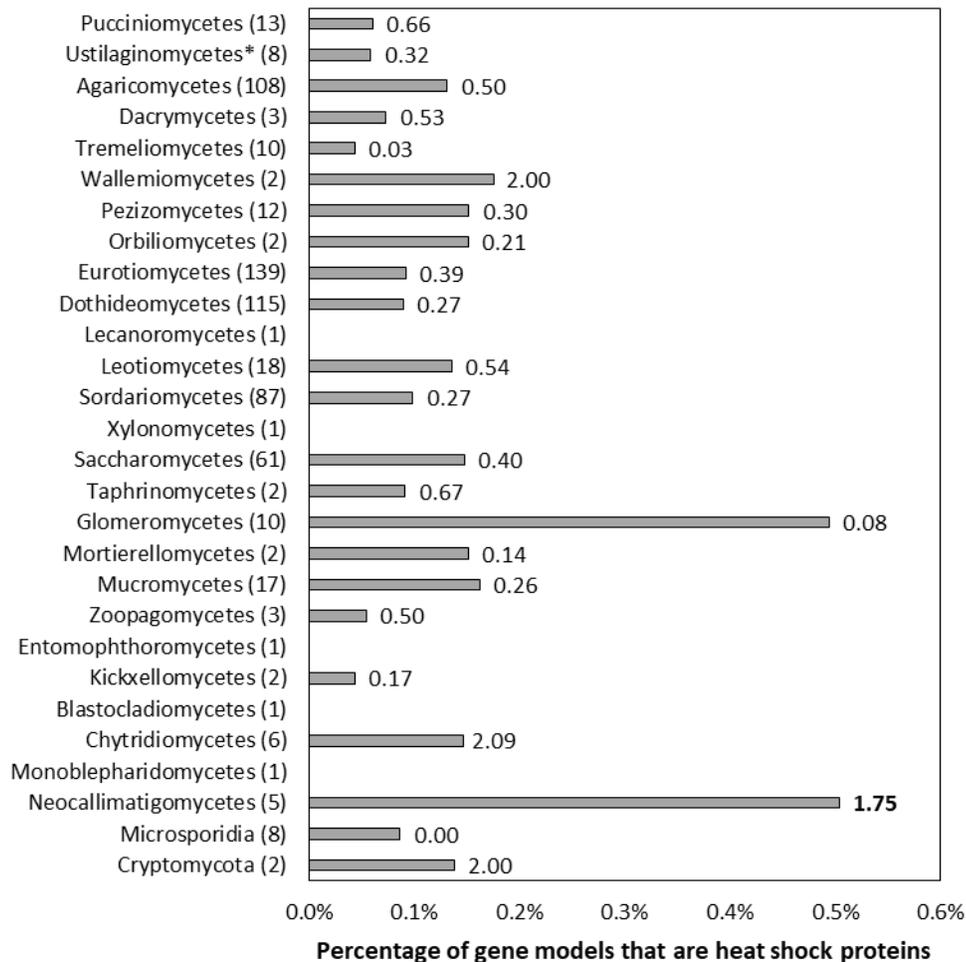
Figure 5.7: The genomes of Neocallimastigomycetes are enriched in heat shock proteins (hsps) and favor small hsps over large hsps compared to other fungi. Percentages given are the number of genes annotated as KOG0710 (HSP26/HSP42) or KOG0101 (HSP70/HSC70,HSP70 superfamily) out of the total number of gene models for each genome, averaged over all published genomes from that clade available from the MycoCosm portal[23]. Each data label is the ratio of small hsps to large hsps, calculated from the number of gene models assigned to KOG0710 or KOG0101 within each genome. Ratios were also averaged across all published genomes within a clade. All clades were significantly difference in hsp percentage relative to the Neocallimastigomycetes, except for Glomeromycetes, as assessed by a two-tailed, two-sample unequal variance student's t-test (alpha level < 0.1). Numbers in parenthesis after each clade name indicate the number of genomes represented from that clade. Clades are listed in order of divergence from the earliest common ancestor.

125
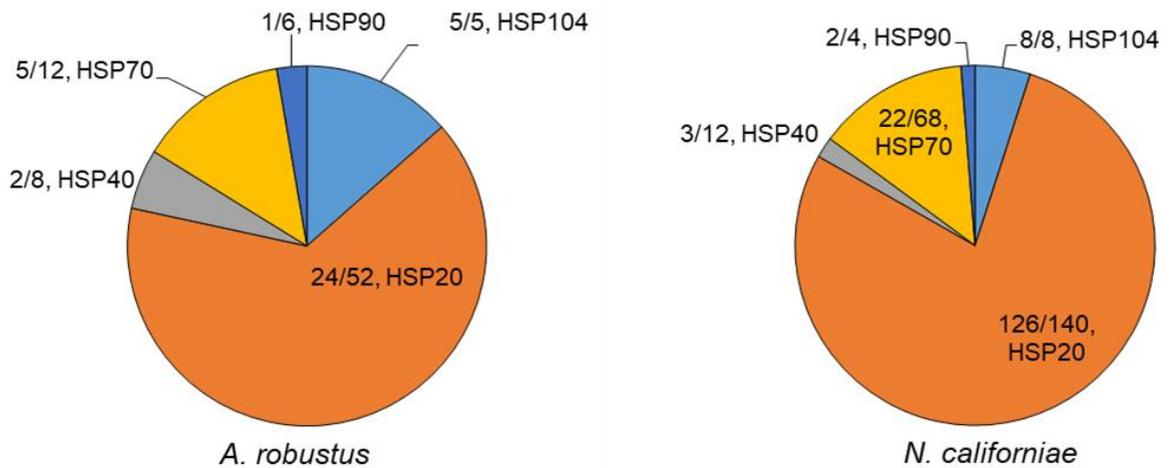
Figure 5.8: *A. robustus* **and** *N. californiae* **upregulate a greater proportion of HSP20 chaperones relative to all other hsps.** Pie charts depict the proportion of upregulated genes of each type out of the total number of upregulated hsps (absolute $\log_2FC\geq1$, p-adjusted≤0.05) at 45 min (*A. robustus*) or 60 min (*N. californiae*) after the completion of the heat shock. The presented fractions are the number of upregulated hsps in each class divided by the total number of MycoCosm[23] gene models with the corresponding KOG annotation. KOG0710 (HSP20/HSP42) was further annotated by InterPro[181] as HSP20.

The genomes of rumen fungi are known to be AT-rich and subsequently plentiful in asparagine repeats[332]. Similarly, the malaria-causing parasite *Plasmodium falciparum* from the Apicomplexa phylum also has a genome rich in asparagine repeats, and the parasite compensates for the propensity of its proteins to agglomerate by using the heat shock protein 110 and likely other chaperones to stabilization its proteome[333,334]. It is possible that rumen fungi may use their hsps to stabilize their asparagine-rich proteomes, similar to *Plasmodium falciparum*. Furthermore, the majority of the small hsps of the Neocallimastigomycetes were constitutively transcribed at greater than 0.5 RPKM during normal laboratory cultivation (Table 5.4).

Table 5.4: Constitutive transcription of small hsps during standard laboratory cultivation. Transcriptomes were previously acquired[6] from a variety of cultivation conditions including growth on grass and soluble sugars**.** The transcriptomes were filtered for small hsps annotated as "hsp20" by InterProScan[335] Number of genes encoding small hsps were estimated by the number of genes annotated as KOG0710 in MycoCosm[23]

Fungus	Small hsp gene count	Median RPKM	Count of transcripts > 0.5 RPKM
Anaeromyces robustus	52	10.22	49
Neocallimastix californiae	140	2.68	65
Piromyces finnis	24	6.19	22

5.3   *Conclusions*

`   By studying Neocallimastigomycetes, as one of the earliest-diverging fungi on the fungal evolutionary tree, we gain important insight into the evolution of broadly conserved cellular responses to stress, such as the HSR and UPR. In this work, we have demonstrated that in the genomes from four genera of Neocallimastigomycetes (*Anaeromyces*, *Caecomyces*, *Neocallimastix*, and *Piromyces*), a putative PERK-like protein is encoded. These predicted transmembrane proteins are highly conserved between genera, and are homologous in the catalytic kinase domain to metazoan PERK proteins. Transmembrane PERK-like proteins been previously identified in apicomplexan parasites, where they respond to different forms of stress, including ER stress, and regulate developmental life stages. The PERK-like proteins from Neocallimastigomycetes were homologous only in the kinase domain to apicomplexan transmembrane kinases. Therefore, early eukaryotes within Apicomplexa and Chytridiomycota appear to have in common PERK-like transmembrane kinases with distinct regulatory regions from each other. The function of these kinases in Neocallimastigomycetes is unknown and requires experimental validation. We also

demonstrated that the genomes of Neocallimastigomycetes are enriched in heat shock proteins (hsps), especially small hsps, compared to all other fungi. The majority of the small hsps are constitutively transcribed during standard laboratory cultivation. These constitutively transcribed hsps may serve to stabilize an asparagine-rich genome, which has been demonstrated to be the case in *Plasmodium falciparum*. These findings establish that the stress response of Neocallimastigomycetes is unique compared with other fungi, with surprising commonalities with Protozoa and Metazoa.

5.4   Materials and Methods

5.4.1   Annotation of genes in the Neocallimastigomycete unfolded protein response

Sequences of components of the UPR in model organisms *Saccharomyces cerevisiae* and *Drosophila melanogaster* were queried against the filtered model proteins of *Anaeromyces robustus, Caecomyces churrovis, Neocallimastix californiae*, and *Piromyces finnis* using protein BLAST[226] within the MycoCosm portal[23]. Query protein sequences were as follows (accession numbers are given in parenthesis): KAR2 (NP_012500), IRE1 (NP_011946.1), PERK (NP_649538), eIF2alpha (NP_001285329), and, ERO1 (NP_013576), CNE1 (NP_009343), and CALR (NP_001262430). Candidate PERK proteins were also identified by searching the MycoCosm portal[23] for filterered model proteins annotated as KOG1033 eIF-2alpha kinase PEK/EIF2AK3. Candidate PERK gene models were checked for RNA coverage in the MycoCosm genomebrowser, using previously published RNA-seq data[6]. *A. robustus* 182287 was selected instead of protein Id 13827 because the associated gene model had better RNA-seq coverage. Each protein sequence was queried for transmembrane helices by TMHMM2.0[225] and for signal peptides by SignalP-

5.0[312]. Only proteins with transmembrane helices were considered as PERK candidates. PAS and kinase domains were established for each PERK proteins using CD-Search[182].

5.4.2   Protein sequence alignment and visualization of putative PERK homologs

Alignment of protein sequences was performed using Clustal Omega[336] with default parameters in the MEGA[337] interface. Protein alignments were visualized using Jalview[314] with Clustalx coloring and sorted by pairwise identity. The Hanging sequences with no conservation, such as was the case for *T. gondii* TgIF2K-A, were removed to the left and right of the alignment in Jalview. Accession numbers for PERK proteins used in alignments were as follows, source organism given in parenthesis: AAS48463 (*Toxoplasma gondii* eIF2α kinases A, also called TgIF2K-A), XP_011239504 (*Mus musculus*), XP_024328881.1 (*Plasmodium falciparum*), XP_953607.1 (*Theileria annulata*), NP_001262283 (*Drosophila melanogaster*), and NP_001300844 (*Homo sapiens*). MycoCosm[23] protein Ids for PERK candidates from Neocallimastigomycetes are presented in Table 5.1.

5.4.3   Construction of gene phylogenies for transmembrane kinases from
Neocallimastigomycetes and Apicomplexa

The following sequences were searched via MMseqs2[338] against the NCBI non-redundant protein database, MycoCosm[23], and MMETPS[162]: MycoCosm protein Ids 13827 (*Anaeromyces robustus*), 543476 (*Caecomyces churrovis*), 503500 (*Neocallimastix californiae*), and 367815 (*Piromyces finnis*), and NCBI accession numbers XP_024328881.1 (*Plasmodium falciparum*), XP_953607.1 (*Theileria annulata*)*,* AAS48463.1 (*Toxoplasma gondii*), and ACA62938.1 (*Toxoplasma gondii*). For searches of Neocallimastigomycete sequences, the class Neocallimastigomycetes was excluded from the results. Phylogenetic

trees were constructed by FastTree and RAxML from these sequences and their top 100 highest-scoring hits.

5.4.4   Survey of heat shock proteins from the fungal tree of life

All published genomes in the MycoCosm portal[23] were searched for gene models with the annotation KOG0710 Molecular chaperone (small heat-shock protein Hsp26/Hsp42) and KOG0101 Molecular chaperones HSP70/HSC70, HSP70 superfamily. The ratio of the count of gene models annotated as KOG0710 to gene models annotated as KOG0101 was calculated for each genome and averaged over all genomes within each clade (class or phylum). A two-tailed, two-sample unequal variance student's t-test (alpha level < 0.1) was used to assess significant differences in average hsp ratios between each clade and the Neocallimastigomycetes. The percentage of hsps for each genome was calculated by dividing the sum of all gene models belonging to KOG0101 or KOG0710 by the total number of gene models.

5.4.5   Routine cultivation of Neocallimastix caliorniae and Anaeromyces robustus

The anaerobic fungal strains *Neocallimastix californiae* and *Anaeromyces robustus* were isolated via reed canary grass enrichment from the fecal matter collected from two ruminants at the Santa Barbara Zoo. *N. californiae* originates from a goat; *A. robustus* originates from a sheep. The isolates were separately grown at 39 °C under anaerobic conditions in Hungate tubes containing 9.0 mL of autoclaved complex media ("MC") with 0.1 g of milled reed canary grass as the substrate and 100% CO_2 in the headspace. The complex media contains 2.5 g/L yeast extract, 6.0 g/L sodium bicarbonate, 10 g/L Bacto™ Casitone, and 15.0 vol% clarified rumen fluid. The fungal strains achieved mid-log phase of growth every 3-4 days

and were aseptically transferred at this time point into fresh complex media with 0.1 g of milled reed canary grass as the substrate. Pressure accumulation in the headspace due to the production of fermentation gases was used as a proxy to quantify and track fungal growth.

5.4.6 Heat-shock procedure

1.0 mL of either *N. californiae* and *A. robustus* was inoculated by sterile syringe into 0.1 g of reed canary grass substrate and 10 mL of complex media[262] ("MC") in each Hungate tube with 100% CO_2 headspace and grown anaerobically at 39 °C for 48 hours. After the growth period, a total of 24 replicates of each species were subjected to a 48 °C water bath for 15 minutes. The fungi were harvested in replicates of four at 15 minute intervals starting immediately before heat shock (control group) up to 60 minutes after the completion of the heat shock. Standard good practices for working with RNA were followed during all steps. The cultures were transferred from the Hungate tubes to 15 mL Falcon tubes at the time of harvest. They were then centrifuged for 7 minutes at 4 °C and 3220 *g* in a swinging bucket rotor (Eppendorf™ A-4-81). 1 mL of RNAlater® (Sigma-Aldrich®) was added to each of the pellets using sterile, filter pipette tips. The samples were then vortexed for 5 seconds to thoroughly mix the pellet and stabilization solution. The Falcon tubes containing the pellets with RNAlater®, were stored at -20°C until extraction.

5.4.7 RNA extraction

For each fungal strain, total RNA from a randomly selected sample from each of four time points was extracted manually to first ensure the presence of high quality RNA. The remainder of the samples were subsequently submitted to automated extraction via a QIAcube (QIAGEN). The frozen cell pellets were thawed from storage on ice and then

centrifuged for 10 minutes at 4°C and 3220 *g*. The RNAlater™ was decanted from each replicate and the remaining pellets were transferred to previously autoclaved 2-mL screw-cap tubes (Fisher Scientific) containing 1 mL of 0.5 mm zirconia beads (Biospec). To each tube, 450 µL (manual extraction) or 600 µL (QIAcube) of a mixture of buffer RLT (QIAGEN) and 14.3 M β-mercaptoethanol (Sigma) in a ratio of 1 mL to 10 uL respectively, was added. The cells were lysed using the Biospec Mini-Beadbeater-16 for 1 minute, placed briefly on ice, and then centrifuged using a microcentrifuge (Eppendorf™ 5424) for 3 minutes at room temperature and 13000 *g*. The lysate was removed using gel-loading pipette tips (Fisher Scientific) and deposited in either round-bottom tubes for total RNA extraction (QIAcube exraction) or QIAshredder tubes (manual extraction). QIAcube extraction was executed following the RNeasy Mini protocol. Manual extraction was completed according to the protocol for "Purification of Total RNA from Plant Cells and Tissues and Filamentous Fungi" outlined in the RNeasy® Mini Handbook. The optional on-column DNAse digest was included in both methods.

5.4.8   *RNA quality assessment, library preparation, sequencing, and data analysis pipeline*

RNA quality was assessed for the two critical metrics that dictate successful sequencing; concentration by Qubit 2.0 Fluorometer (Invitrogen) and degradation by Agilent TapeStation. All samples exhibited a starting amount of total RNA above the minimum threshold for the sequencing protocol (200 ng) and an RNA Integrity Number (RINe) above 7.0. The Illumina® Truseq® Stranded mRNA kit was used to prepare the mRNA libraries for the *N. californiae* and *A. robustus* samples as it isolates eukaryotic polyadenylated mRNA. The resulting libraries were sequenced into 75 bp single-end reads employing a high output kit to generate more than 400 million reads on an Illumina® NextSeq500. HISAT[220] was

used to align the reads of each species to their respective genomes, which are publically available for download from the Joint Genome Institute (JGI) MycoCosm portal[23]. After mapping, featureCounts[221] was used to quantify the number of reads mapped to distinct genes for each of the two fungal strains. Subsequently, the DESeq2[222] package in R version 3.4.3[270] was used to test for differential gene expression between the control and each of the time points following heat shock for *N. californiae* and *A. robustus* respectively. Genes were classified as differentially regulated if the requirements of an absolute $\log_2$ fold change greater than or equal to 1 and a *p*-adjusted value less than 0.05 were met. The resulting dataset was then analyzed using the functional annotations from KEGG[191,] GO, InterPro[181] and KOG[212] publically available via MycoCosm portal[23].

5.4.9   *Data availability*

Protein phylogenies for transmembrane kinases from Apicomplexa and Neocallimastigomycetes are available at the following GitHub repository: https://github.com/cswift3/stress_response_anaerobic_fungi

FASTQ files for all samples sequenced as part of this work available through the National Center for Biotechnology Information (NCBI) BioProject PRJNA665745 at https://dataview.ncbi.nlm.nih.gov/object/PRJNA665745?archive=sra.

6 Fungal and bacterial secondary metabolites shape consortia membership in an enriched goat fecal microbiome

The prior basis for this work is discussed in Peng et al. (*Nature Microbiology,* in press 2020).

6.1 Introduction

Little is known about the role of eukaryotic and prokaryotic secondary metabolites in modulating the dynamics of the microbial communities within the digestive tracts of herbivores. Historically, it was thought unusual for obligate anaerobes to encode biosynthetic gene clusters (BGCs)[339]. However, recent literature[30,53,340] suggests that the biosynthetic potential of anaerobic bacteria is more extensive than previously thought, with the genomes of members of the phylum Firmicutes encoding multiple gene clusters, including polyketide synthases (PKSs), nonribosomal peptide synthetases (NRPSs), and ribosomally synthesized and post-translationally modified peptides (RiPPs)[53]. Although plant natural products have been considered as feed additives for ruminants to improve animal nutrition[341], the role of microbial natural products on shaping the microbiome has not been extensively studied. Expanding our understanding of the role of microbial natural products within the rumen environment could benefit both human and animal health.

The rumen environment may be a valuable source of novel antimicrobials. Bacteriocins, or antimicrobial peptides produced by bacteria[342,343], have been more extensively characterized from rumen microbes compared to other classes of natural products. Butyrivibriocin OR79A was discovered from *Butyrivibrio fibrisolvens* OR79A[17,18] and butyrivibriocin AR10 from *Butyrivibrio fibrosolvens* AR10[17]. Notably, the peptide

Lynronne-1, which was identified from the rumen via a functional metagenomics and computation, decreased the bacterial count of methicillin-resistant *Staphylococcus aureus* (MRSA) in a murine model[19]. Rumen fungi may also synthesize natural products with antimicrobial activity, as suggested by a recent dual-transcriptomic study of a co-culture of rumen fungi and bacteria (Chapter four).

The interplay between human and animal health has been recognized in the spread of antibiotic resistance genes through agricultural practices[344–346]. Subtherapeutic doses of antibiotics are used in agriculture to prevent disease and increase growth, which may contribute to the rise of antibiotic resistance[344]. Engineering the microbiomes of ruminants to achieve the same benefits of disease prevention and increased growth, without increasing antibiotic resistance, is a challenging, but desirable outcome. Knowledge of the role of microbial natural products within the ruminant microbiome will inform engineering efforts to improve microbiome performance in regards to host health. Metagenomic sequencing of enrichment cultures from an inoculum is one possible approach of an overall design-build-test-learn strategy in microbiome engineering to identify microbial natural products[347].

While antibiotic resistance poses a threat to human health, antibiotic resistance genes in BGCs can help identify novel antibiotics[149,348,349]. Microbes producing antibiotics require one or more mechanisms to avoid suicide, such as producing self-resistant variants of the antibiotic target, modifying the antibiotic target, or exporting, inactivating, or sequestering the antibiotic itself, and these mechanisms are detectable by resistance genes[350]. For example, glycopeptide antibiotics inhibit cell wall synthesis by targeting peptidoglycan precursors, but resistance genes for the peptidoglycan precursors in the antibiotic-producer *Enterococcus faecium* BM4147 encode a single base substitution that reduces glycopeptide binding affinity by 1000-fold[351,352]. Computational tools have been developed to detect

confirmed resistance genes and potentially resistant homologues for core housekeeping genes clustered near biosynthetic genes in order to rank BGCs in bacteria[155,156,353] for novel antibiotic potential. By acquiring metagenomic datasets, novel genomes can be analyzed for promising drug targets, which facilitates the identification of antimicrobial compounds from microbes that may be not be cultivable in isolation.

There are numerous metagenomic datasets now available from the rumen microbiomes of several herbivores[31,67,354–356] (Peng et al., *Nature Microbiology*, in press 2020), particularly cows. These various studies have primarily focused on expanding collections of rumen anaerobes[31,354], the primary metabolism of the rumen microbial communities[67,356], or their biomass degradation via carbohydrate-active enzymes (CAZymes)[67,356]. However, few studies have focused on the BGCs found in microbes native to the rumen microbiome, and their role in shaping the microbial community, with some exceptions. Previously, 501 genomes from the Hungate 1000 collection were mined using antiSMASH 3.0, resulting in 6,906 predicted BGCs[357]. However, transcriptomic and proteomic expression of the BGCs from the Hungate 1000 collection was not directly analyzed.

Metatranscriptomics has proven to be a valuable tool with regards to understanding the dynamics of BGC expression in response to environmental stimuli[358]. Recently, this technique was applied to analyze the expression of nonribosomal peptide synthetase (NRPS) and polyketide synthase (PKS) genes from the rumen microbiome[340]. However, this study aligned reads from independent metatranscriptomic datasets to metagenome-assembled genomes (MAGs) from the reference Hungate 1000 collection[359] and NCBI (http://www.ncbi.nlm.nih.gov/genome/). The expression of BGCs from microbes that may be absent from this collection or database cannot be ascertained by this approach. Until now,

no one has assembled reference MAGs from rumen-derived microbial communities and subsequently sequenced the RNA of the same communities.

Although our knowledge of prokaryotic BGCs from the rumen is limited, even less is known about fungal BGCs. To date, no study has previously captured the expression of fungal BGCs in complex microbial communities derived from ruminants. Anaerobic gut fungi are also important members of the microbial rumen community, although they are outnumbered by bacteria in the rumen by orders of magnitude[5]. Previous work has shown that anaerobic gut fungi possess the biosynthetic machinery to synthesize natural products (Chapter two) and that they upregulate NRPS genes in the presence of the rumen bacteria *Fibrobacter* sp. UWB7 (Chapter four). Overall, this suggests that BGCs from anaerobic fungi could regulate microbial membership in the rumen micorbiome if these BGCs are transcriptionally active and making chemical products.

Here, we expand on a previous study where we enriched for different microbial communities from the goat fecal microbiome using a combination of antibiotic and substrate choices (Peng et al., *Nature Microbiology,* in press 2020). By cultivating these communities through 10 consecutive generations of batch passaging, we established how the microbial composition and gene expression of the communities shifted throughout the enrichment process. We generated both metatranscriptomic and metagenomic datasets during this enrichment process and then analyzed these datasets without reliance on exogenous culture collections or databases. Additionally, we assessed the transcription of biosynthetic genes from fungi within the antibiotics-treated communities. This approach allowed us to correlate the BGC expression profiles from both prokaryotes and fungi to the changing membership of microbial communities enriched from the goat fecal microbiome.

6.2   Results and Discussion

6.2.1   Genetic potential of the goat fecal microbiome for the synthesis of natural products

To assess the potential of the rumen-derived microbiome to synthesize natural products, we mined the genomes of 719 novel metagenome-assembled genomes (MAGs) obtained from sequencing of the source fecal pellet as well as the cultivated consortia (Peng et al., *Nature Microbiology*, in press 2020) using the <u>anti</u>biotics and <u>s</u>econdary <u>m</u>etabolites <u>analysis</u> <u>sh</u>ell (antiSMASH) version 4.0[107]. Prior to mining the genomes, contigs less than 10 kbp were removed to reduce the likelihood of BGCs being split between contigs, which would inflate the number of calculated gene clusters. In total, 688 biosynthetic gene clusters (BGCs) were predicted from the 719 MAGs. These BGCs were divided equally between MAGs recovered from the source fecal pellet or the cultivated consortia (Table 6.1). The overall proportion of BGCs in all MAGs with significant similarity to sequenced BGCs was 87% (Table 6.1), as evaluated by ClusterBlast[107], which queries a candidate BGC sequence against ~220,000 predicted BGC sequences within NCBI GenBank[360]. Only 12% of the BGCs from all MAGs returned a KnownClusterBlast significant hit (Table 6.1). KnownClusterBlast[107] queries clusters against known BGCs with experimentally characterized products in the Minimum Information about a Biosynthetic Gene Clusters (MIBiG) repository[134,361]. Product structures were predicted for 18% of the MAGs and monomers obtained from 27% of the MAGs, corresponding to the NRPS and NRPS hybrid BGCs. We noted that neither butyrivibriocin AR10 or OR79A were among the KnownClusterBlast hits, although both were present in the MIBiG database. Taken together, these findings suggest that the goat fecal microbiome is a source of undiscovered BGCs and may contain structurally novel natural products.

138

Table 6.1: Number of BGCs predicted from 719 unique MAGs from the goat fecal microbiome and their similarity to known BGCs. ClusterBlast queries clusters against clusters predicted from NCBI GenBank[360] and KnownClusterBlast queries the MIBiG repository[134,361]. Monomers and structures were predicted for NRPS or hybrids NRPS BGCs only.

MAGs	Number of BGCs	ClusterBlast significant hit, %	KnownClusterBlast significant hit, %	Structure prediction, %	Monomer prediction, %
Fecal Pellet	340	98	10	13	19
Cultivation*	348	76	13	22	35
Total	688	87	12	18	27

*Cultivation MAGs refer to those MAGs recovered from passaged consortia rather than the source microbiome (fecal pellet).

Out of 348 cultivated MAGs, 84 BGCs were identified without ClusterBlast or KnownClusterBlast significant hits (hereafter referred to as novel BGCs). The majority of these 84 novel BGCs were classified as NRPSs or hybrids of NRPSs (e.g. Type I PKS-NRPS), followed by Ribosomally synthesized and Post-translationally modified Peptides (RiPPs) and bacteriocins. Firmicutes encoded 81% of the novel BGCs. In contrast to the enrichment consortia, only seven of the BGCs out of all 345 identified from the pellet MAGs had no ClusterBlast or KnownClusterBlast significant hits. Furthermore, the fact that the majority of the novel BGCs belonged to MAGs classified as Firmicutes suggests that this phyla in particular is a rich source of biosynthetic potential, as has been suggested previously by Letzel and colleagues[53].

The ClusterBlast and KnownClusterBlast results from these MAGs are valuable as a means to estimate BGC redundancy and phylogenetic diversity at the genus level. The combined ClusterBlast results for the BGCs from both pellet and enrichment MAGs revealed 46 hits from the genus *Clostridium*, 22 from *Bacillus*, and 18 from *Streptococcus*. However, a total of 131 unique genera were represented among the ClusterBlast results. Therefore, while only 35 genera were identified from the 719 MAGs (Peng et al., *Nature Microbiology,* in press 2020), characterization of the BGCs implies a much greater diversity at the genus-level for the 719 MAGs. To estimate BGC redundancy, we calculated the

proportion of unique ClusterBlast and KnownClusterBlast significant hits for both pellet and enrichment MAGs. For the BGCs from the pellet MAGs, 36% of KnownClusterBlast hits and 48% of ClusterBlast top significant hits were unique. Similarly, 45% of ClusterBlast significant hits were unique and 59% of ClusterBlast hits were unique from the enrichment consortia. From this we estimate that roughly half of the 688 BGCs may be unique.

Although we have used the KnownClusterBlast and ClusterBlast results here as an estimate for the novelty and redundancy of BGCs, it is worth noting that even the BGCs with KnownClusterBlast significant hits may in fact be novel. We note that in some cases the secondary metabolite class for the putative BGC and the known BGC are not in agreement. For example, BGC ID c00041_NODE_17_c2 from the MAG ag0r3_bin.9 is classified as a bacteriocin, but the top significant KnownClusterBlast hit is the polyketide kijanimicin.

6.2.2 *Enrichment of bacteria from the Bacteroidetes and Firmicutes phyla changes the biosynthetic portfolio of the microbial community*

Firmicutes and Bacteroidetes were the most represented phyla among the MAGs assembled from the goat fecal microbiome (Peng et al., *Nature Microbiology*, in press 2020). These phyla were also dominant in the cultivated consortia. Therefore, we were particularly interested in biosynthetic potential of these phyla, and genomic content and transcriptomic expression of their BGC portfolio. We observed that the genomes of Bacteroidetes are far richer in biosynthetic gene clusters than previously thought. Based on the sequenced genomes available at present time, Letzel and co-workers suggested that the secondary metabolite potential of the phylum Bacteroidetes is limited[30,] with strains having no more than one or two PKS, NRPS, or hybrid genes. Analysis of the MAGs from the goat fecal pellet and subsequent cultivated consortia revealed that some bacteria belonging to this

phylum may also synthesize aryl polyenes, a class of pigmented secondary metabolites not previously reported in Fimicutes[362]. From the goat fecal pellet, 11 BGCs encoding predicted aryl polyenes were recovered and 10 of these were from distinct MAGs.

The overall suite of BGCs identified from the 719 MAGs (Figure 6.1) demonstrated that the source microbiome encoded a greater number of BGCs for each natural product class than the MAGs from the cultivated consortia, excluding antimicrobial peptides (AMPs), aryl polyenes, and NRPSs. Examination of the distribution of BGCs within these classes between phyla for the pellet-derived and cultivated MAGs revealed that the enrichment within these classes was due to the increased proportion of Firmicutes and Bacteroidetes (Figure 6.2). In consortia enriched without antibiotics treatment by growth on bagasse, ~10% of aryl polyene signature genes were transcribed in two or more biological replicates by at least 0.5 RPKM. By this same criterion, 7% of NRPS signature genes and ~20% of antimicrobial peptides (bacteriocins and microcins) were transcribed during cultivation. These data suggest that these natural product classes may contribute to the dominance of Firmicutes and Bacteroidetes during cultivation.
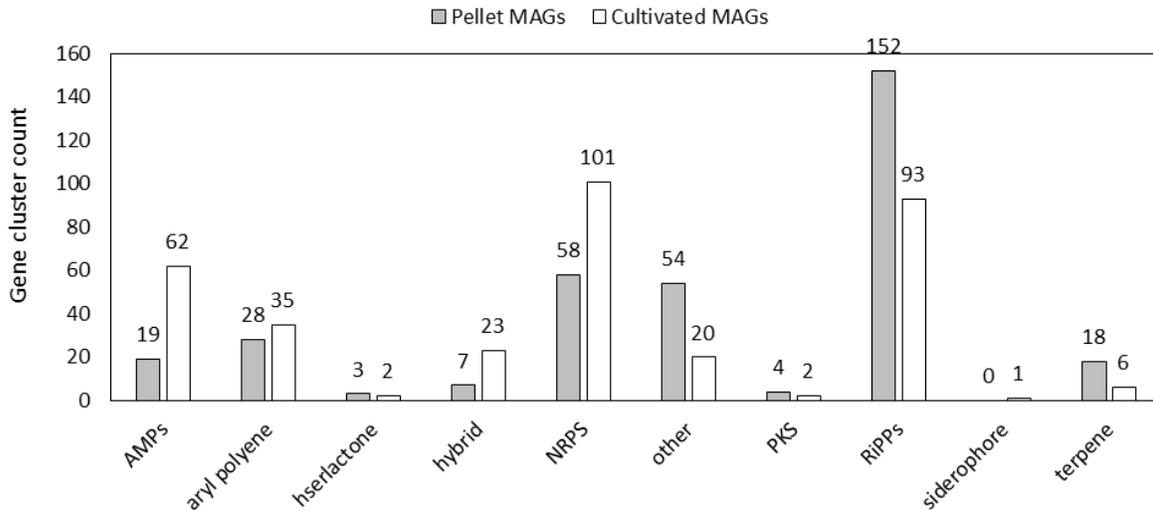
Figure 6.1: Gene cluster count for each natural product class encoded in MAGs recovered from source microbiome ("Pellet MAGs") or cultivation ("Cultivated MAGs"). AMPs=antimicrobial peptides (bacteriocins and microcinis); hserlactone=homoserine lactone; hybrids=combinations of two or more classes; NRPS=nonribosomal peptide synthetase; PKS=polyketide synthase; RiPPs=ribosomally synthesized and post-translationally modified peptides (lanthipeptides, lassopeptides, sactipeptides, thiopeptides, and thiopeptide-sactipeptides).
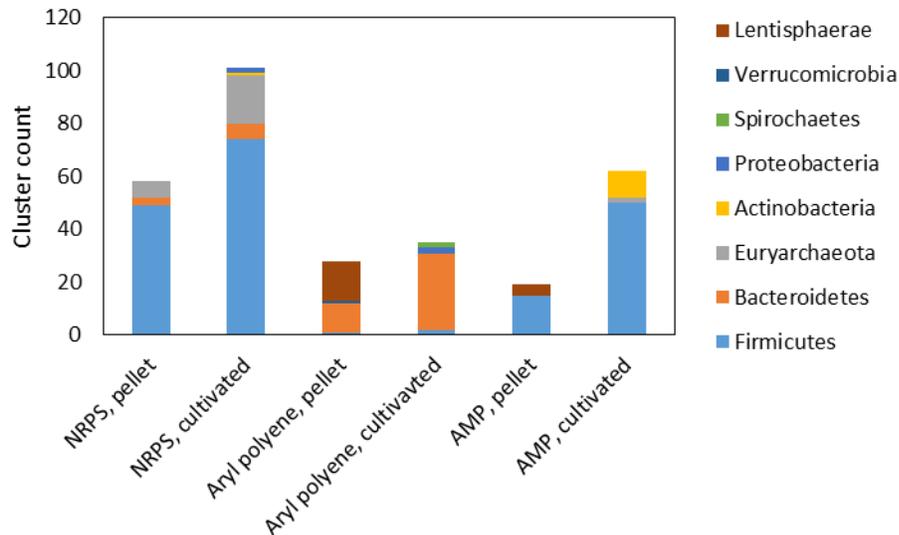
Figure 6.2: Enrichment of Bacteroidetes and Firmicutes changes the proportion of antimicrobial peptides, aryl polyenes, and NRPSs. Gene clusters are divided into those identified from source microbiome MAGs ("pellet") and those from the cultivated MAGs ("cultivated). NRPS=nonribosomal peptide synthetases; AMP=antimicrobial peptide (bacteriocin or microcin).

6.2.3 Biosynthetic genes are expressed, but not differentially regulated, through successive generations of batch passaging

To assess the activity of the BGCs, we sequenced the RNA of consortia from generations two, six and nine of batch passaging (see Methods). Signature genes of each gene cluster are those used in the antiSMASH profile Hidden Markov Model (pHMM)[11] to identify the cluster. In other words, they correspond to the core biosynthetic machinery of the natural product. Comparing the expression (in RPKM) of signature genes for antibiotics-free consortia grown on alfalfa, bagasse, and reed canary grass substrates (Figure 6.3), we ascertained that only a small proportion of the total signature genes was consistently expressed (6-15% depending on the substrate). We defined a transcribed gene as one in which the average expression (in RPKM) across at least two biological replicates was greater than 0.5. This definition accounted for genes that were transcribed consistently at low levels. We then checked whether the signature genes were differentially expressed between batch passage generations in the antibiotics-free bagasse consortia, for which we had three biological replicates each for generations 2, 6, and 9. None of the signature genes were differentially expressed (p-adjusted<0.05). The expression level of the signature genes ranged from 0.5 to 1000 RPKM (Figure 6.3). The constitutive high-level expression of a small portion of biosynthetic genes implies that their natural products may be of functional significance to the producer.
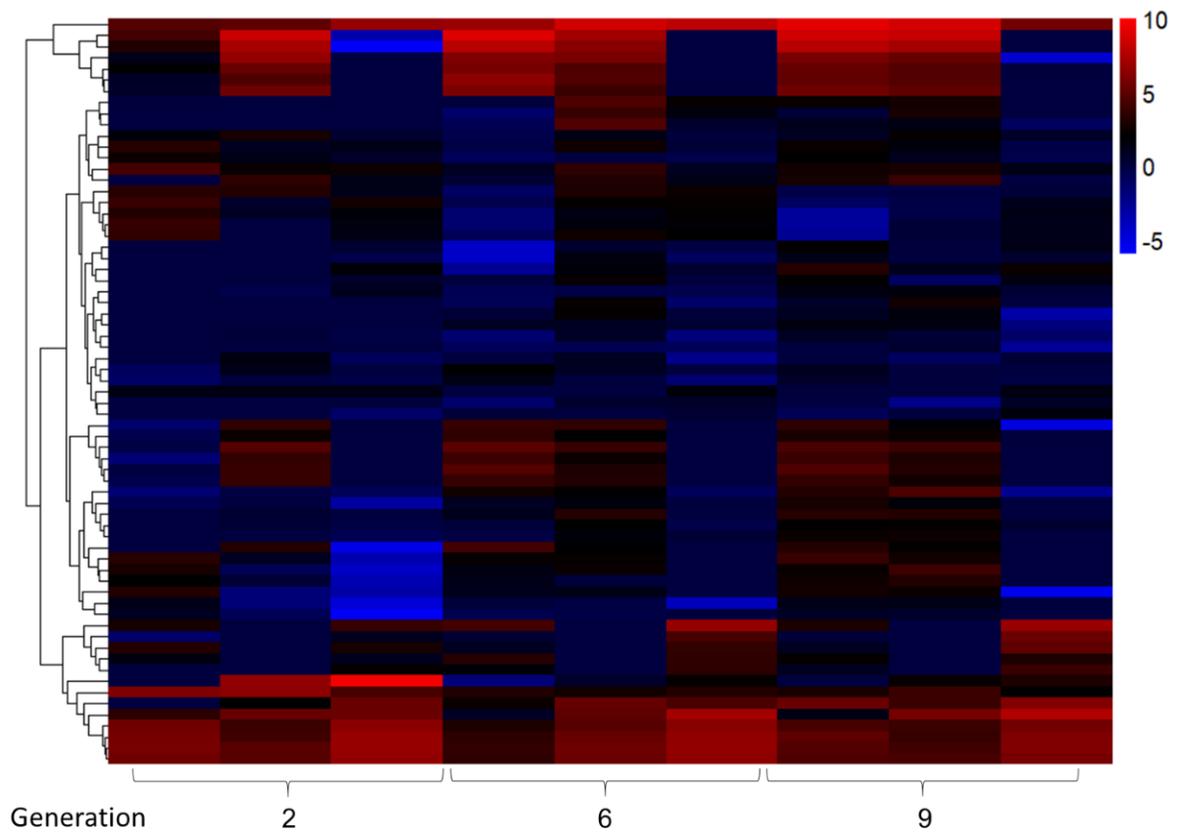
Figure 6.3: Biosynthetic genes are not differentially regulated between batch passage generations in antibiotics-free consortia enriched on bagasse. Log$_2$(RPKM) of signature genes from all MAGs with greater than 0.5 RPKM in 2 or more replicates of generation 9.

6.2.4 *The most abundant MAGs transcribe core biosynthetic genes during enrichment*

Relative abundance of the MAGs within generations zero, five, and ten were previously determined for all 719 MAGs (Peng et al., *Nature Microbiology,* in press 2020). Collectively, the genera of *Butyrivibrio* and *Pseudobutyrivibrio* (phylum Firmicutes) contained 38 BGCs, representing 5% of the total BGCs identified within the 719 MAGs. *Butyrivibrio* was the most abundant MAG within the alfalfa antibiotics-free consortia during the first half of the enrichment, until it was outcompeted by *Streptococcus* in later generations (Table 6.2). Similarly, *Pseudobutyrivibrio* dominated in the bagasse antibiotics-free consortia and was later outcompeted by *Lachonospiraceae* (Table 6.2).

144

Table 6.2. Most abundant and most active MAGs within each of the antibiotics-free consortia.
G=Batch passage generation (defined in Methods). The most active MAGs were determined from the largest RPKM and were measured across two or more replicates from generations 2, 6, and 9. For cases in which the most active MAG differed between biological replicates, all replicates are listed. MAG taxonomy is as defined as follows: Bac=*Bacteroidales,* But=*Butyrivibrio*, Lac=*Lachnospiraceae,* Rum=*Ruminococcus*, Sele=*Selemonas*, Str=*Streptococcus*. Numbers succeeding each taxonomy refer to Supplementary Table 3 of Peng et al. (*Nature Microbiology,* in press 2020).

Substrate Generation, Replicate	Most abundant MAG	Relative abundance, %	Most active MAG
Alfalfa			
G0, R3	But1	20.4	Rum1
G5, R3	But1	12.1	Str1
G10, R3	Str1	21.0	Str1
Bagasse			
G0, R2	Lac3	14.9	Rum2/Bac2/Rum1
G5, R2	Pse1	19.5	Bac2/Bac2/Rum1
G10, R2	Lac3	24.2	Bac2/Bac2/Rum2
Reed canary grass			
G0, R3	But1	15.4	Rum1/Bac2/Lac1
G5, R3	Str2	44.4	Str1/Str1/Lac1
G10, R3	But1	11.1	Bac2*
Xylan			
G0, R1	Sele1	21.5	Sele2/Pre3
G10, R1	Sele1	71.5	Sele1

*One replicate only.

The most abundant MAG in generation 5 of the bagasse antibiotics free consortia (Pse1) was rich in BGCs compared to the majority of the other MAGs. The average number of BGCs encoded per MAG was two, with a standard deviation of 1 BGC. The *Pseudobutyrivibrio* MAG Pse1 encoded 6 BGCs: 3 bacteriocins, 1 sactipeptide, 1 NRPS, and 1 Type 1 PKS-NRPS. The most abundant MAG in G0 and G5 (see Methods) of the alfalfa antibiotics-free consorta, *Butyrivibrio* MAG But1, encoded a thiopeptide-sactipeptide and thiopeptide-bacteriocin. Heatmaps of the relative expression of the signature genes in the clusters of the dominant MAGs (Figure 6.4) revealed that each MAG preferentially expressed one of its gene clusters: a bacteriocin (Pse1) or thiopeptide-sactipeptide (But1). The thiopeptide-sactipeptide was expressed highly compared to the other signature genes of the MAG in both the alfalfa and reed canary grass consortia. The expression of these gene

clusters during the enrichment suggests that antimicrobial peptides and RiPPs assisted the *Butyrivibrio* and *Pseudobutyrivibrio* genera in becoming the dominant microbes in these consortia. In contrast, the most abundant MAG within G0 and G10 of the bagasse antibiotics-free consortia, *Lachnospiraceae* MAG Lac3, had no predicted BGCs. Therefore, it is unlikely that secondary metabolism assisted this organism in become the dominant microbe in the community after 10 generations of batch passaging.

(A)

(B)

(C)
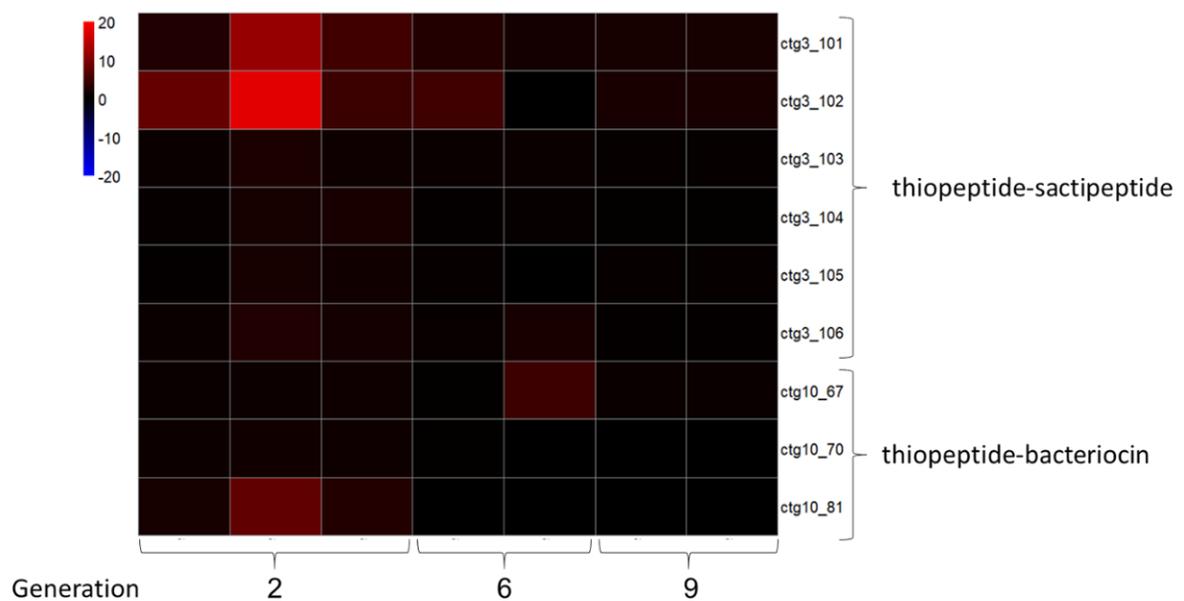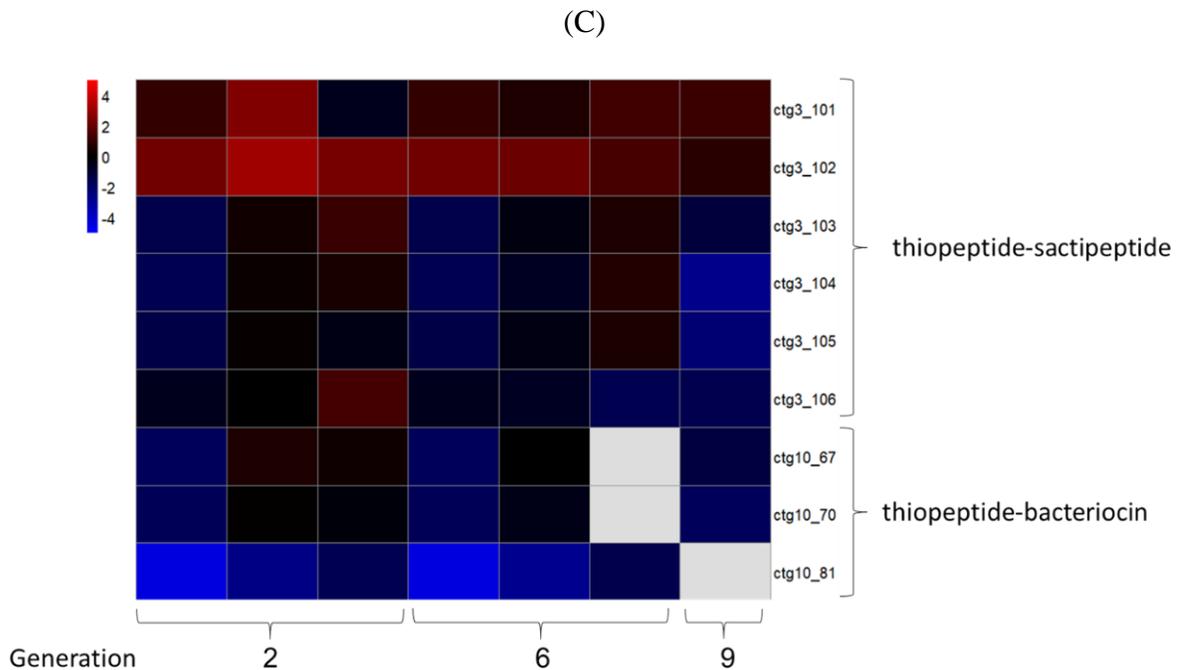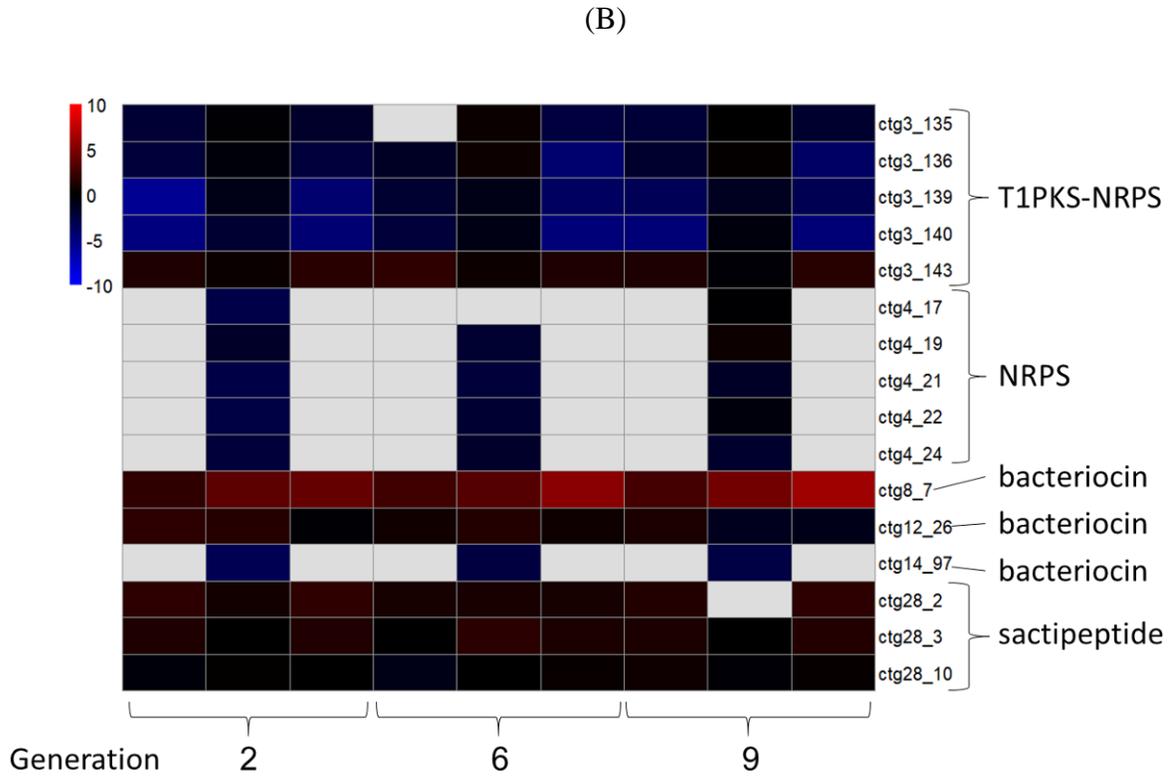
Figure 6.4: Relative expression of signature genes of BGCs from the most abundant MAGs within each substrate. Coloring of cells denotes the log2 fold change of each gene (RPKM) relative to the median expression of the MAG. (A) But1, alfalfa, (B) Pse1, bagasse, (C) But1, reed canary grass. See Table 6.2 for MAG taxonomy. Gray cells indicate no detectable gene expression.

147

The most active MAG, which was considered to be the MAG with the highest RPKM (million reads mapped to genome normalized by genome length in kb), differed in all cases except xylan generation 10 (replicate 1) from the most abundant MAG (Table 6.2). We quantified the expression of their BGCs relative to the median RPKM of all predicted genes (not limited to secondary metabolism). The most active MAG in at least two biological replicates for generations 6 and 9 of the bagasse antibiotics-free consortia was Bac2, belonging to the order *Bacteroidales*. Across generations 2, 6 and 9, two signature genes of an aryl polyene were transcribed 8-fold higher than the median expression of all genes (Figure 6.5). The *Streptococcus* MAG Str1 was the most active MAG the alfalfa antibiotics-free consortia during generations 6 and 9 (Table 6.2). Furthermore, it became the most abundant MAG in generation 10. In one replicate each of generations 2 and 6, Str1 transcribed its sole BGC, a bacteriocin, 32-fold higher than the median RPKM of all predicted genes. Together, these findings suggest that specific BGCs are implicated in the establishment of certain bacteria as abundant and/or active members of the consortia.
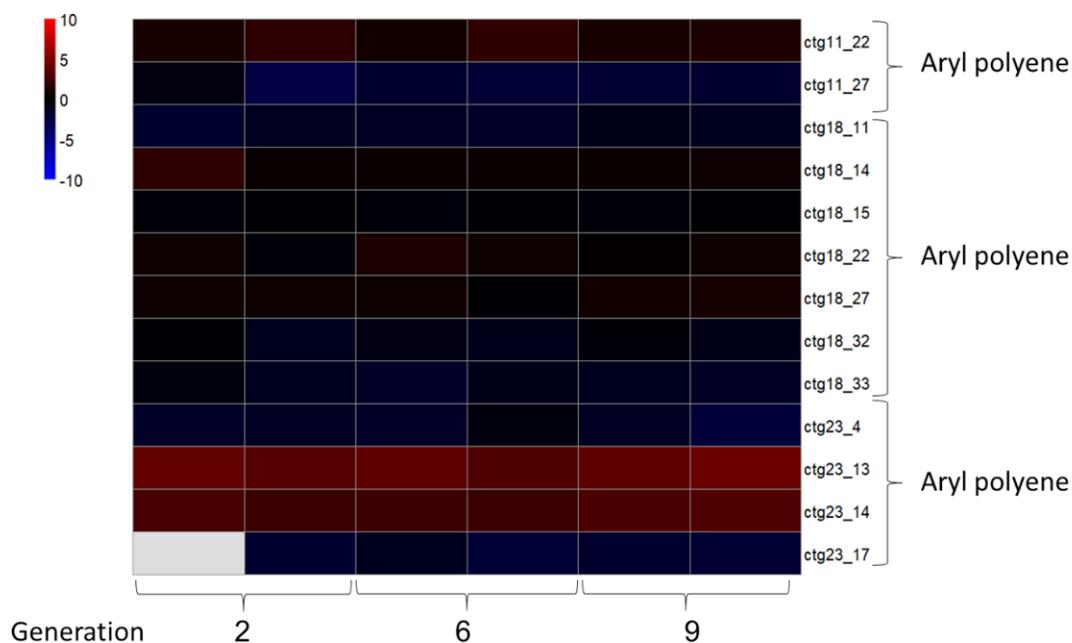
Figure 6.5: Relative expression of signature genes of BGCs from the most active MAG, Bac2, within bagasse antibiotics-free consortia. Coloring of cells denotes the $\log_2$ fold change of each gene (RPKM) relative to the median expression of all predicted genes of the MAG. Gray cell indicates no detectable gene expression.

6.2.5 *The biosynthetic gene clusters of anaerobic bacterial genomes contain putative self-resistance genes*

To assess the antibiotic potential of the BGCs from among the 719 prokaryotic MAGs, we used putative self-resistance genes identified by the Antibiotic Resistant Target Seeker (ARTS)[155,353] as indicators of bioactivity. We prioritized BGCs with a greater number of resistance genes proximal to core biosynthetic genes within the clusters. By this approach, 209 of the BGCs contained putative self-resistance genes, and 51 of the BGCs contained at least one resistance gene, with confirmed function as a self-resistance gene in other species, but putative function in the current work. Out of the 91 novel BGCs (no ClusterBlast or KnownClusterBlast significant hits), 28 had potentially resistant core genes and 7 had resistance genes confirmed in other bacteria. This subset of novel BGCs is most promising

for future antibiotic mining, and genes marked as domains of unknown function in the cluster may indicate novel modes of action.

The ARTS results show that BGCs from Firmicutes spp. comprise a large portion of the high-ranking BGCs for biosynthetic potential. 84% of the BGCs with resistance genes were from Firmicutes MAGs. 82% of the novel BGCs with potentially resistant core genes were from Firmicutes, further demonstrating the untapped antibiotic potential of the Firmicutes phylum. Furthermore, 17% of the inspected BGC ARTS results for cultivated MAGs were from the genus *Pseudobutyrivibrio* or *Butyrivibrio* (both belonging to the Firmicutes phylum). These data give further evidence that the enriched Firmicutes spp. could be producing secondary metabolites to inhibit other microbes. The BGCs including putative self-resistance genes of the most abundant *Butyrivibrio* sp. (But1) are depicted in Figure 6.6.
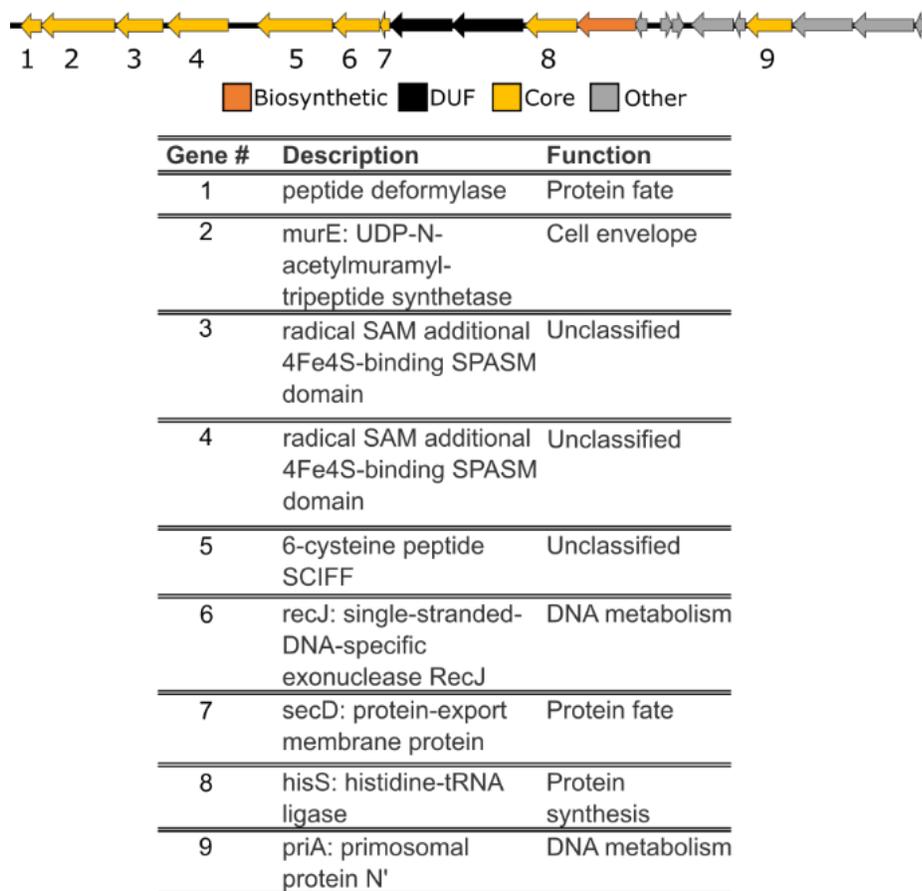
Figure 6.6: Thiopeptide-sactipeptide gene cluster from *Butyrivibrio* sp. contains putative self-resistance genes. Antibiotic Resistant Target Seeker (ARTS)[353] output summary is shown for the BGC with the highest relative expression from a *Butyrivibrio* sp. (But1), the most abundant MAG in generations 0 and 5 of the alfalfa antibiotics-free consortia. The thiopeptide-sactipeptide BGC is located on scaffold 3 at 101791-131137 bp in the genome. This figure was generated by Lazarina Butkovich.

6.2.6 *Neocallimastix californiae upregulate transcription of core biosynthetic genes*

related to secondary metabolism

In addition to the prokaryotic consortia, we also enriched for anaerobic fungi by antibiotics treatment (penicillin-streptomycin or chloramphenicol) on different carbon substrates. Although these communities were predominantly fungal, some bacteria with resistance to the antibiotics remained present in the consortia (Peng et al., *Nature Microbiology,* in press 2020). Based on the promising results from the synthetic co-cultures of the anaerobic gut fungi *Anaeromyces robustus* or *Caecomyces churrovis* with the rumen

bacteria *Fibrobacter* sp. UWB7 (Chapter four), we hypothesized that the fungal MAGs upregulate transcription of their BGCs during enrichment to compete with prokaryotic community members resistant to antibiotics treatment. *Neocallimastix californiae* was identified as the dominant eukaryotic member of the antibiotics-treated consortia (Peng et al., *Nature Microbiology,* in press 2020). From batch passage generation 2 to 9, 20 genes encoding core biosynthetic enzymes were differentially expressed (Table 6.3). Three-quarters of these genes were upregulated at least two-fold in generation 9 compared to generation 2. Notably, polyketide synthase and NRPS genes were among the fungal genes that were upregulated, consistent with observations of *A. robustus* in synthetic co-culture with *Fibrobacter* sp. UWB7 (Chapter four). The resistant bacteria in these fungal-dominated communities, which include members of the *Erysipelotrichaceae* and *Methanobacteriaceae* families. They also transcribed two-thirds of their signature genes in BGCs at levels greater than 10 RPKM in at least one biological replicateThese findings suggest that anaerobic fungi may utilize their secondary metabolism during the stress of enrichment, presumably to compete against or communicate with other microbes, as previously demonstrated for other fungi[143,363]. The simultaneous expression of prokaryotic BGCs supports microbial competition between *N. californiae* and prokaryotic community members resistant to penicillin-streptomycin treatment.

Table 6.3: Fungi in the penicillin/streptomycin-treated bagasse consortia upregulate transcription of biosynthetic and accessory genes between batch passages. Log$_2$ fold change of transcripts encoding core biosynthetic enzymes of natural products in batch passage generation 9 compared to batch passage generation 2. Only transcript with an absolute log$_2$FC>1 and adjusted p-value<0.05 are shown.

Scaffold	Gene start	Gene end	Strand	log2FC	Natural product class
257	174232	208818	-	5.78	NRPS
261	42962	46408	+	3.41	Other
428	49836	56505	+	3.09	T1PKS
257	167502	174029	-	3.08	NRPS
161	219512	224288	-	2.80	Cf fatty acid
177	189990	204620	-	2.42	NRPS
62	374962	386019	+	2.29	NRPS
71	88791	92873	-	2.25	Other
146	53142	63668	+	2.05	T1PKS
82	344910	348800	+	1.94	Cf saccharide
84	563563	565296	-	1.85	NRPS
18	122326	128928	-	1.79	T1PKS
54	90624	103068	+	1.62	T1PKS
90	396252	404816	+	1.17	T1PKS
560	8478	15509	-	1.03	Cf saccharide
191	283454	286403	-	-1.53	Cf saccharide
88	478798	479307	+	-1.55	Terpene
254	162638	165242	-	-1.68	Cf saccharide
239	120548	121745	+	-2.14	Cf saccharide
13	630206	631310	+	-7.00	Cf saccharide

In contrast to the penicillin-streptomycin-treated bagasse consortia, chloramphenicol-treated consortia exhibited only a small number of differentially regulated biosynthetic genes. Two PKS genes located on scaffolds 6 and 137 were significantly upregulated (log2FC>1, p-adjusted<0.05) from generation 2 to generation 9. Neither gene was differentially expressed from generation 2 to 9 of the penicillin-streptomycin-treated bagasse consortia. Differential expression analysis of the penicillin-streptomycin treated consortia (generations 2, 6, and 9) compared to the chloramphenicol-treated consortia resulted in the upregulation of three biosynthetic genes in penicillin-streptomycin-treated cultures (log$_2$FC>1, p-adjusted<0.05): one bacteriocin gene and two genes classified as other or fatty acid-like by ClusterFinder algorithm[144]. The relative abundance of prokaryotes within the chloramphenicol-treated bagasse consortia was less than the penicillin-streptomycin-treated

bagasse consortia (refer to Supplementary Table 3, Peng et al., *Nature Microbiology,* in press 2020).

6.3   Conclusions

The herbivore microbiome has not been sufficiently explored for its potential to provide natural products of great utility to society. We have demonstrated that enrichment is a powerful technique to recover high-quality metagenome-assembled genomes (MAGs) that can be mined for natural products. From 719 prokaryotic genomes, we uncovered 688 biosynthetic gene clusters (BGCs). Enrichment from the goat fecal microbiome changed the microbial communities and the associated suite of BGCs. Although overall there were more BGCs for each natural product class recovered from the source microbiome MAGs than the cultivated MAGs, a greater number of aryl polyenes, antimicrobial peptides (bacteriocins and microcins), and nonribosomal peptide synthetases (NRPSs) were recovered from the cultivated MAGs due to the selection for Firmicutes and Bacteroidetes. Each of the most abundant MAGs within the antibiotics-free consortia grown on different substrates (alfalfa, bagasse, reed canary grass) highly expressed at least one core biosynthetic gene. In penicillin-streptomycin-treated consortia enriched on bagasse, the anaerobic fungus *Neocallimastix californiae* upregulated 15 BGCs from generation to generation 9 of batch passaging, further suggesting that BGCs play a role in shaping the community structure.

6.4   Materials and Methods

6.4.1   Overview

Complete details of the enrichment procedure and subsequent cultivation can be found in Peng et al. (*Nature Microbiology*, in press 2020). Briefly, enrichment consortia from a

source inoculum of a fresh goat fecal pellet from the Santa Barbara Zoo were grown anaerobically on the following carbon substrates: (1) alfalfa, (2) bagasse, (3) reed canary grass, and (4) xylan. For each carbon substrate, triplicate consortia were subjected to one of three treatments: (1) no antibiotics, (2) chloramphenicol, or (3) penicillin/streptomycin. The consortia were passaged in serum bottles every 3-5 days, according to Figure 6.7.
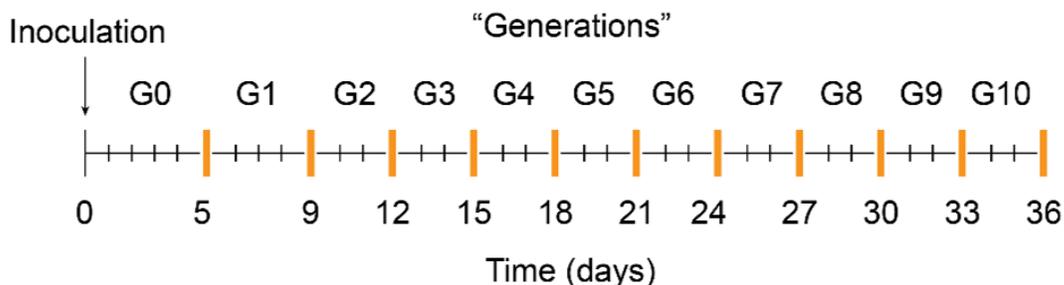
Figure 6.7: Schematic of the passaging of enrichment cultures from the fecal inoculum (Day 0) through the final cultures (Generation 10).

6.4.2   *RNA extraction and QC*

The frozen enrichment consortia that had been treated with RNAlater (QIAGEN) were thawed on ice and then centrifuged at 10,000 *g* using a fixed angle rotor (Eppendorf F-34-6-38) at 4 °C to remove the RNAlater. Cells were lysed by liquid nitrogen grinding. Total RNA was extracted using the RNeasy Plant Mini kit (QIAGEN) on a QIAcube (QIAGEN), following the RNeasy Mini protocol for animal cells with QIAshredder homogenizenization and optional on-column DNase digest. Since each sample was composed of ~0.5 g of plant material plus biomass, multiple spins were performed in the microcentrifuge in order to concentrate the total RNA onto a single RNA spin column for each sample before loading into the QIAcube. The quantity and quality of RNA were assessed by a QuBit 2.0 fluorometer (Invitrogen) and TapeStation (Agilent), respectively.

155

6.4.3   Library preparation, sequencing, and quality control

For antibiotics-free consortia, ribosomal RNA was depleted using the Illumina Ribo-Zero rRNA Removal Kit (Yeast) spiked into the Illumina Ribo-Zero Gold rRNA Removal Kit (Epidemiology). To obtain both prokaryotic and eukaryotic libraries from antibiotics-treated consortia, each sample was divided and 200 ng was used as the input into each alternative pipeline: 1) polyA selection for the fungal library using polyT beads, or 2) ribosomal depletion by Illumina Ribo-Zero rRNA Removal Kit (Yeast) spiked into the Illumina Ribo-Zero Gold rRNA Removal Kit (Epidemiology) for the library enriched in bacterial mRNA. Stranded RNA-seq libraries were created by the Joint Genome Institute and quantified by qPCR. Libraries were sequenced by paired-end dual-indexed150 bp reads using the NovaSeq S4 (Illumina, San Diego, CA). Quality control was performed for prokaryotic libraries following pipeline version 3.4.0 from bbtools[264] (version 38.20) for prokaryotic libraries, and following pipeline version 3.4.2 from bbtools[264] (version 38.22) for polyA libraries.

6.4.4   Mining 719 prokaryotic metagenome-assembled genomes (MAGs) for biosynthetic gene clusters (BGCs) and antibiotic resistance genes

719 prokaryotic metagenome-assembled genomes (MAGs) were filtered to exclude all contigs less than 10 kbp. antiSMASH 4.2 was then used to mine the filtered MAGs with the following options: subclusterblast, knownclusterblast, clusterblast, smcogs, and borderpredict. The taxon option was set to bacteria and the input option set to nucleotide.

FASTA files for 155 cultivated MAGs and 229 goat fecal pellet MAGs were uploaded to the Antibiotic Resistant Target Seeker (ARTS) version 2[22] web-interface, in groups of 20 or fewer files per run, based on the phylum reference set for comparison.

Default settings were used with the appropriate phylum reference set selected. Core hits are genes with matches to genes with essential housekeeping annotations in the TIGRFAMs[43] database. ResModel hits are genes with matches in the ResFams[44] database. CoreRes hits are genes that are both Core and ResModel. Individual BGCs were then manually inspected for at least one hit (Core, ResModel, or CoreRes) proximal to a biosynthetic gene (if specified). For the MAGs with novel BGCs, ARTS jobs were run with the domain of unknown function option selected, and any MAGs without a phylum reference (Euryarchaeota) were run with the metagenome reference set.

6.4.5   RNA-seq data analysis

RNA-seq reads from each sample were aligned to a concatenated fasta file for all 719 prokaryotic MAG plus one high-quality eukaryotic MAG (described in Peng et al. (*Nature Microbiology,* in press 2020). using BBMap[264] version 38.63 with the parameter 'minid' equal to 0.95. Biosynthetic gene cluster predictions by antiSMASH were converted to simple annotation format (SAF) and then FeatureCounts[221] was used to count the number of reads for each gene with primary read filtering (count primary alignments only). Counts were normalized to RPKM by dividing by the gene length (kbp) and the total number of reads mapped to the concatenated genome in FASTA format.

Analysis of eukaryotic reads from the polyA libraries was performed similarly by aligning the reads from the antibiotics-treated samples to 719 prokaryotic MAGs and the *Neocallimastix californiae* genome, available from the MycoCosm portal[23]. Predicted secondary metabolite gene clusters previously acquired by antiSMASH 3.0[106] (Chapter two). The signature genes were counted by featureCounts[221] using the options described above.

Differential expression analysis was performed using DESeq2[222] version 1.18.1 in R[270] version 3.4.3.

7 Conclusions

7.1 Perspectives

7.1.1 The rumen microbiome is a source of novel natural products

For many years, researchers have focused on the discovery of novel natural products with therapeutic and antimicrobial activities from environments such as the soil and ocean. The rumen microbiome has been neglected as a valuable reservoir of microbial natural products. With the advent and improvement of next-generation sequencing technologies as well as the refinement of bioinformatic techniques, researchers can now capture high-quality genomes of low abundance organisms such as anaerobic fungi. Mining metagenome-assembled genomes for biosynthetic gene clusters of natural products has revealed that the potential of rumen microbes has been underappreciated.

7.1.2 Resolving complex microbial partnerships: a crucial step towards microbiome engineering

The success of many processes is attributed to the performance of systems of microbes. The rumen microbiome is directly responsible for the breakdown of lignocellulosic matter into nutrients necessary for the survival of the host and also for the resulting methane emissions. The ability to tune the performance of such a complex system relies on our ability to discern complex microbial partnerships. The advantage of unraveling complex microbial interactions over using techniques that do not require mechanistic understanding, such as machine learning, is that such an understanding can inform the rationale design of synthetic

microbiomes to achieve different objectives or to tune the performance of different microbiomes. The human microbiome differs from the rumen microbiome in the input and microbial community membership, but the process of fermenting sugars into short chain fatty acids is the same. Experiments are much easier to conduct in ruminants than humans for many reasons, including the ability to control diet for an extended period of time. The human gut microbiome has become well-recognized as an integral part of human health[364–366]. Therefore, a mechanistic understanding of the complex network of microbial interactions in the rumen microbiome will inform strategies to treat human diseases linked to the health of the gut microbiome.

7.2   Future directions

7.2.1   Dual transcriptomics as a tool towards understanding the functions of fungal secondary metabolites in fungal-bacterial co-cultures

Co-cultivation of rumen fungi with known strains of anaerobic bacteria is a promising strategy towards deciphering the functions of their secondary metabolites. Dual transcriptomics of *Anaeromyces robustus* or *Caecomyces churrovis* with *Fibrobacter* sp. UWB7 (Chapter four) suggested that both the growth substrate and the fungal strain affect fungal gene expression when exposed to bacterial challenge, since the percentage of differentially expressed genes greatly varied between pairings and substrates. Literature[121,123] suggests that anaerobic fungi may also have an antagonistic relationship with *Ruminococcus* spp. Dual transcriptomics of strains of cultivable fungi (*A. robustus*, *C. churrovis, N. californiae*, and *P. finnis*) paired with *Ruminococcus* spp. (e.g. *Ruminococcus flavefaciens*) will shed light on whether the biosynthetic enzymes are activated by specific bacterial genera

or whether they are transcribed when the fungus is exposed to any rumen bacteria. More broadly, the native environment of the bacteria and its characteristics (e.g. Gram negative, Gram positive, spore-forming) may influence fungal BGC expression. Understanding the scope of fungal BGC expression when the fungus is exposed to different bacteria will inform whether fungal natural products could serve as narrow- or broad-spectrum antibiotics.

7.2.2 Strategies towards linking biosynthetic gene clusters to their pathway products

The results described in Chapter two report the biosynthetic potential of anaerobic fungi for the production of novel secondary metabolites. One of the greatest challenges in the field of natural products discovery as well as for the anaerobic fungi remains how to link the BGCs of an organism to their pathway products. In many cases, heterologous expression of native BGCs in a model host, such as *Escherichia coli* or *Saccharomyces cerevisiae* has proven to be a successful strategy to address this challenge[367–371]. We have several strains of *E. coli*, *S. cerevisiae*, and *A. nidulans* harboring the biosynthetic genes of anaerobic fungi (Appendix C). Systematic comparisons of the metabolites of the heterologous host to the wild-type host and the anaerobic fungus from which the genes originated will enable the connection of anaerobic fungal BGCs to their pathway products.

A third approach is to use genetic engineering tools such as RNA interference (RNAi) or clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9 (CRISPR-Cas9) gene editing. Some success towards genetic engineering of anaerobic fungi has been attained by RNAi[372]. However, the strains of anaerobic fungi discussed in this thesis are currently not genetically tractable. Nevertheless, the ability to systematically knock out selected biosynthetic genes of anaerobic fungi would enable the direct linkage of genes to their pathway products. A similar approach was used to knock out the BGCs of commonly

rediscovered antibiotics from strains of actinomycetes, leading to the discovery of novel secondary metaboliltes[136].

7.2.3 Characterizing the extent of species-level biosynthetic diversity

Two new strains of *Piromyces* (*Piromyces macparvus* and *Piromyces* sp. G2) are in the process of genomic sequencing. Once their genomes are successfully assembled, they can be mined using antiSMASH and their biosynthetic potential compared to *Piromyces finnis*. Although much work has already been done comparing the BGCs between genera, as detailed in chapter two, this will be the first time that high-quality genomes of multiple species of the same genus can be studied simultaneously. How conserved BGCs are between different species will deepen our understanding of the biosynthetic potential of anaerobic fungi. Comparisons of other fungi have revealed significant differences in the biosynthetic potential even at the species-level. For example, *Aspergillus fumigatus* and *Aspergillus novofumigatus* had 24 BGCs in common out of 90 total combined BGCs[373].

7.2.4 Further investigation of epigenetic regulation of the secondary metabolism of anaerobic fungi using ChIP-Seq or related techniques

The secondary metabolism of several organisms is regulated epigenetically. The application of the histone deacetylase inhibitor suberoylanilide hydroxamic acid (SAHA) to *Aspergillus niger* activated the synthesis of nigerone A[374] and the treatment of *A. niger* with DNA methyltransferase inhibitor 5-azacytidine transcriptionally upregulated many BGCs[375]. Evidence provided in Chapter four supports that epigenetic modifications such as histone 3 lysine 27 trimethylation (H3K27me3) together with homologs of the putative methyltransferase LaeA may regulate the secondary metabolism of anaerobic fungi.

Techniques such as chromatin immunoprecipitation (ChIP)-seq and Assay for Transposase-Accessible Chromatin (ATAC)-seq could be valuable in further characterizing epigenetic regulation of BGCs. However, both ChIP-seq and ATAC-seq require intact chromatin, which has not yet been successfully isolated from anaerobic fungi.

7.2.5 *Structural elucidation of novel secondary metabolites produced by anaerobic fungi*

Secondary metabolites are often produced in trace amounts during standard laboratory cultivation[33]. Prior to scale-up of fungal cultures to produce sufficient material for structural elucidation by 1D- and2D-NMR, one approach is to develop a separation pipeline for purification of bioactive compounds. For example, HPLC can be used to separate fractions, which can then be solidified by a Speed-Vac and reconstituted in a solvent. These concentrated fractions can be used for antimicrobial assays. This approach has been used successfully by others in natural products research[102]. Although laborious, preparation of at least semi-pure fractions is vital for the structural characterization of novel compounds of potential therapeutic or antibiotic value from anaerobic fungi.

7.2.6 *Characterization of fungal terpenes via GC-MS*

The metabolic characterization of the natural products of anaerobic fungi (discussed in chapters two and four) was limited to LC-MS/MS. For certain classes of natural products, such as terpenes, GC-MS is the best method of detection. The genomes of *A. robustus*, *C. churrovis*, and *P. finnis* all encode a single squalene synthase, an intermediate enzyme in the production of terpenoids and steroids, whereas *N. californiae* encodes two copies of the squalene synthase. Since this gene is conserved in four genera of anaerobic fungi, its

metabolic product may have an important biological function. Further experimentation is needed to decipher its function and structure.

7.2.7 *Leveraging self-resistance genes in biosynthetic gene clusters to discover antimicrobial compounds*

Discovering antibiotics with new modes of action is desirable but challenging. As the number of predicted BGCs has grown in parallel with the number of sequenced microbial genomes, new and innovative ways have been developed to prioritize the characterization of BGCs likely to produce antimicrobials. The Antibiotic Resistant Target Seeker[155] is a tool developed to prioritize bacterial BGCs based on the presence of duplicate, sequence-variant housekeeping genes co-localized in the gene cluster that confer self-resistance to the production of a toxic molecule. Chapter six highlighted the discovery of 688 BGCs from the sequencing of the goat fecal microbiome and subsequent cultivation of consortia enriched on different carbon substrates. These and other BGCs recently acquired through mining the Hungate 1000 collection[67] are rich datasets for further exploration with tools such as ARTS. A complementary approach to ARTS was developed by Vandova and colleagues to mine BGCs for self-resistance genes[156]. Their approach, while applied to bacterial Type I PKS genes, can be generalized to eukaryotic gene clusters and is thus an attractive method to apply to the BGCs of anaerobic fungi.

7.3 *Overall conclusions*

The capability of anaerobic fungi to synthesize natural products has been overlooked for decades. We established that the genomes of anaerobic fungi are rich in biosynthetic gene clusters encoding numerous classes of natural products including polyketides, nonribosomal

164

peptides, and antimicrobial peptides. The biosynthetic portfolio varied between genera of anaerobic fungi, although some gene clusters were conserved. The conservation of a Type I PKS gene between four different genera of anaerobic fungi implied an important biological function for the polyketide. Further sequencing of other anaerobic fungi will inform whether the polyketide is necessary for all anaerobic fungi. Sequence similarity and phylogeny of NRPS condensation domains and bacteriocin-like genes suggested that horizontal gene transfer was the mechanism of acquisition for some fungal biosynthetic genes. A quarter of the total biosynthetic gene clusters of four strains of anaerobic fungi were transcriptionally expressed during laboratory cultivation. Therefore, a significant portion of the biosynthetic potential of anaerobic fungi can be characterized without activation by abiotic or biotic stimuli. We detected by LC-MS/MS the presence of small, nonpolar molecules without matches to known compounds in spectral libraries, which further supported that anaerobic fungi synthesize novel natural products.

The functions that natural products synthesized by anaerobic fungi serve are not yet known, although we postulate that some of these natural products may modulate the interactions of anaerobic fungi with other microorganisms. Anaerobic fungi live with many other microorganisms within the digestive tracts of large herbivores. Their relationships with other microbial members of the community are complex, ranging from synergistic to antagonistic interactions. Synergistic interactions were demonstrated by the upregulation of fungal CAZymes when *Anaeromyces robustus* was co-cultured with *Methanobacterium bryantii* (Chapter three). Antagonistic interactions were revealed by the dual transcriptomic characterization of *A. robustus* or *C. churrovis* co-cultured with *Fibrobacter* sp. UWB7 (Chapter four), which upregulated genes encoding putative bacterial drug efflux pumps and fungal NRPS and PKS-like genes.

165

Upregulation of specialized metabolism is one mechanism that anaerobic fungi use to respond to stressors such as the presence of microbial competitors. However, anaerobic fungi also face other kinds of environmental stress and respond with conserved pathways such as the heat shock and unfolded protein responses. We also demonstrated that the basal position of anaerobic fungi on the fungal tree of life results in a unique stress response sharing elements of both the fungal and metazoan unfolded protein responses. Cellular stress responses are often triggered by the overexpression of native proteins or the production of recombinant proteins in a heterologous host. By characterizing the stress response of anaerobic fungi, we have paved the way for further development of anaerobic fungi as a biotechnology platform for the production of valuable biomolecules such as biomass-degrading enzymes.

Within complex microbial communities containing both fungi and bacteria enriched from a goat fecal microbiome via substrate selection and antibiotics treatment, anaerobic fungi upregulated the biosynthetic genes of their secondary metabolism in later generations of batch passaging relative to earlier cultivation. In antibiotics-free consortia enriched from the same source microbiome, the most abundant bacteria constitutively transcribed one or more biosynthetic genes of secondary metabolites at high levels compared to the rest of their transcriptome throughout 10 generations of batch passaging. Therefore, fungal and bacterial biosynthetic genes displayed different expression profiles during cultivation. In addition, the genomes of rumen bacteria are rich in ribosomally synthesized and post-translationally modified peptides (RiPPs) compared to the genomes of anaerobic fungi. More work is needed to determine the function of bacterial and fungal natural products in the native system, but these *in vitro* results suggest that natural products play an important role in shaping the rumen microbial community.

166

Engineers can leverage the secondary metabolism of the complex rumen ecosystem to develop novel therapeutic drugs and antibiotics, design synthetic consortia for industrial biomass degradation, or tune the native microbial system in order to benefit host health. Secondary metabolites are powerful tools for future microbiome engineering efforts. The rumen microbiome offers a unique environment to develop microbiome engineering strategies that leverage secondary metabolites.

8   Appendices

8.1   Appendix A: Supplementary figures for chapter two
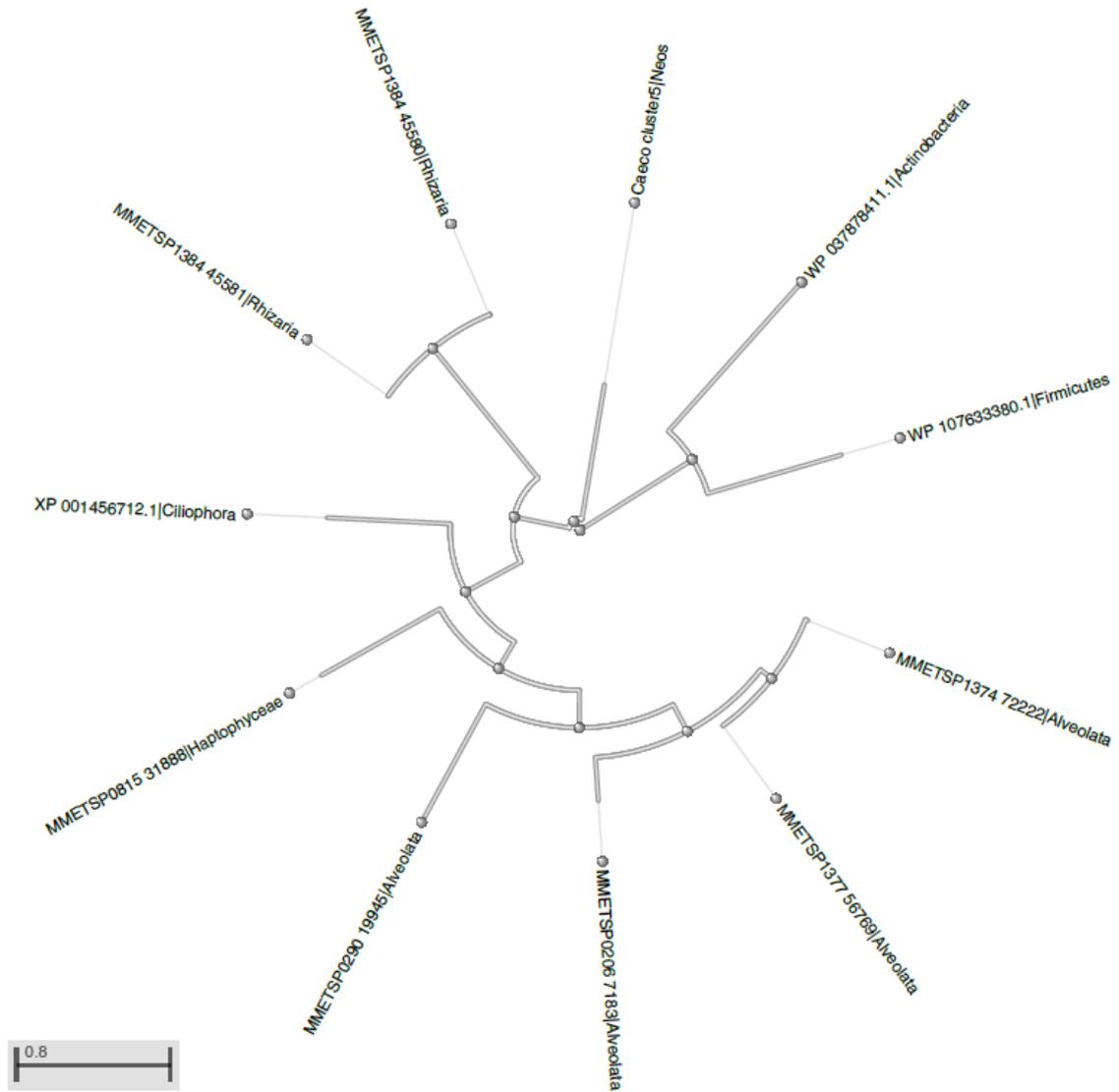
Supplementary Figure 8.1.1: Phylogeny of *C. churrovis* bacteriocin located on scaffold 83 (antiSMASH cluster 5), represented as "Caeco cluster 5|Neos." Tree rendered using NCBI Tree Viewer 1.17.5.
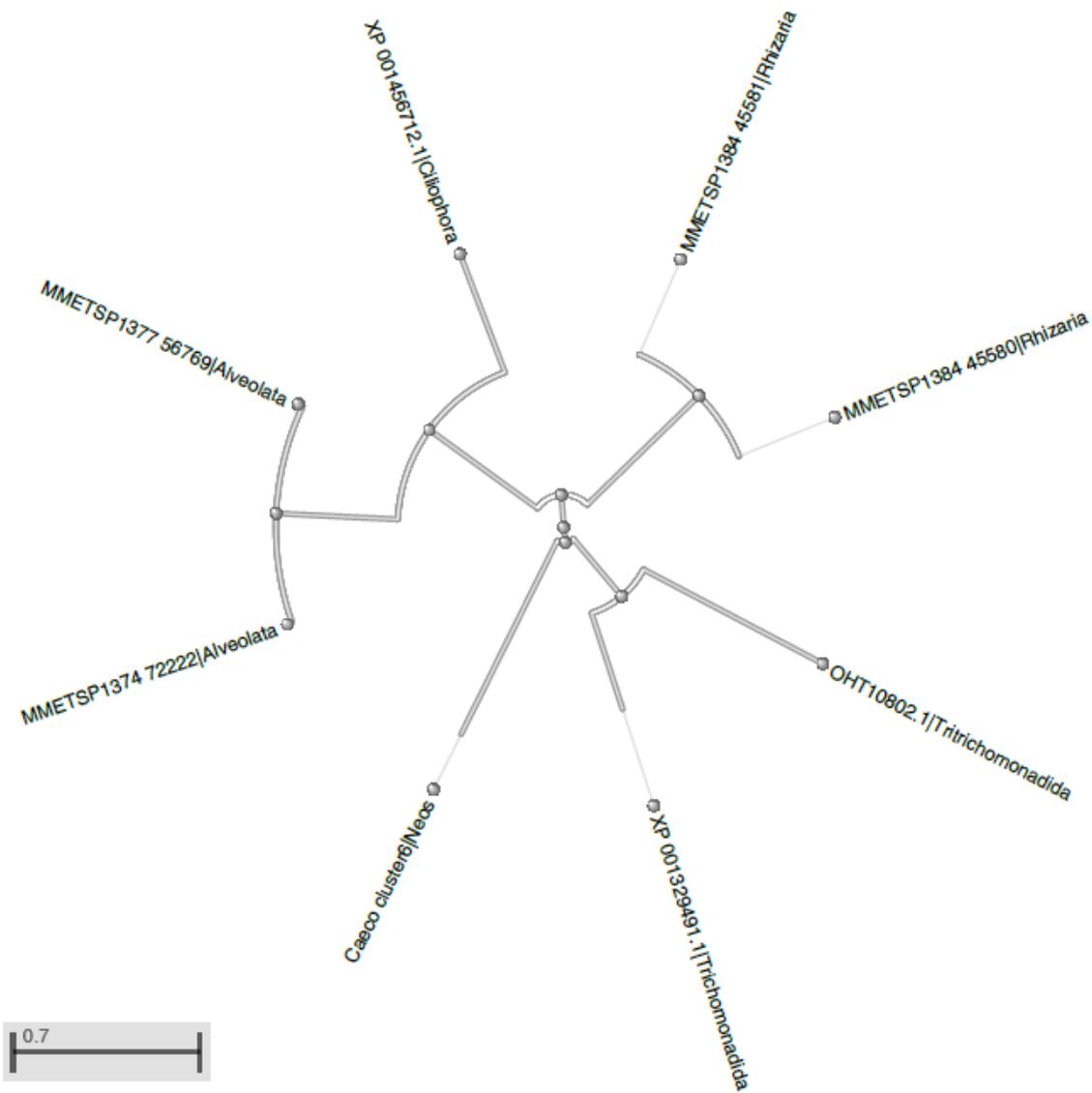
Supplementary Figure 8.1.2: Phylogeny of *C. churrovis* bacteriocin located on scaffold 90 (antiSMASH cluster 6), represented as "Caeco cluster 6|Neos." Tree rendered using NCBI Tree Viewer 1.17.5.
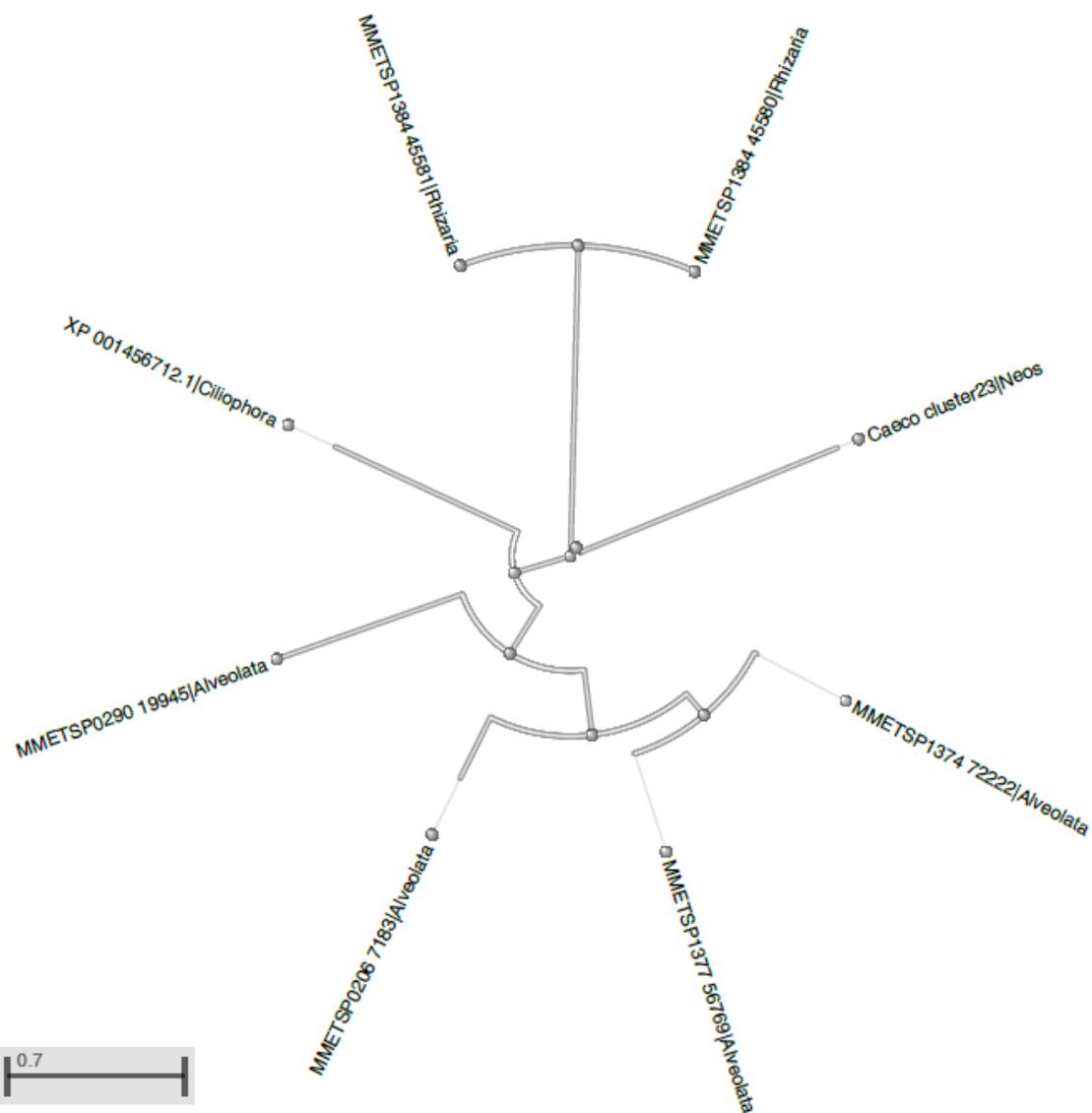
Supplementary Figure 8.1.3: Phylogeny of *C. churrovis* bacteriocin located on scaffold 616 (antiSMASH cluster 23), represented as "Caeco cluster23|Neos." Tree rendered using NCBI Tree Viewer 1.17.5.
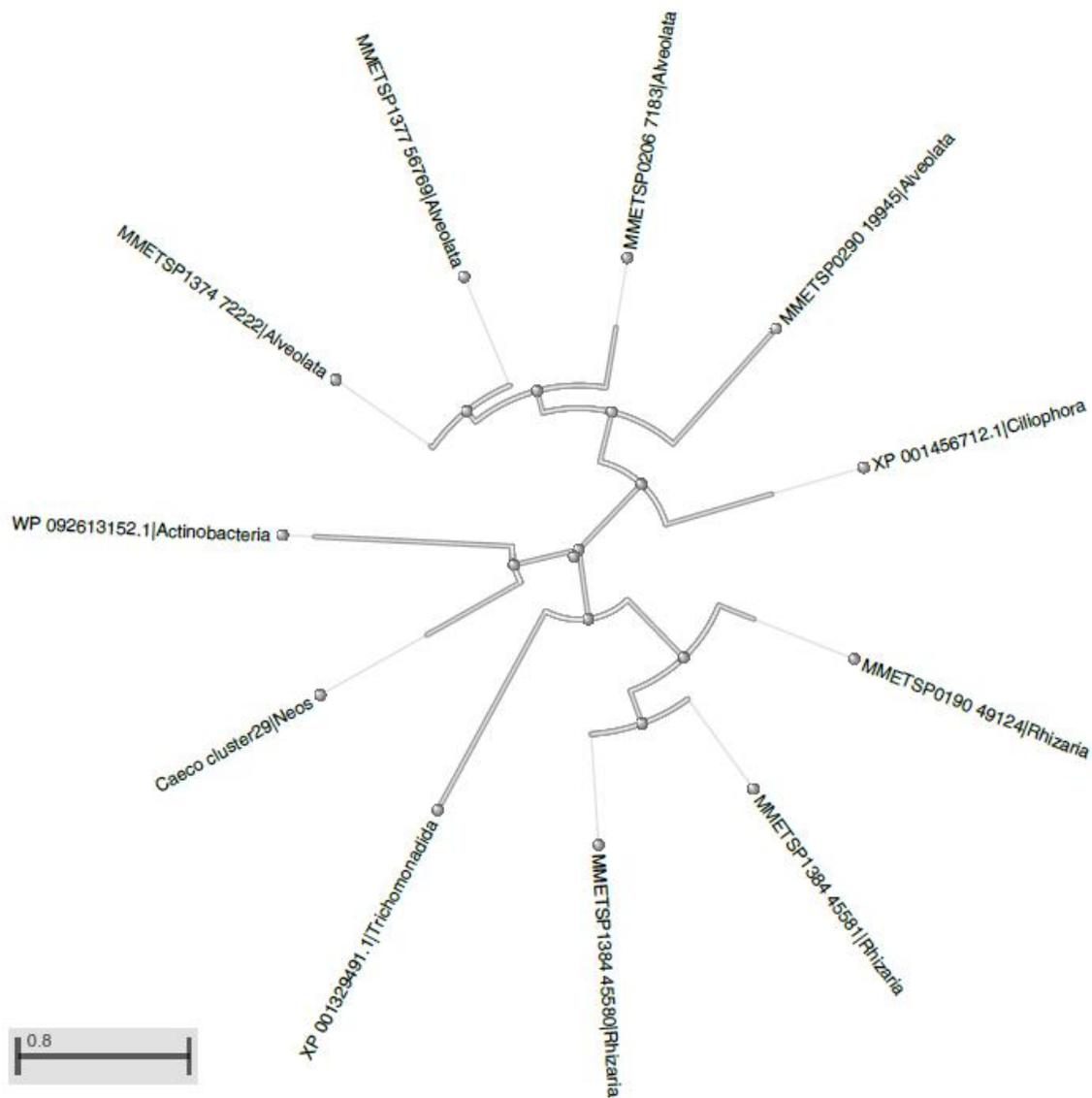
Supplementary Figure 8.1.4: Phylogeny of *C. churrovis* bacteriocin located on scaffold 1501 (antiSMASH cluster 29), represented as "Caeco cluster 29|Neos." Tree rendered using NCBI Tree Viewer 1.17.5.

171
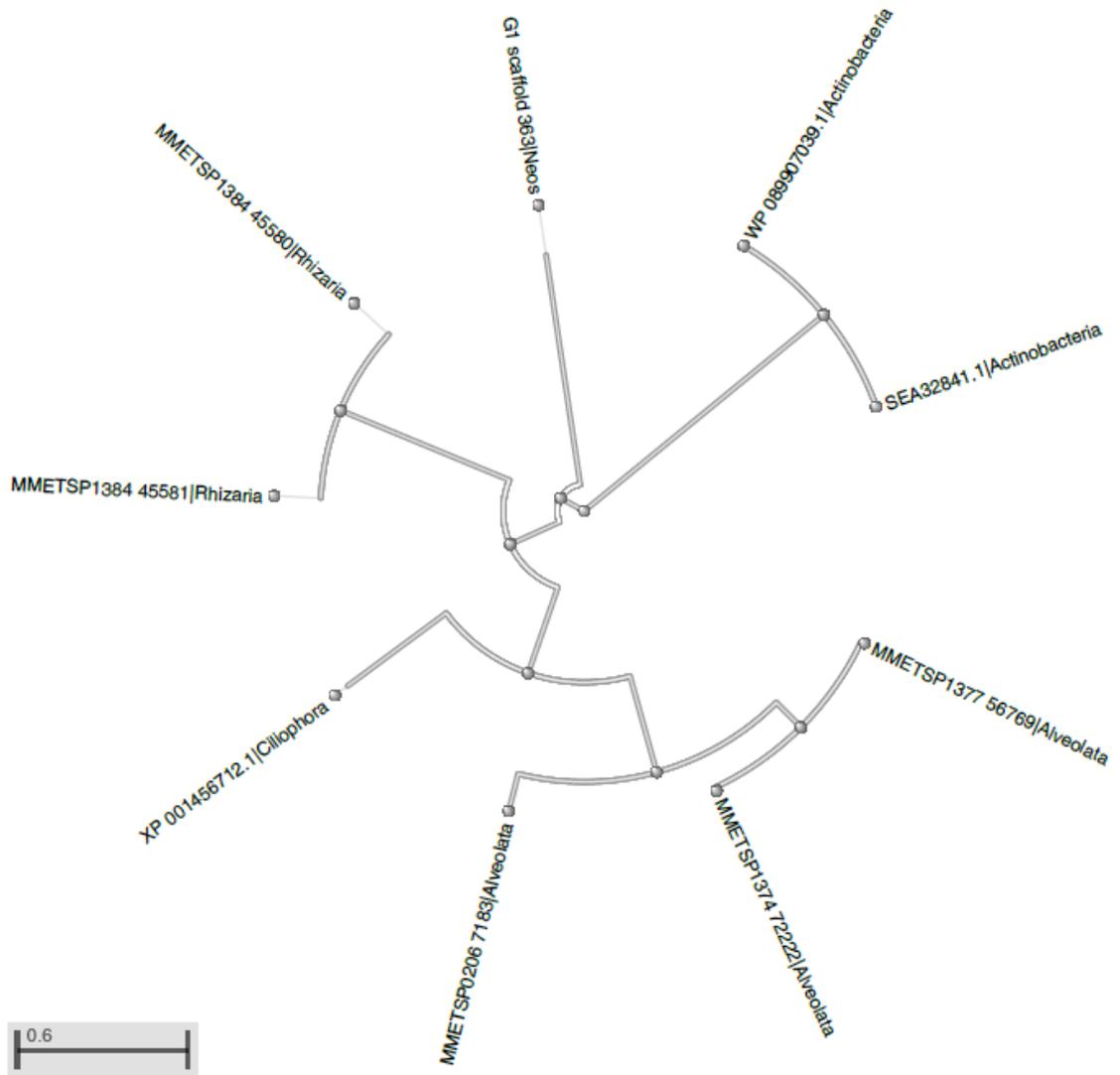
Supplementary Figure 8.1.5: Phylogeny of *N. californiae* bacteriocin located on scaffold 363 (antiSMASH cluster 4), represented as G1 scaffold 363|Neos." Tree rendered using NCBI Tree Viewer 1.17.5.
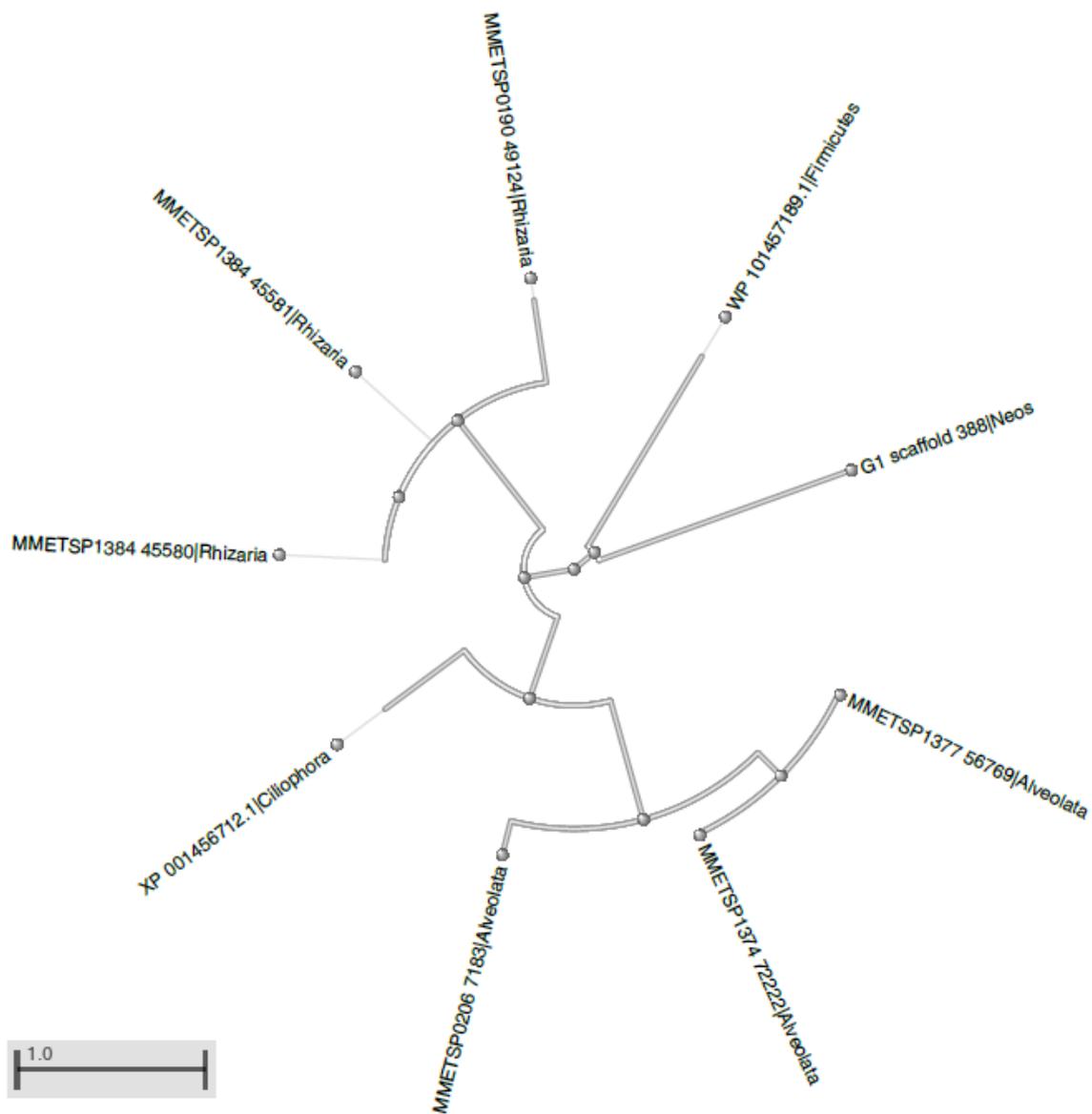
Supplementary Figure 8.1.6: Phylogeny of *N. californiae* bacteriocin located on scaffold 388 (antiSMASH cluster 5), represented as "G1 scaffold 388|Neos." Tree rendered using NCBI Tree Viewer 1.17.5.
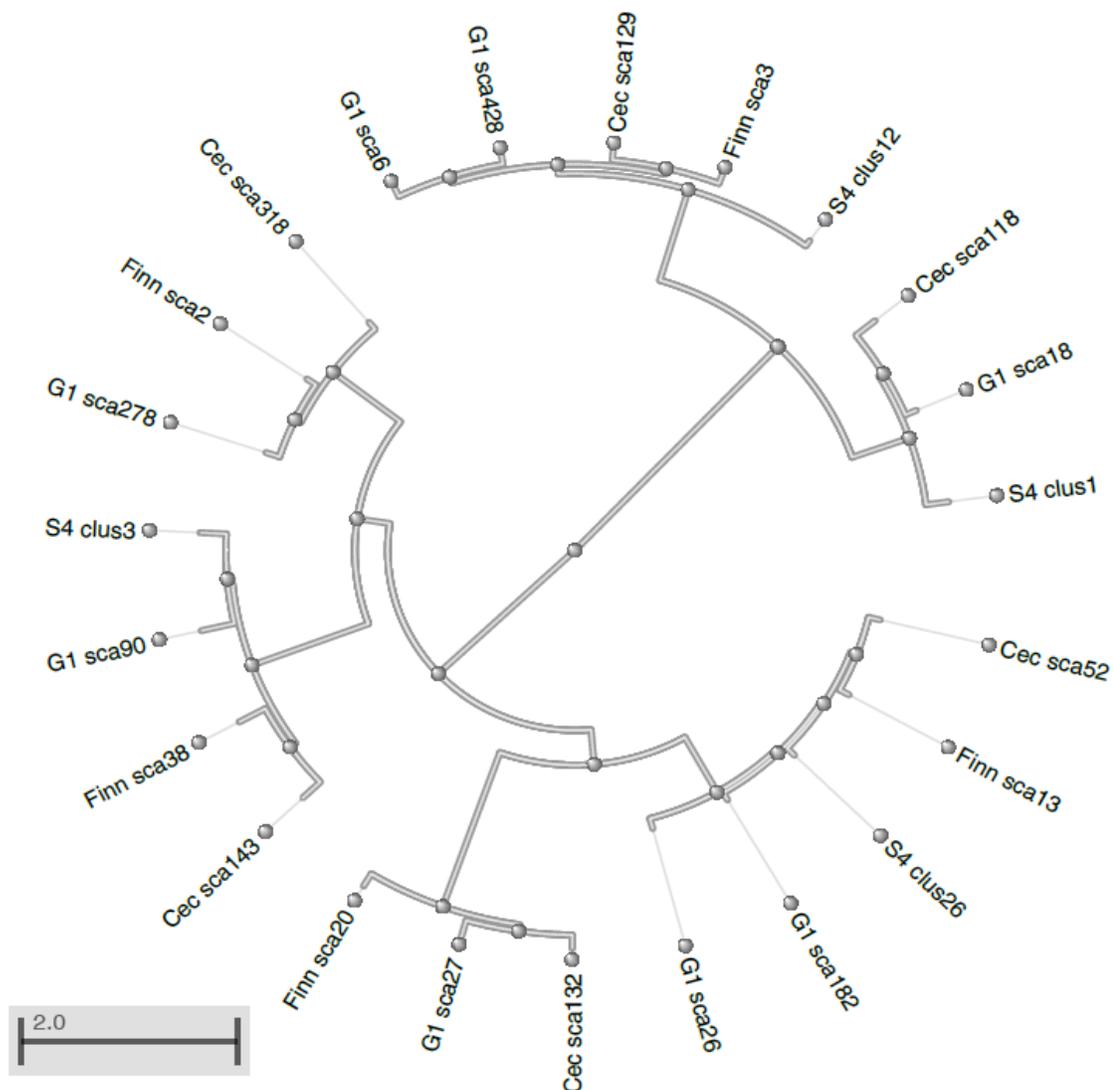
Supplementary Figure 8.1.7: Maximum likelihood phylogenetic tree of type I PKS genes from
A. robustus, *C. churrovis*, *N. californiae*, **and** *P. finnis*. Each gene is represented by its scaffold or
cluster number (e.g. "sca2" signifies the PKS gene on scaffold 2, "clus3" is the PKS gene in cluster
3). Cec=C. churrovis, G1=*N. californiae*, S4=*A. robustus*, Finn=*P. finnis*. All genes were aligned
using ClustalW[313,336]. The resulting alignment file was used as input to construct a maximum
likelihood phylogenetic tree by the RAxML [179]. HPC2 tool on XSEDE. RAxML Default input
parameters were used for both ClustalW and RAxML. Tree rendered using NCBI Tree Viewer
1.17.5.

8.2   Appendix B: Supplementary figures for chapter four

H3 Blot   H3K4me3 Blot   H3K27me3 Blot

39.216 (s)   151.147 (s)   455.833 (s)
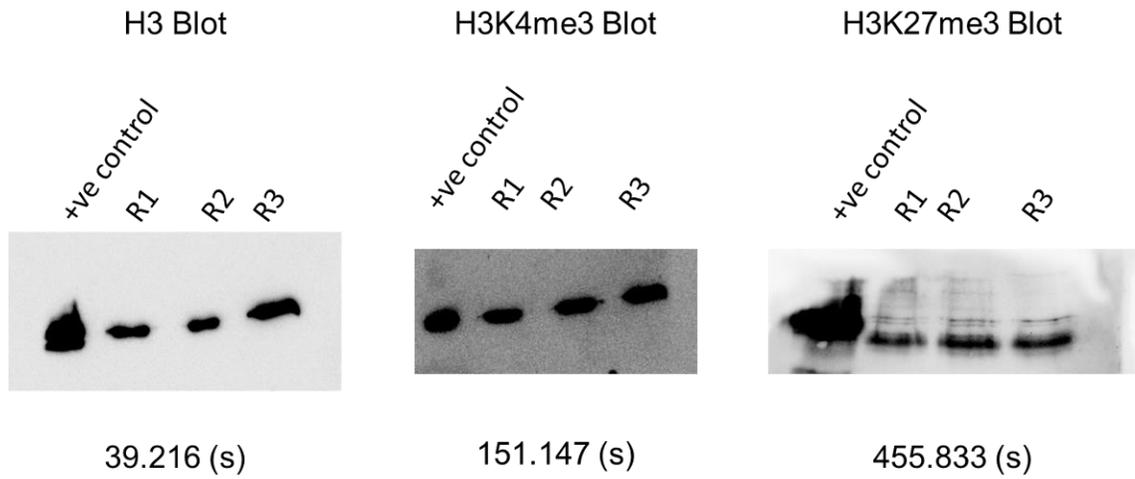
Supplementary Figure 8.2.1: Western blot images of *A. robustus* monocultures grown on Avicel®.

H3 Blot   H3K4me3 Blot   H3K27me3 Blot

832.151 (s)   3211.121 (s)   613.07(s)
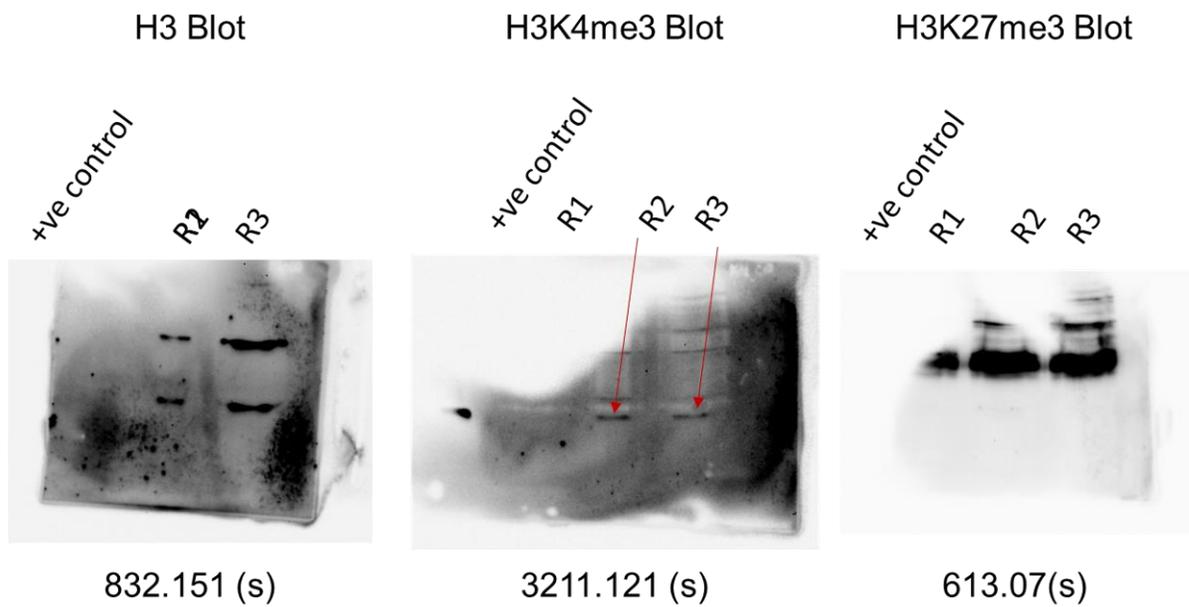
Supplementary Figure 8.2.2: Western blot images of *A. robustus* co-culture with *F.* sp. UWB7 grown on Avicel®.
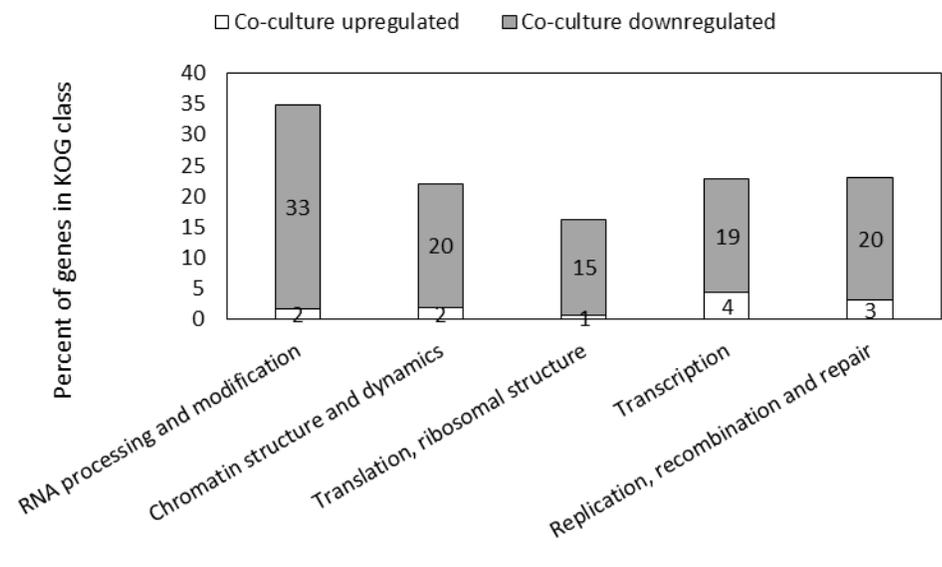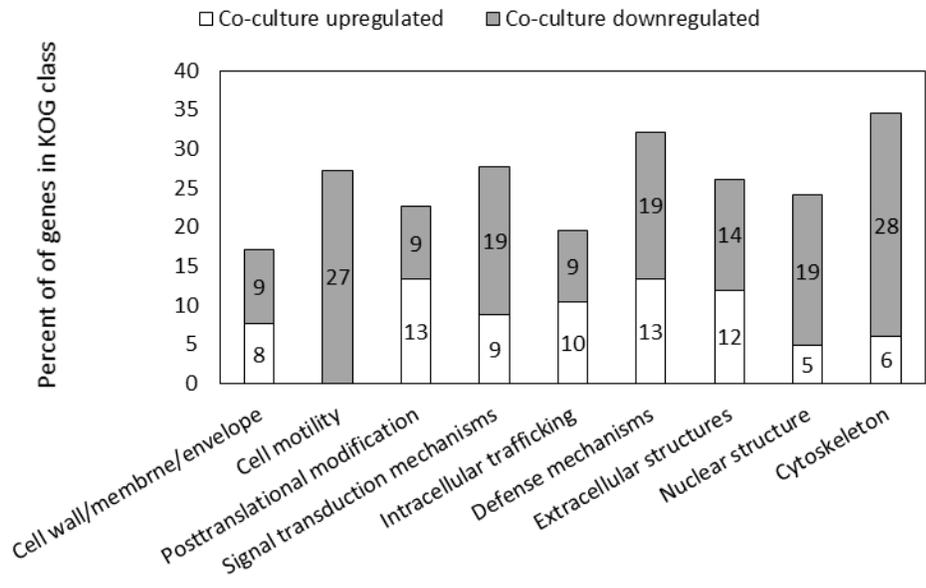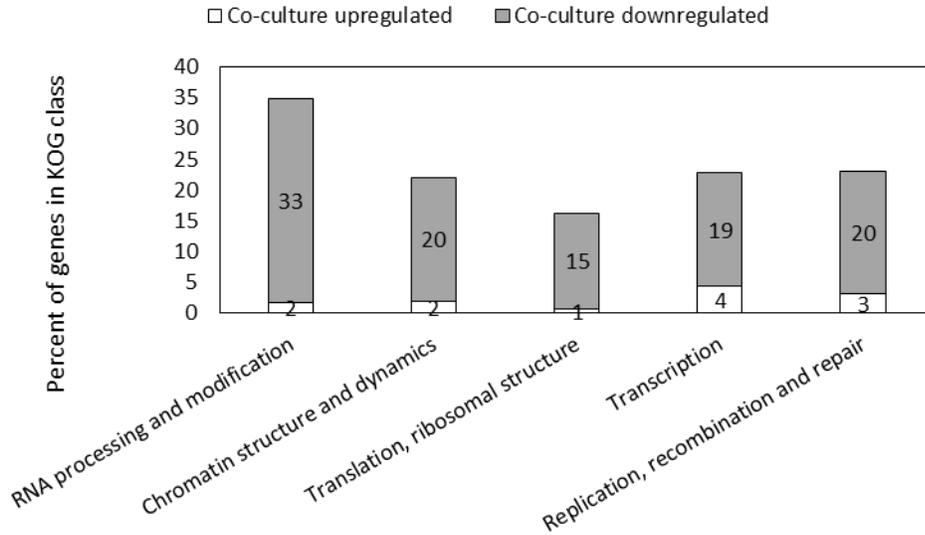
Supplementary Figure 8.2.3: Proportion of differentially expressed genes in each eukaryotic Orthologous Group (KOG)[212] class for *A. robustus* in co-culture with *F.* sp. UWB7 on Avicel® relative to *A. robustus* monoculture on Avicel®. KOG classes are organized into three plots: cellular processes and signaling (top), information storage and processing (middle), and metabolism (bottom). All carbohydrate active enzymes (CAZymes) with catalytic domains were binned into the KOG class Carbohydrate transport and metabolism and all CAZymes without a catalytic domain were excluded from this analysis.
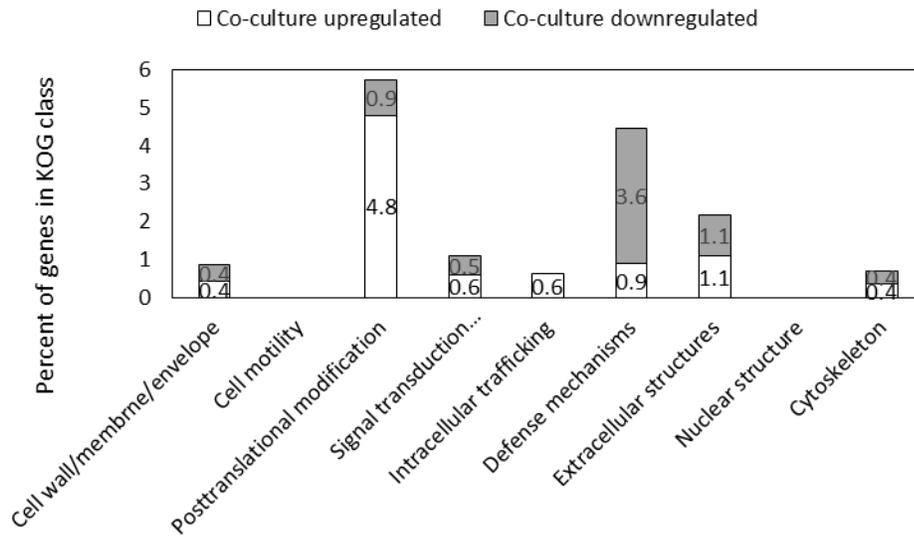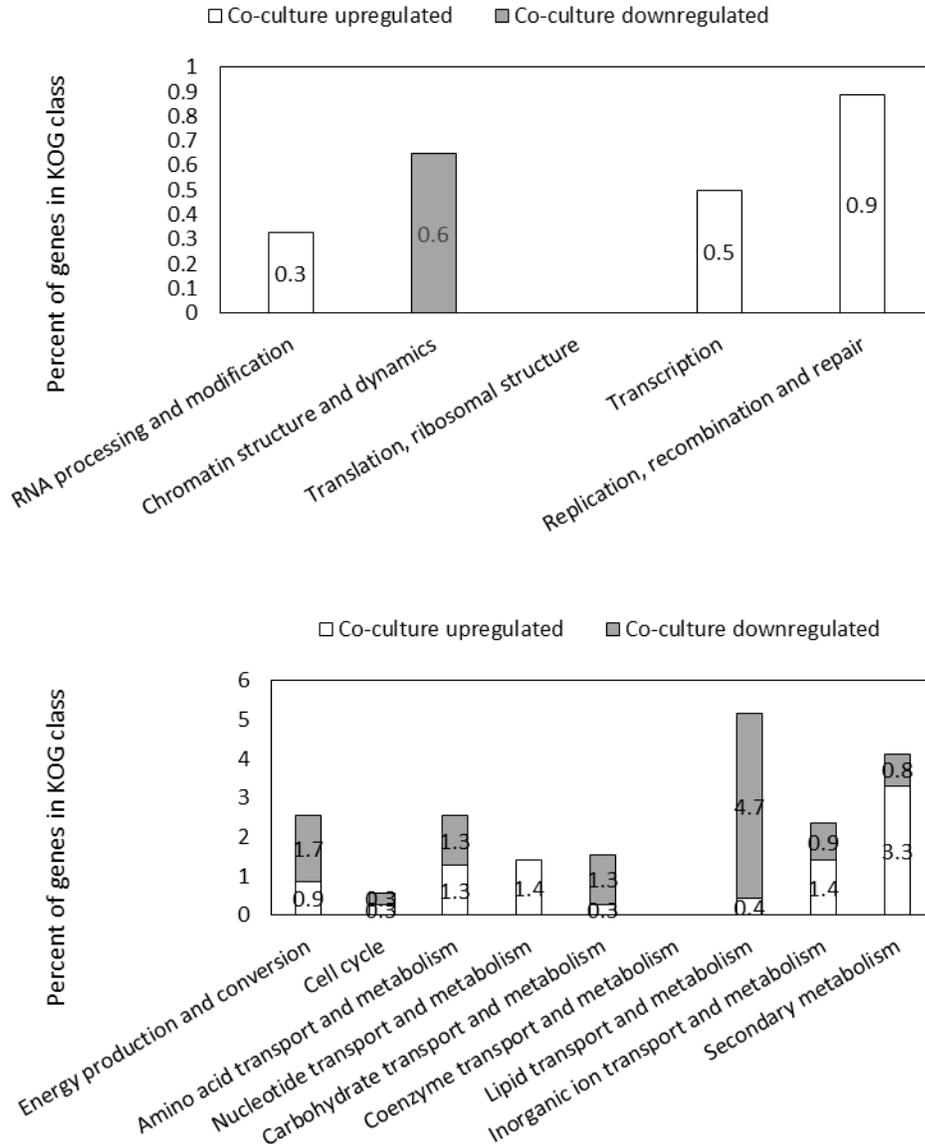
Supplementary Figure 8.2.4: Proportion of differentially expressed genes in each euKaryotic Orthologous Group (KOG)[212] class for *A. robustus* in co-culture with *F.* sp. UWB7 on switchgrass relative to *A. robustus* monoculture on switchgrass. KOG classes are organized into three plots: cellular processes and signaling (top), information storage and processing (middle), and metabolism (bottom). All carbohydrate active enzymes (CAZymes) with catalytic domains were binned into the KOG class Carbohydrate transport and metabolism and all CAZymes without a catalytic domain were excluded from this analysis.
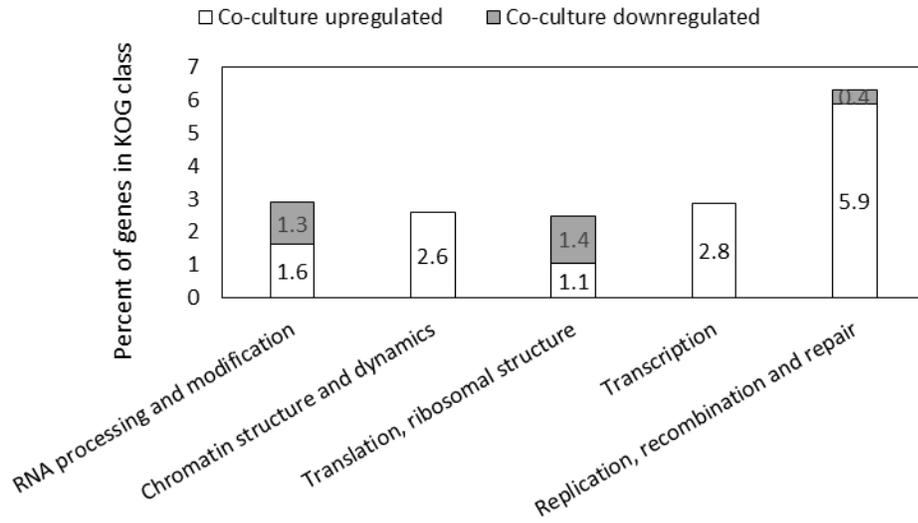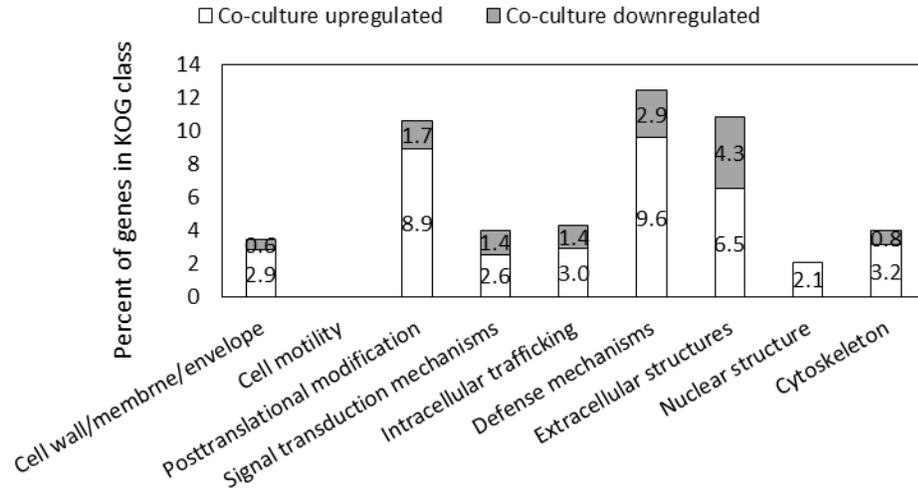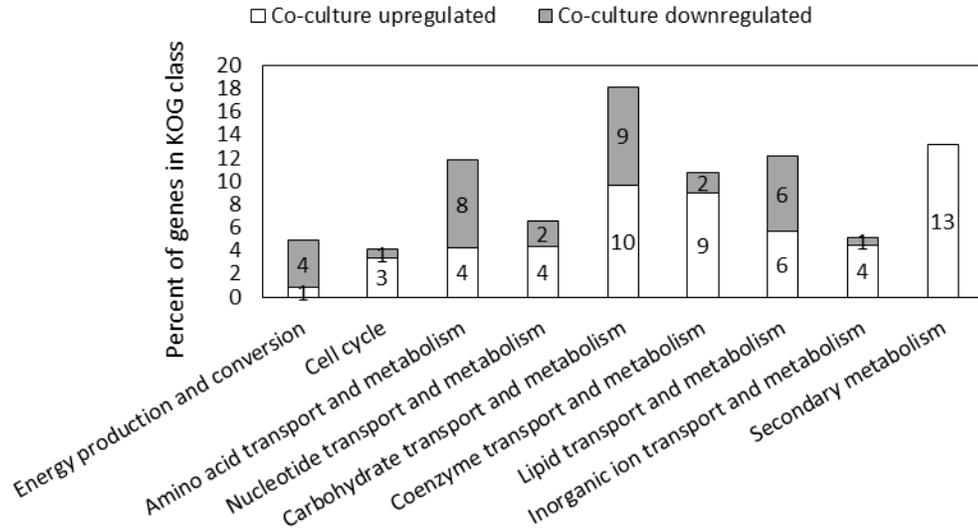
Supplementary Figure 8.2.5: Proportion of differentially expressed genes in each euKaryotic Orthologous Group (KOG)[212] class for *C. churrovis* in co-culture with *F.* sp. UWB7 on switchgrass relative to *C. churrovis* monoculture on switchgrass. KOG classes are organized into three plots: cellular processes and signaling (top), information storage and processing (middle), and metabolism (bottom). All carbohydrate active enzymes (CAZymes) with catalytic domains were binned into the KOG class Carbohydrate transport and metabolism and all CAZymes without a catalytic domain were excluded from this analysis.
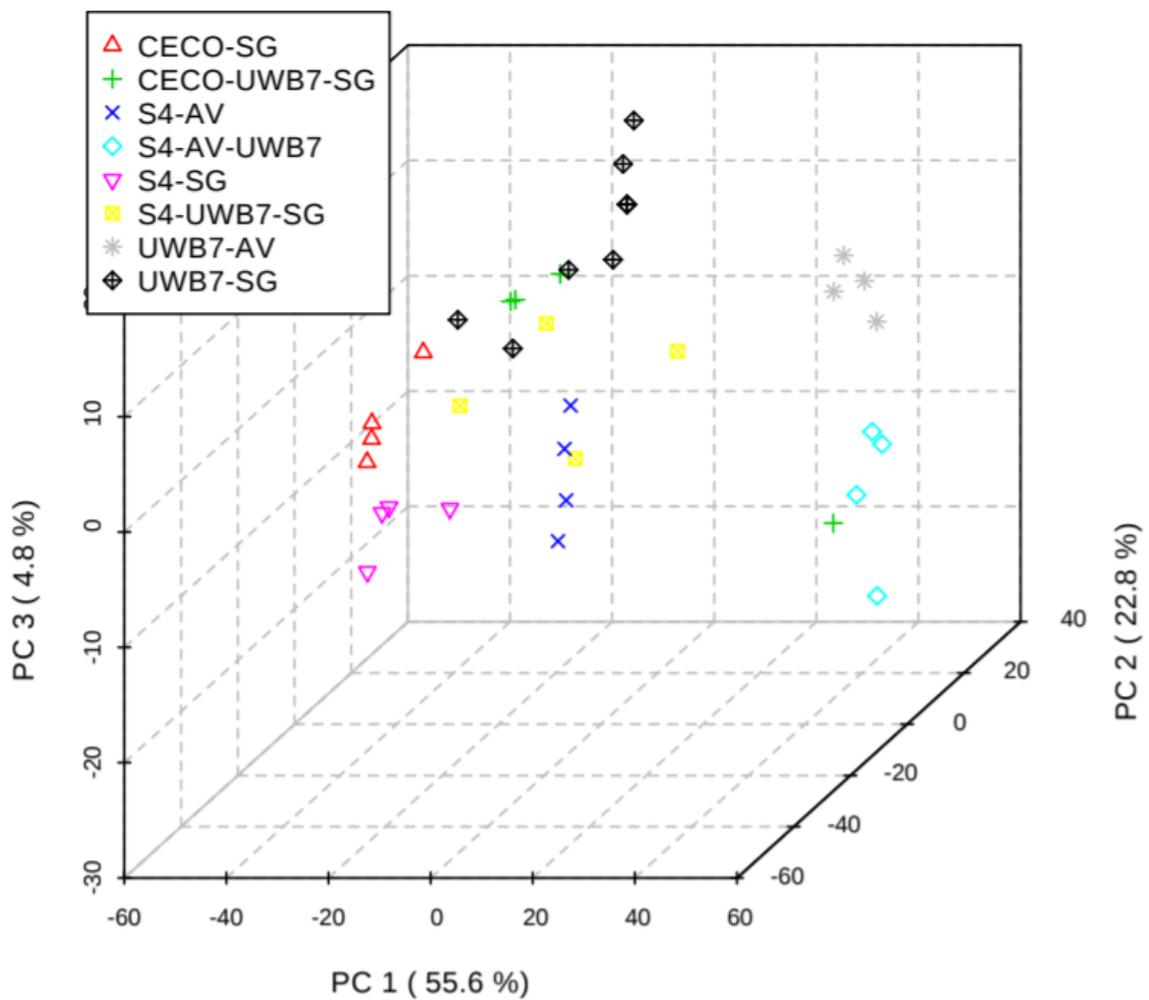
Supplementary Figure 8.2.6: Three-dimensional principal component analysis (PCA) plots of the untargeted nonpolar metabolomics data for co-cultures and monocultures of *A. robustus*, *C. churrovis*, and *F*. sp. UWB7. S4=*A. robustus*, Ceco=*C. churrovis*, UWB7=*F*. sp. UWB7, AV=Avicel®, SG=switchgrass. Plots were rendered by MetaboAnalyst[258].

8.3 Appendix C: Heterologous expression of the biosynthetic genes of anaerobic Gut fungi

Table 8.3.1: *Saccharomyces cerevisiae* **BJ5464-NpgA strains harboring PKS or NRPS genes from anaerobic fungi.** Vectors contain uracil auxotrophic marker and ADH2 promoter.

Strain name	Native fungus	SM Type	PKS family*	MycoCosm Protein Ids
B211.1	*A. robustus*	NRPS	NA	**
B211.2	*A. robustus*	NRPS	NA	328851, 270536, 270537, 295505, 207703
B211.3	*A. robustus*	PKS	1	270773, 295743, 283391, 295774, 329005, 248107, 329007, 270778
B211.4	*A. robustus*	PKS	4	
B227.1	*C. churrovis*	PKS	4	27727
B227.2	*C. churrovis*	PKS	1	489529
B227.3	*C. churrovis*	PKS	3	547760
B227.4	*P. finnis*	PKS	6	316229

*PKS families refer to those described in Chapter two.

**No gene model exists in MycoCosm. See antiSMASH prediction for scaffold 240 (ctg240_allorf000208, cluster 24) from Swift et al., manuscript in preparation.

Table 8.3.2: *Aspergillus nidulans* **strains harboring NRPS genes from anaerobic fungi.**

Strain name	Native fungus	SM Type	MycoCosm Protein Ids	Host
B227B_p001	*C. churrovis*	NRPS	554372	pCC1FOS_AMA1_pAlcA
B227B_p002	*N. californiae*	NRPS	701295	pCC1FOS_AMA1_pAlcA
B227B_p005	*A. robustus*	NRPS	330657	pCC1FOS_AMA1_pAlcA
B227B_p007	*C. churrovis*	NRPS	554372	pCC1FOS_AMA1_gpdA
B227B_p008	*N. californiae*	NRPS	701295	pCC1FOS_AMA1_gpdA
B227B_p011	*A. robustus*	NRPS	330657	pCC1FOS_AMA1_gpdA

Table 8.3.2: *Escherichia coli* strains harboring PKS genes from anaerobic fungi. Vector is pESC-Ura with Kanamycin resistance marker for all strains. The same PKS genes are also available within pET28Y in *E. coli*, with the exception of PKS_003.

Strain name	Native fungus	SM Type	PKS family*	antiSMASH reading frame**	scaffold
PKS_001	*Anaeromyces robustus*	Type I PKS	1	ctg258_allorf000323	258
PKS_002	*Neocallimastix californiae*	Type I PKS	1	ctg26_allorf000372	26
PKS_003	*Neocallimastix californiae*	Type I PKS	1	ctg41_allorf002517	182
PKS_004	*Neocallimastix californiae*	Type I PKS	4	ctg118_orf000000	428
PKS_005	*Anaeromyces robutsus*	Type I PKS	4	ctg127_allorf000908	127
PKS_006	*Neocallimastix californiae*	Type I PKS	4	ctg6_orf20	6

*PKS families refer to those described in Chapter two.
**See Supplementary Table S3 of Swift et al., manuscript in preparation for sequences.

8.4   Appendix D: Heterologous expression of a bacteriocin from *Caecomyces churrovis* in *Saccharomyces cerevisiae*

Bacteriocins are a class of natural products typically produced by bacteria, but surprisingly both *Neocallimastix californiae* and *Caecomyces churrovis* genomes encoded putative bacteriocins (see Chapter two). We cloned a histidine-tagged putative bacteriocin gene located on scaffold 90 of the *C. churrovis* genome into pITY with a prepro signal system using a double digest by the restriction enzymes SacII and EagI. The pITY construct was transformed into *Saccharomyces cerevisiae* BJ5464. The histidine-tagged bacteriocin was under the control of a galactose inducible promoter. The putative bacteriocins was predicted to be an 11 kDa peptide, although it is not known whether the bacteriocin is post-translationally modified in the native system, although it is predicted to be glycosylated. Production of the bacteriocin by the heterologous strain was verified by Western blotting. A 50 kDa protein was observed in the heterologous strain, but not in the BJ5464 control (Figure 8.4.1). The 50 kDa protein may be a hyperglycosylated form of the predicted 11 kDa

bacteriocin. The bacteriocin was also verified at the Joint Bioenergy Energy Institute by mass spectrometry, although the level of soluble bacteriocin detected was low.
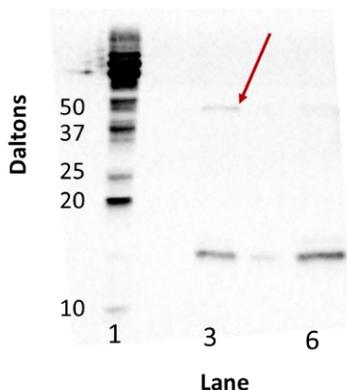
Figure 8.4.1: Western blot to verify production of *C. churrovis* bacteriocin in BJ5464. Lane 1: Precision Plus Protein™ WesternC™ standard (Bio-Rad), Lane 3: cell pellet from BJ5464 harboring bacteriocin gene from *C. churrovis*, Lane 6: BJ5464 control. A native protein with non-specific binding to anti-His was observed at 15 kDa. The non-native protein is visible at 50 kDa (red arrow).

8.5   *Appendix E: Plate-based assays of fungal bioactivity against Fibrobacter sp. UWB7*

Bioactivity assays are powerful complementary approaches to computational tools such as molecular networking in the quest to discover compounds with desired bioactivities[101]. In order to assess for antibiosis between anaerobic fungi and *Fibrobacter* sp. UWB7, we grew lawns of *F.* sp. UWB7 in an anaerobic chamber (AS-580, Anaerobe Systems) with a headspace of 5% hydrogen, 20% carbon dioxide, and balance nitrogen. For each plate (1 w/v% agar in 20 mL of complex medium C mins, "MC-", supplemented with volatile fatty acids, 10 mM glucose and 20 mM cellobiose), we then added autoclaved Whatman filter paper discs (0.6 μm pore size, 25 mm diameter circles, VWR catalog #28158-158) dipped into fungal monoculture supernatant, fungal-*F*.sp. UWB7 co-culture supernatant, or blank growth medium as a control.

Although initial results were promising (Figure 8.4.1), they were not reproducible in separate experiments (results not shown). We varied several parameters including the amount and concentration of bacterial inoculum for the plate, duration of the cultivation prior to harvesting the supernatant and the fungal strain (experiments were conducted with *Anaeromyces robustus*, *Caecomyces churrovis*, *Neocallimastix californiae*). Growth of *F. sp.* UWB7 was inconsistent on plates in the chamber and often did not form a lawn, which further complicated the experiments. It is possible that the initial rings of inhibition were the effect of bacterial cell dilution from the spread of residual supernatant seeping out from the filter disc, although every effort was made to shake off excess supernatant from the disc before placement.

In the future, other types of bioactivity assays should be tried without the use of tiler paper discs. Some possibilities include measuring the growth through optical density of *F. sp.* UWB7 on a soluble sugar with and without the addition of filtered fungal supernatant. Another alternative is to use a soft agar overlay inoculated with bacteria and spotting on the fungal supernatant, which is a method that has been used successfully in the assessment of albicidins[102]. Furthermore, the bioactivity assays should be validated with a known antibiotic, such as penicillin-streptomycin or chloramphenicol.

(A)

(B)
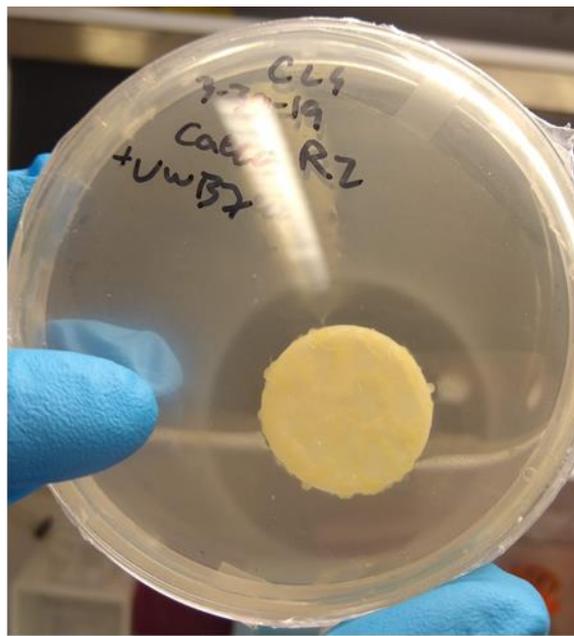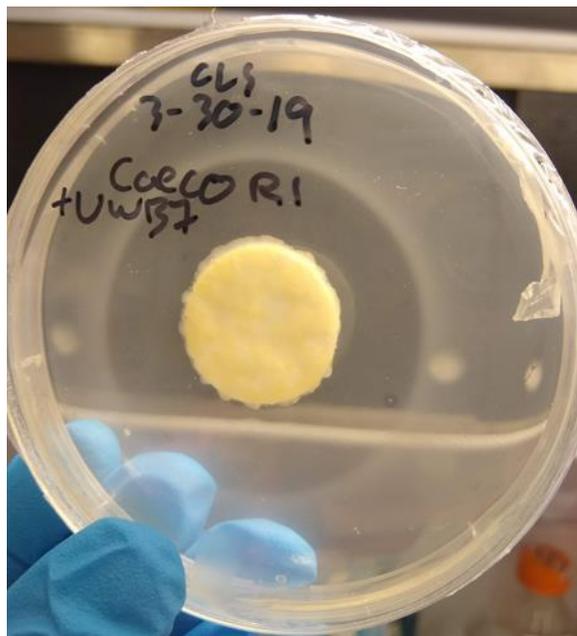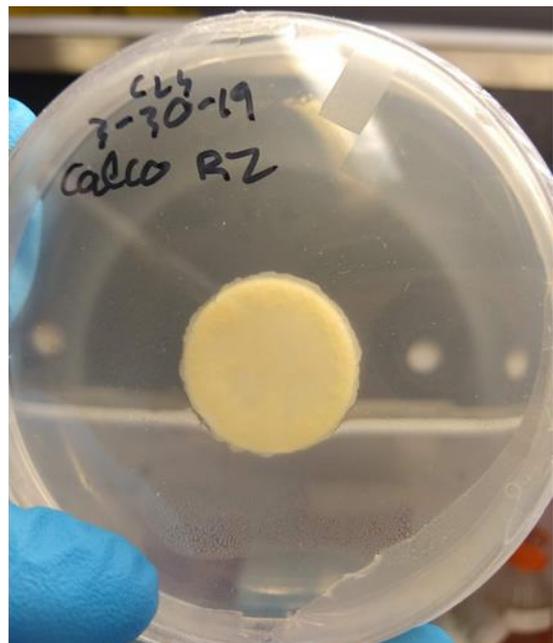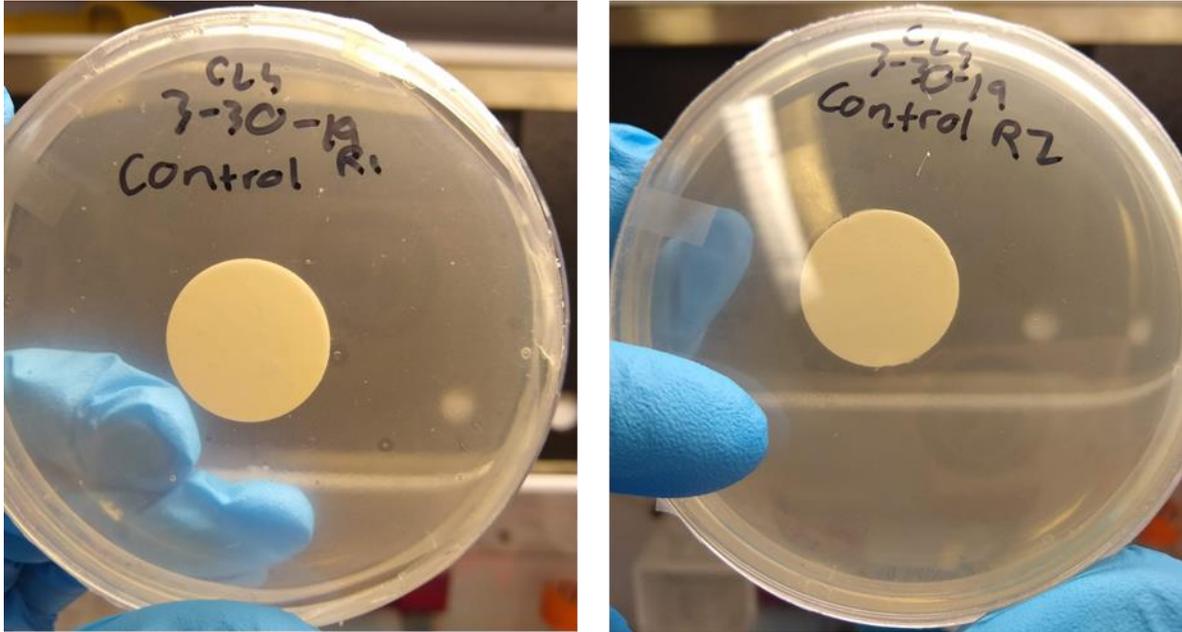
Figure 8.5.1: Rings of inhibition were observed in lawns of *F*. sp. UWB7 when filter paper discs were added that had been dipped in *Caecomyces churrovis* supernatant. Filter paper discs were dipped in (A) *C. churrovis-F.* sp. UWB7 co-culture supernatant, (B) *C. churrovis* monoculture supernatant, and (C) blank growth medium. Biological replicates 1 (left) and 2 (right).

9   References

(1)   Patridge, E.; Gareiss, P.; Kinch, M. S.; Hoyer, D. An Analysis of FDA-Approved Drugs: Natural Products and Their Derivatives. *Drug Discov. Today* **2016**, *21* (2), 204–207. https://doi.org/10.1016/j.drudis.2015.01.009.

(2)   Lewis, K. Platforms for Antibiotic Discovery. *Nat. Rev. Drug Discov.* **2013**, *12* (5), 371–387. https://doi.org/10.1038/nrd3975.

(3)   Mohimani, H.; Gurevich, A.; Mikheenko, A.; Garg, N.; Nothias, L. F.; Ninomiya, A.; Takada, K.; Dorrestein, P. C.; Pevzner, P. A. Dereplication of Peptidic Natural Products through Database Search of Mass Spectra. *Nat. Chem. Biol.* **2017**, *13* (1), 30–37. https://doi.org/10.1038/nchembio.2219.

(4)   Gross, H. Genomic Mining - A Concept for the Discovery of New Bioactive Natural Products. *Current Opinion in Drug Discovery and Development*. 2009, pp 207–219.

(5)   Trinci, A. P. J.; Davies, D. R.; Gull, K.; Lawrence, M. I.; Bonde Nielsen, B.; Rickers, A.; Theodorou, M. K. Anaerobic Fungi in Herbivorous Animals. *Mycol. Res.* **1994**, *98* (2), 129–152. https://doi.org/10.1016/S0953-7562(09)80178-0.

(6)   Solomon, K. V.; Haitjema, C. H.; Henske, J. K.; Gilmore, S. P.; Borges-Rivera, D.; Lipzen, A.; Brewer, H. M.; Purvine, S. O.; Wright, A. T.; Theodorou, M. K.; et al. Early-Branching Gut Fungi Possess a Large, Comprehensive Array of Biomass-Degrading Enzymes. *Science (80-.).* **2016**, *351* (6278), 1192–1195. https://doi.org/10.1126/science.aad1431.

(7)   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.; et al. A Parts List for Fungal Cellulosomes Revealed by Comparative Genomics. *Nat. Microbiol.* **2017**, *2* (May),

17087. https://doi.org/10.1038/nmicrobiol.2017.87.

(8)   Karlovsky, P. *Secondary Metabolites in Soil Ecology*; 2008; Vol. 14.

(9)   Dinan, L. *Natural Products Isolation*; 1998. https://doi.org/10.1007/978-1-59259-256-2_10.

(10)   Dias, D. A.; Urban, S.; Roessner, U. A Historical Overview of Natural Products in Drug Discovery. *Metabolites* **2012**, *2* (4), 303–336. https://doi.org/10.3390/metabo2020303.

(11)   Medema, M. H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M. A.; Weber, T.; Takano, E.; Breitling, R. AntiSMASH: Rapid Identification, Annotation and Analysis of Secondary Metabolite Biosynthesis Gene Clusters in Bacterial and Fungal Genome Sequences. *Nucleic Acids Res.* **2011**, *39* (Web Server issue), W339-46. https://doi.org/10.1093/nar/gkr466.

(12)   Weissman, K. J. Introduction to Polyketide Biosynthesis. *Methods Enzymol.* **2009**, *459*, 3–16. https://doi.org/10.1016/S0076-6879(09)04601-1.

(13)   Berkov, S.; Mutafova, B.; Christen, P. Molecular Biodiversity and Recent Analytical Developments: A Marriage of Convenience. *Biotechnol. Adv.* **2014**, *32* (6), 1102–1110. https://doi.org/10.1016/j.biotechadv.2014.04.005.

(14)   Michael, C. A.; Dominey-Howes, D.; Labbate, M. The Antimicrobial Resistance Crisis: Causes, Consequences, and Management. *Front. Public Heal.* **2014**, *2* (SEP), 1–8. https://doi.org/10.3389/fpubh.2014.00145.

(15)   Ersoy, S. C.; Heithoff, D. M.; Barnes, L.; Tripp, G. K.; House, J. K.; Marth, J. D.; Smith, J. W.; Mahan, M. J. Correcting a Fundamental Flaw in the Paradigm for Antimicrobial Susceptibility Testing. *EBioMedicine* **2017**, *20*, 173–181.

https://doi.org/10.1016/j.ebiom.2017.05.026.

(16) Martin, J. K.; Wilson, M. Z.; Moore, G. M.; Sheehan, J. P.; Mateus, A.; Li, S. H.-J.; Bratton, B. P.; Kim, H.; Rabinowitz, J. D.; Typas, A.; et al. A Dual-Mechanism Antibiotic Targets Gram-Negative Bacteria and Avoids Drug Resistance. *bioRxiv* **2020**, 2020.03.12.984229. https://doi.org/10.1101/2020.03.12.984229.

(17) Kalmokoff, M. L.; Cyr, T. D.; Hefford, M. A.; Whitford, M. F.; Teather, R. M. Butyrivibriocin AR10, a New Cyclic Bacteriocin Produced by the Ruminal Anaerobe Butyrivibrio Fibrisolvens AR10: Characterization of the Gene and Peptide. *Can. J. Microbiol.* **2003**, *49* (12), 763–773. https://doi.org/10.1139/w03-101.

(18) Kalmokoff, M. L.; Lu, D.; Whitford, M. F.; Teather, R. M. Evidence for Production of a New Lantibiotic (Butyrivibriocin OR79A) by the Ruminal Anaerobe Butyrivibrio Fibrisolvens OR79: Characterization of the Structural Gene Encoding Butyrivibriocin OR79A. *Appl. Environ. Microbiol.* **1999**, *65* (5), 2128–2135. https://doi.org/10.1128/aem.65.5.2128-2135.1999.

(19) Oyama, L. B.; Girdwood, S. E.; Cookson, A. R.; Fernandez-Fuentes, N.; Privé, F.; Vallin, H. E.; Wilkinson, T. J.; Golyshin, P. N.; Golyshina, O. V.; Mikut, R.; et al. The Rumen Microbiome: An Underexplored Resource for Novel Antimicrobial Discovery. *npj Biofilms Microbiomes* **2017**, *3* (1), 1–10. https://doi.org/10.1038/s41522-017-0042-1.

(20) Bolites, M.; Vining, L. C. FUNCTIONS OF SECONDARY Structural Complexity of Secondary Metabolites Waste Products. *Annu. Rev. Microbiol* **1990**, *44*, 395–427.

(21) Keller, N. P.; Turner, G.; Bennett, J. W. Fungal Secondary Metabolism - from Biochemistry to Genomics. *Nat. Rev. Microbiol.* **2005**, *3* (12), 937–947.

https://doi.org/10.1038/nrmicro1286.

(22)   Peralta-Yahya, P. P.; Zhang, F.; del Cardayre, S. B.; Keasling, J. D. Microbial
        Engineering for the Production of Advanced Biofuels. *Nature* **2012**, *488* (7411), 320–
        328. https://doi.org/10.1038/nature11478.

(23)   Grigoriev, I. V.; Nikitin, R.; Haridas, S.; Kuo, A.; Ohm, R.; Otillar, R.; Riley, R.;
        Salamov, A.; Zhao, X.; Korzeniewski, F.; et al. MycoCosm Portal: Gearing up for
        1000 Fungal Genomes. *Nucleic Acids Res.* **2014**, *42* (D1), 699–704.
        https://doi.org/10.1093/nar/gkt1183.

(24)   Jenke-Kodama, H.; Sandmann, A.; Müller, R.; Dittmann, E. Evolutionary
        Implications of Bacterial Polyketide Synthases. *Mol. Biol. Evol.* **2005**, *22* (10), 2027–
        2039. https://doi.org/10.1093/molbev/msi193.

(25)   Li, Y. F.; Tsai, K. J. S.; Harvey, C. J. B.; Li, J. J.; Ary, B. E.; Berlew, E. E.; Boehman,
        B. L.; Findley, D. M.; Friant, A. G.; Gardner, C. A.; et al. Comprehensive Curation
        and Analysis of Fungal Biosynthetic Gene Clusters of Published Natural Products.
        Fungal Genet. Biol. **2016**, *89*, 18–28. https://doi.org/10.1016/j.fgb.2016.01.012.

(26)   Kubicek, C.; Druzhinina, I. *Environmental and Microbial Relationships. THE
        MYCOTA IV.*; 2007; Vol. 4.

(27)   Gilmore, S. P.; Henske, J. K.; O'Malley, M. A. Driving Biomass Breakdown through
        Engineered Cellulosomes. *Bioengineered* **2015**, *6* (4), 204–208.
        https://doi.org/10.1080/21655979.2015.1060379.

(28)   Solomon, K. V; Haitjema, C. H.; Thompson, D. A.; Malley, M. A. O. Extracting Data
        from the Muck : Deriving Biological Insight from Complex Microbial Communities
        and Non-Model Organisms with next Generation Sequencing. *Curr. Opin.*

Biotechnol. **2014**, *28*, 103–110. https://doi.org/10.1016/j.copbio.2014.01.007.

(29)   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. Transcriptomic Characterization of Caecomyces Churrovis: A Novel, Non-
       Rhizoid-Forming Lignocellulolytic Anaerobic Fungus. *Biotechnol. Biofuels* **2017**, *10*
       (1), 305. https://doi.org/10.1186/s13068-017-0997-4.

(30)   Letzel, A.-C.; Pidot, S. J.; Hertweck, C. A Genomic Approach to the Cryptic
       Secondary Metabolome of the Anaerobic World. *Nat. Prod. Rep.* **2013**, *30* (3), 392–
       428. https://doi.org/10.1039/c2np20103h.

(31)   Stewart, R. D.; Auffret, M. D.; Warr, A.; Walker, A. W.; Roehe, R.; Watson, M.
       Compendium of 4,941 Rumen Metagenome-Assembled Genomes for Rumen
       Microbiome Biology and Enzyme Discovery. *Nat. Biotechnol.* **2019**, *37* (8), 953–961.
       https://doi.org/10.1038/s41587-019-0202-3.

(32)   Deduke, C.; Timsina, B.; Piercey-normore, M. D. Effect of Environmental Change on
       Secondary Metabolite Production in Lichen-Forming Fungi. *Internatinal Perpectives
       Glob. Environ. Chang.* **2012**, No. 1983, 197–230. https://doi.org/10.5772/26954.

(33)   Drew, S. W.; Demain, a L. Effect of Primary Metabolites on Secondary Metabolism.
       Annu. Rev. Microbiol. **1977**, *31*, 343–356.
       https://doi.org/10.1146/annurev.mi.31.100177.002015.

(34)   Amos, G. C. A.; Awakawa, T.; Tuttle, R. N.; Letzel, A.-C.; Kim, M. C.; Kudo, Y.;
       Fenical, W.; S. Moore, B.; Jensen, P. R. Comparative Transcriptomics as a Guide to
       Natural Product Discovery and Biosynthetic Gene Cluster Functionality. *Proc. Natl.
       Acad. Sci.* **2017**, *114* (52), E11121–E11130.

https://doi.org/10.1073/pnas.1714381115.

(35) Doull, J. L.; Singh, A. K.; Hoare, M.; Ayer, S. W. Conditions for the Production of Jadomycin B by Streptomyces Venezuelae ISP5230: Effects of Heat Shock, Ethanol Treatment and Phage Infection. *J. Ind. Microbiol.* **1994**, *13* (2), 120–125. https://doi.org/10.1007/BF01584109.

(36) Doull, J. L.; Ayer, S. W.; Singh, A. K.; Thibault, P. Production of a Novel Polyketide Antibiotic, Jadomycin B, by Streptomyces Venezuelae Following Heat Shock. *J. Antibiot. (Tokyo).* **1993**, *46* (5), 869–871. https://doi.org/10.7164/antibiotics.46.869.

(37) Schroeckh, V.; Scherlach, K.; Nützmann, H.-W.; Shelest, E.; Schmidt-Heck, W.; Schuemann, J.; Martin, K.; Hertweck, C.; Brakhage, A. A. Intimate Bacterial-Fungal Interaction Triggers Biosynthesis of Archetypal Polyketides in Aspergillus Nidulans. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (34), 14558–14563. https://doi.org/10.1073/pnas.0901870106.

(38) Lawrence, J. Selfish Operons: The Evolutionary Impact of Gene Clustering in Prokaryotes and Eukaryotes. *Curr. Opin. Genet. Dev.* **1999**, *9* (6), 642–648. https://doi.org/10.1016/S0959-437X(99)00025-8.

(39) Rokas, A.; Wisecaver, J. H.; Lind, A. L. The Birth, Evolution and Death of Metabolic Gene Clusters in Fungi. *Nat. Rev. Microbiol.* **2018**. https://doi.org/10.1038/s41579-018-0075-3.

(40) Zhou, H.; Qiao, K.; Gao, Z.; Meehan, M. J.; Li, J. W. H.; Zhao, X.; Dorrestein, P. C.; Vederas, J. C.; Tang, Y. Enzymatic Synthesis of Resorcylic Acid Lactones by Cooperation of Fungal Iterative Polyketide Synthases Involved in Hypothemycin Biosynthesis. *J. Am. Chem. Soc.* **2010**, *132* (13), 4530–4531.

https://doi.org/10.1021/ja100060k.

(41)   Walsh, C. T.; Fischbach, M. A. Natural Products Version 2.0: Connecting Genes to
        Molecules. *Journal of the American Chemical Society*. American Chemical Society
        2010, pp 2469–2493. https://doi.org/10.1021/ja909118a.

(42)   Katz, L. The DEBS Paradigm for Type I Modular Polyketide Synthases and Beyond.
        Methods Enzymol. **2009**, *459*, 113–142. https://doi.org/10.1016/S0076-
        6879(09)04606-0.

(43)   Minowa, Y.; Araki, M.; Kanehisa, M. Comprehensive Analysis of Distinctive
        Polyketide and Nonribosomal Peptide Structural Motifs Encoded in Microbial
        Genomes. *J. Mol. Biol.* **2007**, *368* (5), 1500–1517.
        https://doi.org/10.1016/j.jmb.2007.02.099.

(44)   Cox, R. J. Polyketides, Proteins and Genes in Fungi: Programmed Nano-Machines
        Begin to Reveal Their Secrets. *Org. Biomol. Chem.* **2007**, *5* (13), 2010–2026.
        https://doi.org/10.1039/b704420h.

(45)   Li, Y.; Zhong, Z.; Hou, P.; Zhang, W.-P.; Qian, P.-Y. Resistanc to Nonribosomal
        Peptide Antibiotics Mediated by D-Stereospecific Peptidases. *Nat. Chem. Biol.* **2018**,
        14 (April), 381–387.

(46)   Caboche, S.; Pupin, M.; Leclère, V.; Fontaine, A.; Jacques, P.; Kucherov, G.
        NORINE: A Database of Nonribosomal Peptides. *Nucleic Acids Res.* **2008**, *36*
        (SUPPL. 1), 326–331. https://doi.org/10.1093/nar/gkm792.

(47)   Marahiel, M. A.; Essen, L.-O. Chapter 13. Nonribosomal Peptide Synthetases
        Mechanistic and Structural Aspects of Essential Domains. *Methods Enzymol.* **2009**,
        458, 337–351. https://doi.org/10.1016/S0076-6879(09)04813-7.

(48) Gaudelli, N. M.; Long, D. H.; Townsend, C. A. β-Lactam Formation by a Non-Ribosomal Peptide Synthetase during Antibiotic Biosynthesis. *Nature* **2015**, *520* (7547), 383–387. https://doi.org/10.1038/nature14100.

(49) Keller, N. P. Fungal Secondary Metabolism: Regulation, Function and Drug Discovery. *Nat. Rev. Microbiol.* **2018**, *In press*. https://doi.org/10.1038/s41579-018-0121-1.

(50) Van Heel, A. J.; De Jong, A.; Song, C.; Viel, J. H.; Kok, J.; Kuipers, O. P. BAGEL4: A User-Friendly Web Server to Thoroughly Mine RiPPs and Bacteriocins. *Nucleic Acids Res.* **2018**, *46* (W1), W278–W281. https://doi.org/10.1093/nar/gky383.

(51) Ortega, M. A.; Van Der Donk, W. A. New Insights into the Biosynthetic Logic of Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products. *Cell Chem. Biol.* **2016**, *23* (1), 31–44. https://doi.org/10.1016/j.chembiol.2015.11.012.

(52) Goto, Y.; Li, B.; Claesen, J.; Shi, Y.; Bibb, M. J.; van der Donk, W. A. Discovery of Unique Lanthionine Synthetases Reveals New Mechanistic and Evolutionary Insights. *PLoS Biol.* **2010**, *8* (3), 4–13. https://doi.org/10.1371/journal.pbio.1000339.

(53) Letzel, A. C.; Pidot, S. J.; Hertweck, C. Genome Mining for Ribosomally Synthesized and Post-Translationally Modified Peptides (RiPPs) in Anaerobic Bacteria. *BMC Genomics* **2014**, *15* (1), 1–21. https://doi.org/10.1186/1471-2164-15-983.

(54) Lubelski, J.; Rink, R.; Khusainov, R.; Moll, G. N.; Kuipers, O. P. Biosynthesis, Immunity, Regulation, Mode of Action and Engineering of the Model Lantibiotic Nisin. *Cell. Mol. Life Sci.* **2008**, *65* (3), 455–476. https://doi.org/10.1007/s00018-007-7171-2.

(55)   Kodani, S.; Hudson, M. E.; Durrant, M. C.; Buttner, M. J.; Nodwell, J. R.; Willey, J. M. The SapB Morphogen Is a Lantibiotic-like Peptide Derived from the Product of the Developmental Gene RamS in Streptomyces Coelicolor. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (31), 11448–11453. https://doi.org/10.1073/pnas.0404220101.

(56)   Kodani, S.; Lodato, M. A.; Durrant, M. C.; Picart, F.; Willey, J. M. SapT, a Lanthionine-Containing Peptide Involved in Aerial Hyphae Formation in the Streptomycetes. *Mol. Microbiol.* **2005**, *58* (5), 1368–1380. https://doi.org/10.1111/j.1365-2958.2005.04921.x.

(57)   Mary C. Rea, R. Paul Ross, Paul D. Cotter, and C. H. Classification of Bacteriocins From Gram-Positive Bacteria. In *Prokaryotic Antimicrobial Peptides*; Springer New York: New York, NY, 2011; pp 29–53. https://doi.org/10.1007/978-1-4419-7692-5_20.

(58)   Arnison, P. G.; Bibb, M. J.; Bierbaum, G.; Bowers, A. A.; Bugni, T. S.; Bulaj, G.; Camarero, J. A.; Campopiano, D. J.; Challis, G. L.; Clardy, J.; et al. Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Products: Overview and Recommendations for a Universal Nomenclature. *Nat. Prod. Rep.* **2013**, *30* (1), 108–160. https://doi.org/10.1039/c2np20085f.

(59)   Thennarasu, S.; Lee, D. K.; Poon, A.; Kawulka, K. E.; Vederas, J. C.; Ramamoorthy, A. Membrane Permeabilization, Orientation, and Antimicrobial Mechanism of Subtilosin A. *Chem. Phys. Lipids* **2005**, *137* (1–2), 38–51. https://doi.org/10.1016/j.chemphyslip.2005.06.003.

(60)   Hashimoto, M.; Murakami, T.; Funahashi, K.; Tokunaga, T.; Nihei, K. ichi; Okuno, T.; Kimura, T.; Naoki, H.; Himeno, H. An RNA Polymerase Inhibitor,

Cyclothiazomycin B1, and Its Isomer. *Bioorganic Med. Chem.* **2006**, *14* (24), 8259–8270. https://doi.org/10.1016/j.bmc.2006.09.006.

(61) Mizuhara, N.; Kuroda, M.; Ogita, A.; Tanaka, T.; Usuki, Y.; Fujita, K. I. Antifungal Thiopeptide Cyclothiazomycin B1 Exhibits Growth Inhibition Accompanying Morphological Changes via Binding to Fungal Cell Wall Chitin. *Bioorganic Med. Chem.* **2011**, *19* (18), 5300–5310. https://doi.org/10.1016/j.bmc.2011.08.010.

(62) Holmes, D. J.; Caso, J. L.; Thompson, C. J. Autogenous Transcriptional Activation of a Thiostrepton-Induced Gene in Streptomyces Lividans. *EMBO J.* **1993**, *12* (8), 3183–3191. https://doi.org/10.1002/j.1460-2075.1993.tb05987.x.

(63) Lin, P. F.; Samanta, H.; Bechtold, C. M.; Deminie, C. A.; Patick, A. K.; Alam, M.; Riccardi, K.; Rose, R. E.; White, R. J.; Colonno, R. J. Characterization of Siamycin I, a Human Immunodeficiency Virus Fusion Inhibitor. *Antimicrob. Agents Chemother.* **1996**, *40* (1), 133–138. https://doi.org/10.1128/aac.40.1.133.

(64) Iwatsuki, M.; Uchida, R.; Takakusagi, Y.; Matsumoto, A.; Jiang, C. L.; Takahashi, Y.; Arai, M.; Kobayashi, S.; Matsumoto, M.; Inokoshi, J.; et al. Lariatins, Novel Anti-Mycobacterial Peptides with a Lasso Structure, Produced by Rhodococcus Jostii K01-B0171. *J. Antibiot. (Tokyo).* **2007**, *60* (6), 357–363. https://doi.org/10.1038/ja.2007.48.

(65) Wilson, K. A.; Kalkum, M.; Ottesen, J.; Yuzenkova, J.; Chait, B. T.; Landick, R.; Muir, T.; Severinov, K.; Darst, S. A. Structure of Microcin J25, a Peptide Inhibitor of Bacterial RNA Polymerase, Is a Lassoed Tail. *J. Am. Chem. Soc.* **2003**, *125* (41), 12475–12483. https://doi.org/10.1021/ja036756q.

(66) Rosengren, K. J.; Clark, R. J.; Daly, N. L.; Göransson, U.; Jones, A.; Craik, D. J.

Microcin J25 Has a Threaded Sidechain-to-Backbone Ring Structure and Not a Head-to-Tail Cyclized Backbone. *J. Am. Chem. Soc.* **2003**, *125* (41), 12464–12474. https://doi.org/10.1021/ja0367703.

(67)   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.; et al. Cultivation and Sequencing of Rumen Microbiome Members from the Hungate1000 Collection. *Nat. Biotechnol.* **2018**, 1–27. https://doi.org/10.1038/nbt.4110.

(68)   Hetrick, K. J.; van der Donk, W. A. Ribosomally Synthesized and Post-Translationally Modified Peptide Natural Product Discovery in the Genomic Era. *Curr. Opin. Chem. Biol.* **2017**, *38*, 36–44. https://doi.org/10.1016/j.cbpa.2017.02.005.

(69)   Keller, N. P.; Hohn, T. M. Metabolic Pathway Gene Clusters in Filamentous Fungi. *Fungal Genet. Biol.* **1997**, *21* (1), 17–29. https://doi.org/10.1006/fgbi.1997.0970.

(70)   Wisecaver, J. H.; Slot, J. C.; Rokas, A. The Evolution of Fungal Metabolic Pathways. *PLoS Genet.* **2014**, *10* (12). https://doi.org/10.1371/journal.pgen.1004816.

(71)   Kominek, J.; Doering, D. T.; Opulente, D. A.; Shen, X. X.; Zhou, X.; DeVirgilio, J.; Hulfachor, A. B.; Groenewald, M.; Mcgee, M. A.; Karlen, S. D.; et al. Eukaryotic Acquisition of a Bacterial Operon. *Cell* **2019**, *176* (6), 1356-1366.e10. https://doi.org/10.1016/j.cell.2019.01.034.

(72)   Nora Khldi, Jerome Collemare, Marc-Henri Lebrun, K. H. W. Evidence for Horizontal Transfer of a Secondary Metabolite Gene Cluster between Fungi. *Genome Biol.* **2008**, *9* (1). https://doi.org/10.1186/gb-2008-9-1-r18.

(73)   Schmitt, I.; Lumbsch, H. T. Ancient Horizontal Gene Transfer from Bacteria Enhances Biosynthetic Capabilities of Fungi. *PLoS One* **2009**, *4* (2), e4437.

https://doi.org/10.1371/journal.pone.0004437.

(74) Landan, G.; Cohen, G.; Aharonowitz, Y.; Shuali, Y.; Graur, D.; Shiffman, D.
Evolution of Isopenicillin N Synthase Genes May Have Involved Horizontal Gene
Transfer. *Mol. Biol. Evol.* **1990**, *7* (5), 399–406.

(75) Kroken, S.; Glass, N. L.; Taylor, J. W.; Yoder, O. C.; Turgeon, B. G. Phylogenomic
Analysis of Type I Polyketide Synthase Genes in Pathogenic and Saprobic
Ascomycetes. *Proc. Natl. Acad. Sci.* **2003**, *100* (26), 15670–15675.
https://doi.org/10.1073/pnas.2532165100.

(76) Shwab, E. K.; Keller, N. P. Regulation of Secondary Metabolite Production in
Filamentous Ascomycetes. *Mycol. Res.* **2008**, *112* (Pt 2), 225–230.
https://doi.org/10.1016/j.mycres.2007.08.021.

(77) Bussink, H. J.; Clark, A.; Oliver, R. The Cladosporium Fulvum Bap1 Gene: Evidence
for a Novel Class of Yap-Related Transcription Factors with Ankyrin Repeats in
Phytopathogenic Fungi. *Eur. J. Plant Pathol.* **2001**, *107* (6), 655–659.
https://doi.org/10.1023/A:1017932721007.

(78) Pedley, K. F.; Walton, J. D. Regulation of Cyclic Peptide Biosynthesis in a Plant
Pathogenic Fungus by a Novel Transcription Factor. *Proc. Natl. Acad. Sci.* **2001**, *98*
(24), 14174–14179. https://doi.org/10.1073/pnas.231491298.

(79) Niehaus, E. M.; Janevska, S.; Von Bargen, K. W.; Sieber, C. M. K.; Harrer, H.;
Humpf, H. U.; Tudzynski, B. Apicidin F: Characterization and Genetic Manipulation
of a New Secondary Metabolite Gene Cluster in the Rice Pathogen Fusarium
Fujikuroi. *PLoS One* **2014**, *9* (7). https://doi.org/10.1371/journal.pone.0103336.

(80) Tilburn, J.; Sarkar, S.; Widdick, D. A.; Espeso, E. A.; Orejas, M.; Mungroo, J.;

Peñalva, M. A.; Arst, H. N. The Aspergillus PacC Zinc Finger Transcription Factor Mediates Regulation of Both Acid- and Alkaline-Expressed Genes by Ambient PH. *EMBO J.* **1995**, *14* (4), 779–790.

(81)   Zhibing Luo, Hui Ren, Jarrod J. Mousa, Drauzio Range, Yongjun Zhang, Steven D. Bruner, and Nemat O. Keyhani1ing Luo, Hui Ren, Jarrod J. Mousa, Drauzio Rangel, Yongjun Zhang, Steven D. Bruner, and N. O. K. The PacC Transcription Factor Regulates Secondary Metabolite Production, Stress Response, but Has Only Minor Effects on Virulence in the Insect Pathogenic Fungus Beauveria Bassiana. **2017**, *00*, 1–42. https://doi.org/10.1111/1462-2920.13648.

(82)   Dowzer, C. E.; Kelly, J. M. Cloning of the CreA Gene from Aspergillus Nidulans: A Gene Involved in Carbon Catabolite Repression. *Curr. Genet.* **1989**, *15* (6), 457–459. https://doi.org/10.1007/BF00376804.

(83)   Hynes, M. J. Studies on the Role of the Area Gene in the Regulation of Nitrogen Catabolism in Aspergillus Nidulans. *Aust. J. Biol. Sci.* **1975**, *28* (3), 301–313. https://doi.org/10.1071/BI9750301.

(84)   Marion Brodhagen, N. P. K. Signalling Pathways Connecting Mycotoxin Production and Sporulation. *Mol. Plant Pathol.* **2006**, *7* (4), 285–301. https://doi.org/10.1111/j.1364-3703.2006.00338.x.

(85)   Takano, Y.; Kikuchi, T.; Kubo, Y.; Hamer, J. E.; Mise, K.; Furusawa, I. The Colletotrichum Lagenarium MAP Kinase Gene CMK1 Regulates Diverse Aspects of Fungal Pathogenesis. *Mol. Plant-Microbe Interact.* **2000**, *13* (4), 374–383. https://doi.org/10.1094/MPMI.2000.13.4.374.

(86)   Shim, W. B.; Woloshuk, C. P. Regulation of Fumonisin B1 Biosynthesis and

Conidiation in Fusarium Verticillioides by a Cyclin-Like (C-Type) Gene, FCC1 †.
Appl. Environ. Microbiol. **2001**, *67* (4), 1607–1612.
https://doi.org/10.1128/AEM.67.4.1607-1612.2001.

(87)   Bok, J. W.; Keller, N. P. LaeA, a Regulator of Secondary Metabolism in Aspergillus
Spp. *Eukaryot. Cell* **2004**, *3* (2), 527–535. https://doi.org/10.1128/EC.3.2.527-
535.2004.

(88)   Butchko, R. A. E.; Brown, D. W.; Busman, M.; Tudzynski, B.; Wiemann, P. Lae1
Regulates Expression of Multiple Secondary Metabolite Gene Clusters in Fusarium
Verticillioides. *Fungal Genet. Biol.* **2012**, *49* (8), 602–612.
https://doi.org/10.1016/j.fgb.2012.06.003.

(89)   Perrin, R. M.; Fedorova, N. D.; Bok, J. W.; Cramer, R. A.; Wortman, J. R.; Kim, H.
S.; Nierman, W. C.; Keller, N. P. Transcriptional Regulation of Chemical Diversity in
Aspergillus Fumigatus by LaeA. *PLoS Pathog.* **2007**, *3* (4), e50.
https://doi.org/10.1371/journal.ppat.0030050.

(90)   Bayram, Ö.; Krappmann, S.; Ni, M.; Bok, J. W.; Helmstaedt, K.; Yu, J.; Braus, G. H.
VelB/VeA/LaeA Complex Coordinates Light Signal with Fungal Development and
Secondary Metabolism. *Science (80-.).* **2008**, *320* (June), 1504–1507.

(91)   Bayram, Ö.; Braus, G. H. Coordination of Secondary Metabolism and Development
in Fungi: The Velvet Family of Regulatory Proteins. *FEMS Microbiol. Rev.* **2012**, *36*
(1), 1–24. https://doi.org/10.1111/j.1574-6976.2011.00285.x.

(92)   Kato, N.; Brooks, W.; Calvo, A. M. The Expression of Sterigmatocystin and
Penicillin Genes in Aspergillus Nidulans Is Controlled by VeA, a Gene Required for
Sexual Development. *Am. Soc. Microbiol.* **2003**, *2* (6), 1178–1186.

https://doi.org/10.1128/EC.2.6.1178–1186.2003.

(93)   Jenuwein, T.; Allis, C. D. Translating the Histone Code. *Science (80-.).* **2001**, *293* (5532), 1074–1080. https://doi.org/10.1126/science.1063127.

(94)   Freitas, L. H.; Hernandez-Rivas, R.; Ralph, S. A.; Montiel-Condado, D.; Ruvalcaba-Salazar, O. K.; Rojas-Meza, A. P.; Mâncio-Silva, L.; Leal-Silvestre, R. J.; Gontijo, A. M.; Shorte, S.; et al. Telomeric Heterochromatin Propagation and Histone Acetylation Control Mutually Exclusive Expression of Antigenic Variation Genes in Malaria Parasites. *Cell* **2005**, *121* (1), 25–36. https://doi.org/10.1016/j.cell.2005.01.037.

(95)   Horn, D.; Barry, J. D. The Central Roles of Telomeres and Subtelomeres in Antigenic Variation in African Trypanosomes. *Chromosom. Res.* **2005**, *13* (5), 525–533. https://doi.org/10.1007/s10577-005-0991-8.

(96)   Ralph, S. A.; Scherf, A. The Epigenetic Control of Antigenic Variation in Plasmodium Falciparum. *Curr. Opin. Microbiol.* **2005**, *8* (4), 434–440. https://doi.org/10.1016/j.mib.2005.06.007.

(97)   Reyes-Dominguez, Y.; Bok, J. W.; Berger, H.; Shwab, E. K.; Basheer, A.; Gallmetzer, A.; Scazzocchio, C.; Keller, N.; Strauss, J. Heterochromatic Marks Are Associated with the Repression of Secondary Metabolism Clusters in Aspergillus Nidulans. *Mol. Microbiol.* **2010**, *76* (6), 1376–1386. https://doi.org/10.1111/j.1365-2958.2010.07051.x.

(98)   Reyes-Dominguez, Y.; Boedi, S.; Sulyok, M.; Wiesenberger, G.; Stoppacher, N.; Krska, R.; Strauss, J. Heterochromatin Influences the Secondary Metabolite Profile in the Plant Pathogen Fusarium Graminearum. *Fungal Genet. Biol.* **2012**, *49* (1), 39–47. https://doi.org/10.1016/j.fgb.2011.11.002.

(99)   Cichewicz, R. H. Epigenome Manipulation as a Pathway to New Natural Product

Scaffolds and Their Congeners. *Nat. Prod. Rep.* **2010**, *27* (1), 11–22.

https://doi.org/10.1039/b920860g.

(100) Ziemert, N.; Lechner, A.; Wietz, M.; Millan-Aguinaga, N.; Chavarria, K. L.; Jensen,

P. R. Diversity and Evolution of Secondary Metabolism in the Marine Actinomycete

Genus Salinispora. *Proc. Natl. Acad. Sci.* **2014**, *111* (12), E1130–E1139.

https://doi.org/10.1073/pnas.1324161111.

(101) Nothias, L. F.; Nothias-Esposito, M.; Da Silva, R.; Wang, M.; Protsyuk, I.; Zhang, Z.;

Sarvepalli, A.; Leyssen, P.; Touboul, D.; Costa, J.; et al. Bioactivity-Based Molecular

Networking for the Discovery of Drug Leads in Natural Product Bioassay-Guided

Fractionation. *J. Nat. Prod.* **2018**, *81* (4), 758–767.

https://doi.org/10.1021/acs.jnatprod.7b00737.

(102) von Eckardstein, L.; Petras, D.; Dang, T.; Cociancich, S.; Sabri, S.; Grätz, S.; Kerwat,

D.; Seidel, M.; Pesic, A.; Dorrestein, P. C.; et al. Total Synthesis and Biological

Assessment of Novel Albicidins Discovered by Mass Spectrometric Networking.

Chem. - A Eur. J. **2017**, *23* (61), 15316–15321.

https://doi.org/10.1002/chem.201704074.

(103) Cavill, R.; Jennen, D.; Kleinjans, J.; Briedé, J. J. Transcriptomic and Metabolomic

Data Integration. *Brief. Bioinform.* **2016**, *17* (5), 891–901.

https://doi.org/10.1093/bib/bbv090.

(104) Hautbergue, T.; Jamin, E. L.; Debrauwer, L.; Puel, O.; Oswald, I. P. From Genomics

to Metabolomics, Moving toward an Integrated Strategy for the Discovery of Fungal

Secondary Metabolites. *Nat. Prod. Rep.* **2018**, *35* (2), 147–173.

https://doi.org/10.1039/c7np00032d.

(105) Blin, K.; Medema, M. H.; Kazempour, D.; Fischbach, M. A.; Breitling, R.; Takano, E.; Weber, T. AntiSMASH 2.0--a Versatile Platform for Genome Mining of Secondary Metabolite Producers. *Nucleic Acids Res.* **2013**, *41* (Web Server issue), W204-12. https://doi.org/10.1093/nar/gkt449.

(106) Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Bruccoleri, R.; Lee, S. Y.; Fischbach, M. A.; Muller, R.; Wohlleben, W.; et al. AntiSMASH 3.0--a Comprehensive Resource for the Genome Mining of Biosynthetic Gene Clusters. *Nucleic Acids Res.* **2015**, W237–W243. https://doi.org/10.1093/nar/gkv437.

(107) Blin, K.; Wolf, T.; Chevrette, M. G.; Lu, X.; Schwalen, C. J.; Kautsar, S. A.; Suarez Duran, H. G.; de los Santos, E. L. C.; Kim, H. U.; Nave, M.; et al. AntiSMASH 4.0—Improvements in Chemistry Prediction and Gene Cluster Boundary Identification. *Nucleic Acids Res.* **2017**, *1854* (1), 1019–1037. https://doi.org/10.1093/nar/gkx319.

(108) Blin, K.; Shaw, S.; Steinke, K.; Villebro, R.; Ziemert, N.; Lee, S. Y.; Medema, M. H.; Weber, T. AntiSMASH 5.0: Updates to the Secondary Metabolite Genome Mining Pipeline. *Nucleic Acids Res.* **2019**, *47* (W1), W81–W87. https://doi.org/10.1093/nar/gkz310.

(109) Khaldi, N.; Seifuddin, F. T.; Turner, G.; Haft, D.; Nierman, W. C.; Wolfe, K. H.; Fedorova, N. D. SMURF: Genomic Mapping of Fungal Secondary Metabolite Clusters. *Fungal Genet. Biol.* **2010**, *47* (9), 736–741. https://doi.org/10.1016/j.fgb.2010.06.003.

(110) Traxler, M. F.; Watrous, J. D.; Alexandrov, T.; Dorrestein, P. C.; Kolter, R.

Interspecies Interactions Stimulate Diversification of the Streptomyces Coelicolor Secreted Metabolome. *MBio* **2013**, *4* (4), 1–12. https://doi.org/10.1128/mBio.00459-13.

(111) König, C. C.; Scherlach, K.; Schroeckh, V.; Horn, F.; Nietzsche, S.; Brakhage, A. A.; Hertweck, C. Bacterium Induces Cryptic Meroterpenoid Pathway in the Pathogenic Fungus Aspergillus Fumigatus. *ChemBioChem* **2013**, *14* (8), 938–942. https://doi.org/10.1002/cbic.201300070.

(112) Bertrand, S.; Bohni, N.; Schnee, S.; Schumpp, O.; Gindro, K.; Wolfender, J. L. Metabolite Induction via Microorganism Co-Culture: A Potential Way to Enhance Chemical Diversity for Drug Discovery. *Biotechnology Advances*. 2014, pp 1180–1204. https://doi.org/10.1016/j.biotechadv.2014.03.001.

(113) Luti, K. J. K.; Mavituna, F. Elicitation of Streptomyces Coelicolor with Dead Cells of Bacillus Subtilis and Staphylococcus Aureus in a Bioreactor Increases Production of Undecylprodigiosin. *Appl. Microbiol. Biotechnol.* **2011**, *90* (2), 461–466. https://doi.org/10.1007/s00253-010-3032-2.

(114) Abdelmohsen, U. R.; Grkovic, T.; Balasubramanian, S.; Kamel, M. S.; Quinn, R. J.; Hentschel, U. Elicitation of Secondary Metabolism in Actinomycetes. *Biotechnol. Adv.* **2015**, *33* (6), 798–811. https://doi.org/10.1016/j.biotechadv.2015.06.003.

(115) Zuck, K. M.; Shipley, S.; Newman, D. J. Induced Production of N-Formyl Alkaloids from Aspergillus Fumigatus by Co-Culture with Streptomyces Peucetius. *J. Nat. Prod.* **2011**, *74* (7), 1653–1657. https://doi.org/10.1021/np200255f.

(116) Bauchop, T.; Mountfort, D. O. Cellulose Fermentation by a Rumen Anaerobic Fungus in Both the Absence and the Presence of Rumen Methanogens. *Appl. Environ.*

Microbiol. **1981**, *42* (6), 1103–1110.

(117) K.N. Joblin, Gillian P. Campbell, A.J. Richardson, C. S. S. Fermentation of Barley
Straw by Anaerobic Rumen Bacteria and Fungi in Axenic Culture and in Co-Culture
with Methanogens. *Lett. Appl. Microbiol.* **1989**, *9* (5), 195–197.

(118) Mountfort, D. O.; Asher, R. A.; Bauchop, T. Fermentation of Cellulose to Methane
and Carbon Dioxide by a Rumen Anaerobic Fungus in a Triculture with
Methanobrevibacter Sp. Strain RA1 and Methanosarcina Barkeri. *Appl. Environ.
Microbiol.* **1982**, *44* (1), 128–134.

(119) Teunissen, M. J.; Kets, E. P.; Op den Camp, H. J.; Huis in't Veld, J. H.; Vogels, G. D.
Effect of Coculture of Anaerobic Fungi Isolated from Ruminants and Non-Ruminants
with Methanogenic Bacteria on Cellulolytic and Xylanolytic Enzyme Activities. *Arch
Microbiol* **1992**, *157* (2), 176–182.

(120) Marvin-Sikkema, F. D.; Richardson, A. J.; Stewart, C. S.; Gottschal, J. C.; Prins, R.
A. Influence of Hydrogen-Consuming Bacteria on Cellulose Degradation by
Anaerobic Fungi. *Appl. Environ. Microbiol.* **1990**, *56* (12), 3793–3797.

(121) Bernalier, A.; Fonty, G.; Bonnemoy, F.; Gouet, P. Degradation and Fermentation of
Cellulose by the Rumen Anaerobic Fungi in Axenic Cultures or in Association with
Cellulolytic Bacteria. *Curr. Microbiol.* **1992**, *25* (3), 143–148.
https://doi.org/10.1007/BF01571022.

(122) Bernalier, A.; Fonty, G.; Gouet, P. Cellulose Degradation by Two Rumen Anaerobic
Fungi in Monoculture or in Coculture with Rumen Bacteria. *Anim. Feed Sci. Technol.*
1991, *32* (1–3), 131–136. https://doi.org/10.1016/0377-8401(91)90016-L.

(123) Roger, V.; Bernalier, A.; Grenet, E.; Fonty, G.; Jamot, J.; Gouet, P. Degradation of

Wheat Straw and Maize Stem by a Monocentric and a Polycentric Rumen Fungi, Alone or in Association with Rumen Cellulolytic Bacteria. *Anim. Feed Sci. Technol.* **1993**, *42* (1–2), 69–82. https://doi.org/10.1016/0377-8401(93)90024-E.

(124) Dehority, B. A.; Tirabasso, P. A. Antibiosis between Ruminal Bacteria and Ruminal Fungi. *Appl. Environ. Microbiol.* **2000**, *66* (7), 2921–2927. https://doi.org/10.1128/AEM.66.7.2921-2927.2000.

(125) Joblin, K. N.; Matsui, H.; Naylor, G. E.; Ushida, K. Degradation of Fresh Ryegrass by Methanogenic Co-Cultures of Ruminal Fungi Grown in the Presence or Absence of Fibrobacter Succinogenes. *Curr. Microbiol.* **2002**, *45* (1), 46–53. https://doi.org/10.1007/s00284-001-0078-5.

(126) Bacon, C. W.; White, J. *Microbial Endophytes*.

(127) Mackie, A. E.; Wheatley, R. E. Effects and Incidence of Volatile Organic Compound Interactions between Soil Bacterial and Fungal Isolates. *Soil Biol. Biochem.* **1999**, *31* (3), 375–385. https://doi.org/10.1016/S0038-0717(98)00140-0.

(128) Kellner, H.; Zak, D. R. Detection of Expressed Fungal Type I Polyketide Synthase Genes in a Forest Soil. *Soil Biol. Biochem.* **2009**, *41* (6), 1344–1347. https://doi.org/10.1016/j.soilbio.2009.02.033.

(129) Haas, H. Molecular Genetics of Fungal Siderophore Biosynthesis and Uptake: The Role of Siderophores in Iron Uptake and Storage. *Appl. Microbiol. Biotechnol.* **2003**, *62* (4), 316–330. https://doi.org/10.1007/s00253-003-1335-2.

(130) Sheth, R. U.; Cabral, V.; Chen, S. P.; Wang, H. H. Manipulating Bacterial Communities by in Situ Microbiome Engineering. *Trends Genet.* **2016**, *32* (4), 189–200. https://doi.org/10.1016/j.tig.2016.01.005.

(131) Van Den Ende, W.; Peshev, D.; De Gara, L. Disease Prevention by Natural
Antioxidants and Prebiotics Acting as ROS Scavengers in the Gastrointestinal Tract.
Trends Food Sci. Technol. **2011**, *22* (12), 689–697.
https://doi.org/10.1016/j.tifs.2011.07.005.

(132) Hanke, F. J. Natural Products as a Resource for Biologically Active Compounds.
Pharm. Res. **1986**, *13* (8), 1133–1141.

(133) Wilbon, P. A.; Chu, F.; Tang, C. Progress in Renewable Polymers from Natural
Terpenes, Terpenoids, and Rosin. *Macromol. Rapid Commun.* **2013**, *34* (1), 8–37.
https://doi.org/10.1002/marc.201200513.

(134) Medema, M. H.; Kottmann, R.; Yilmaz, P.; Cummings, M.; Biggins, J. B.; Blin, K.;
De Bruijn, I.; Chooi, Y. H.; Claesen, J.; Coates, R. C.; et al. Minimum Information
about a Biosynthetic Gene Cluster. *Nat. Chem. Biol.* **2015**, *11* (9), 625–631.
https://doi.org/10.1038/nchembio.1890.

(135) Raffa, N.; Keller, N. P. A Call to Arms: Mustering Secondary Metabolites for Success
and Survival of an Opportunistic Pathogen. *PLoS Pathog.* **2019**, *15* (4), 1–9.
https://doi.org/10.1371/journal.ppat.1007606.

(136) Culp, E. J.; Yim, G.; Waglechner, N.; Wang, W.; Pawlowski, A. C.; Wright, G. D.
Hidden Antibiotics in Actinomycetes Can Be Identified by Inactivation of Gene
Clusters for Common Antibiotics. *Nat. Biotechnol. 2*. https://doi.org/10.1038/s41587-
019-0241-9.

(137) Zipperer, A.; Konnerth, M. C.; Laux, C.; Berscheid, A.; Janek, D.; Weidenmaier, C.;
Burian, M.; Schilling, N. A.; Slavetinsky, C.; Marschal, M.; et al. Human
Commensals Producing a Novel Antibiotic Impair Pathogen Colonization. *Nature*

2016, *535* (7613), 511–516. https://doi.org/10.1038/nature18634.

(138) Ling, L. L.; Schneider, T.; Peoples, A. J.; Spoering, A. L.; Engels, I.; Conlon, B. P.; Mueller, A.; Schäberle, T. F.; Hughes, D. E.; Epstein, S.; et al. A New Antibiotic Kills Pathogens without Detectable Resistance. *Nature* **2015**, *517* (7535), 455–459. https://doi.org/10.1038/nature14098.

(139) Orpin, C. G.; Ho, Y. W. Ecology and Function of the Anaerobic Rumen Fungi. In *Recent Advances on the Nutrition of Herbivores*; Serdang, Malaysia, Malaysian Society of Animal Production, 1991; pp 163–170.

(140) Podolsky, I. A.; Seppälä, S.; Lankiewicz, T. S.; Brown, J. L.; Swift, C. L.; O'Malley, M. A. Harnessing Nature's Anaerobes for Biotechnology and Bioprocessing. *Annu. Rev. Chem. Biomol. Eng.* **2019**, *10* (1), annurev-chembioeng-060718-030340. https://doi.org/10.1146/annurev-chembioeng-060718-030340.

(141) Shabuer, G.; Ishida, K.; Pidot, S. J.; Roth, M.; Dahse, H.-M.; Hertweck, C. Plant Pathogenic Anaerobic Bacteria Use Aromatic Polyketides to Access Aerobic Territory. *Science* **2015**, *350* (6261), 670–674. https://doi.org/10.1126/science.aac9990.

(142) Calvo, A. M.; Wilson, R. A.; Bok, J. W.; Keller, N. P. Relationship between Secondary Metabolism and Fungal Development. *Microbiol. Mol. Biol. Rev.* **2002**, *66* (3), 447–459. https://doi.org/10.1128/MMBR.66.3.447-459.2002.

(143) Bärenstrauch, M.; Mann, S.; Jacquemin, C.; Bibi, S.; Sylla, O.-K.; Baudouin, E.; Buisson, D.; Prado, S.; Kunz, C. Molecular Crosstalk between the Endophyte Paraconiothyrium Variabile and the Phytopathogen Fusarium Oxysporum – Modulation of Lipoxygenase Activity and Beauvericin Production during the

Interaction. *Fungal Genet. Biol.* **2020**, *139*, 103383.
https://doi.org/10.1016/J.FGB.2020.103383.

(144) Cimermancic, P.; Medema, M. H.; Claesen, J.; Kurita, K.; Wieland Brown, L. C.;
Mavrommatis, K.; Pati, A.; Godfrey, P. A.; Koehrsen, M.; Clardy, J.; et al. Insights
into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene
Clusters. *Cell* **2014**, *158* (2), 412–421. https://doi.org/10.1016/j.cell.2014.06.034.

(145) Cotter, P. D.; Ross, R. P.; Hill, C. Bacteriocins - a Viable Alternative to Antibiotics?
Nat. Rev. Microbiol. **2013**, *11* (2), 95–105. https://doi.org/10.1038/nrmicro2937.

(146) Wang, G.; Li, X.; Wang, Z. APD3: The Antimicrobial Peptide Database as a Tool for
Research and Education. *Nucleic Acids Res.* **2016**, *44* (D1), D1087–D1093.
https://doi.org/10.1093/nar/gkv1278.

(147) Waghu, F. H.; Barai, R. S.; Idicula-thomas, S. Leveraging Family-Specific Signatures
for AMP Discovery and High-Throughput Annotation. *Nat. Publ. Gr.* **2016**, 1–7.
https://doi.org/10.1038/srep24684.

(148) Martínez, B.; Suárez, J. E.; Rodríguez, A. Lactococcin 972: A Homodimeric
Lactococcal Bacteriocin Whose Primary Target Is Not the Plasma Membrane.
Microbiology **1996**, *142* (9), 2393–2398. https://doi.org/10.1099/00221287-142-9-
2393.

(149) Tran, P. N.; Yen, M. R.; Chiang, C. Y.; Lin, H. C.; Chen, P. Y. Detecting and
Prioritizing Biosynthetic Gene Clusters for Bioactive Compounds in Bacteria and
Fungi. *Appl. Microbiol. Biotechnol.* **2019**, *103* (8), 3277–3287.
https://doi.org/10.1007/s00253-019-09708-z.

(150) Majoros, W. H.; Pertea, M.; Salzberg, S. L. TigrScan and GlimmerHMM: Two Open

Source Ab Initio Eukaryotic Gene-Finders. *Bioinformatics* **2004**, *20* (16), 2878–2879. https://doi.org/10.1093/bioinformatics/bth315.

(151) Lukashin, A.; Borodovsky, M. GeneMark.Hmm: New Solutions for Gene Finding. Nucleic Acids Res. *Nucleic Acids Res.* **1998**, *26* (4), 1107–15.

(152) Ter-Hovhannisyan, V.; Lomsadze, A.; Chernoff, Y. O.; Borodovsky, M. Gene Prediction in Novel Fungal Genomes Using an Ab Initio Algorithm with Unsupervised Training. *Genome Res.* **2008**, *18* (12), 1979–1990. https://doi.org/10.1101/gr.081612.108.

(153) Solovyev, V.; Kosarev, P.; Seledsov, I.; Vorobyev, D. Automatic Annotation of Eukaryotic Genes, Pseudogenes and Promoters. *Genome Biol.* **2006**, *7* (Suppl 1), S10. https://doi.org/10.1186/gb-2006-7-s1-s10.

(154) Jia, L. J.; Tang, H. Y.; Wang, W. Q.; Yuan, T. L.; Wei, W. Q.; Pang, B.; Gong, X. M.; Wang, S. F.; Li, Y. J.; Zhang, D.; et al. A Linear Nonribosomal Octapeptide from Fusarium Graminearum Facilitates Cell-to-Cell Invasion of Wheat. *Nat. Commun.* **2019**, *10* (1). https://doi.org/10.1038/s41467-019-08726-9.

(155) Alanjary, M.; Kronmiller, B.; Adamek, M.; Blin, K.; Weber, T.; Huson, D.; Philmus, B.; Ziemert, N.; Kronmiller, B.; Blin, K.; et al. The Antibiotic Resistant Target Seeker (ARTS), an Exploration Engine for Antibiotic Cluster Prioritization and Novel Drug Target Discovery. *Nucleic Acids Res.* **2017**, *45* (W1), W42–W48. https://doi.org/10.1093/nar/gkx360.

(156) Vandova, G. A.; Nivina, A.; Khosla, C.; Davis, R. W. Identification of Polyketide Biosynthetic Gene Clusters That Harbor Self- Resistance Target Genes. *bioRxiv* **2020**, 0–36.

(157) Brown, L. C. W.; Acker, M. G.; Clardy, J.; Walsh, C. T.; Fischbach, M. A. Thirteen Posttranslational Modifications Convert a 14-Residue Peptide into the Antibiotic Thiocillin. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (8), 2549–2553. https://doi.org/10.1073/pnas.0900008106.

(158) Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T. L. BLAST plus: Architecture and Applications. *BMC Bioinformatics* **2009**, *10* (421), 1. https://doi.org/Artn 421\nDoi 10.1186/1471-2105-10-421.

(159) Nielsen, J. C.; Grijseels, S.; Prigent, S.; Ji, B.; Dainat, J.; Nielsen, K. F.; Frisvad, J. C.; Workman, M.; Nielsen, J. Global Analysis of Biosynthetic Gene Clusters Reveals Vast Potential of Secondary Metabolite Production in Penicillium Species. *Nat. Microbiol.* **2017**, *2* (April), 17044. https://doi.org/10.1038/nmicrobiol.2017.44.

(160) Rausch, C.; Hoof, I.; Weber, T.; Wohlleben, W.; Huson, D. H. Phylogenetic Analysis of Condensation Domains in NRPS Sheds Light on Their Functional Evolution. *BMC Evol. Biol.* **2007**, *7* (1), 78. https://doi.org/10.1186/1471-2148-7-78.

(161) Khaldi, N.; Collemare, J.; Lebrun, M. H.; Wolfe, K. H. Evidence for Horizontal Transfer of a Secondary Metabolite Gene Cluster between Fungi. *Genome Biol.* **2008**, *9* (1), 1–10. https://doi.org/10.1186/gb-2008-9-1-r18.

(162) Keeling, P. J.; Burki, F.; Wilcox, H. M.; Allam, B.; Allen, E. E.; Amaral-Zettler, L. A.; Armbrust, E. V.; Archibald, J. M.; Bharti, A. K.; Bell, C. J.; et al. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. *PLoS Biol.* **2014**, *12* (6), e1001889. https://doi.org/10.1371/journal.pbio.1001889.

(163) Ni, M.; Gao, N.; Kwon, N.-J.; Shin, K.-S.; Yu, J.-H. Regulation of Aspergillus Conidiation. In *Cellular and Molecular Biology of Filamentous Fungi*; 2010; pp 559–576.

(164) Ni, M.; Yu, J. H. A Novel Regulator Couples Sporogenesis and Trehalose Biogenesis in Aspergillus Nidulans. *PLoS One* **2007**, *2* (10), 1–9. https://doi.org/10.1371/journal.pone.0000970.

(165) Emms, D. M.; Kelly, S. OrthoFinder: Solving Fundamental Biases in Whole Genome Comparisons Dramatically Improves Orthogroup Inference Accuracy. *Genome Biol.* **2015**, *16* (1), 1–14. https://doi.org/10.1186/s13059-015-0721-2.

(166) Sullivan, M. J.; Petty, N. K.; Beatson, S. A. Easyfig: A Genome Comparison Visualizer. *Bioinformatics* **2011**, *27* (7), 1009–1010. https://doi.org/10.1093/bioinformatics/btr039.

(167) Mondo, S. J.; Dannebaum, R. O.; Kuo, R. C.; Louie, K. B.; Bewick, A. J.; LaButti, K.; Haridas, S.; Kuo, A.; Salamov, A.; Ahrendt, S. R.; et al. Widespread Adenine N6-Methylation of Active Genes in Fungi. *Nat. Genet.* **2017**, *49* (6), 964–968. https://doi.org/10.1038/ng.3859.

(168) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; et al. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34* (8), 828–837. https://doi.org/10.1038/nbt.3597.

(169) Lee, I.-K.; Han, M.-S.; Lee, M.-S.; Kim, Y.-S.; Yun, B.-S. Styrylpyrones from the Medicinal Fungus Phellinus Baumii and Their Antioxidant Properties. *Bioorg. Med.*

Chem. Lett. **2010**, *20* (18), 5459–5461. https://doi.org/10.1016/j.bmcl.2010.07.093.

(170) Lee, I.-K.; Yun, B.-S. Styrylpyrone-Class Compounds from Medicinal Fungi Phellinus and Inonotus Spp., and Their Medicinal Importance. *J. Antibiot. (Tokyo).* **2011**, *64* (5), 349–359. https://doi.org/10.1038/ja.2011.2.

(171) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. SIRIUS 4: A Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information. *Nat. Methods* **2019**, *16* (4), 299–302. https://doi.org/10.1038/s41592-019-0344-8.

(172) Dührkop, K.; Nothias, L. F.; Fleischauer, M.; Ludwig, M.; Hoffmann, M.; Rousu, J.; Dorrestein, P.; Böcker, S. Classes for the Masses: Systematic Classification of Unknowns Using Fragmentation Spectra. *bioRxiv* **2020**. https://doi.org/10.1101/2020.04.17.046672.

(173) Hobson, P. N. Rumen Micro-Organisms. *Prog. Ind. Microbiol.* **1971**, *9*, 41–77.

(174) Theodorou, M. K.; Brookman, J.; Trinci, A. P. J. Anaerobic Fungi. In *Methods in Gut Microbial Ecology for Ruminants*; Springer-Verlag: Berlin/Heidelberg, 2005; pp 55–66. https://doi.org/10.1007/1-4020-3791-0_5.

(175) Finn, R. D.; Coggill, P.; Eberhardt, R. Y.; Eddy, S. R.; Mistry, J.; Mitchell, A. L.; Potter, S. C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; et al. The Pfam Protein Families Database: Towards a More Sustainable Future. *Nucleic Acids Res.* **2016**, *44* (D1), D279–D285. https://doi.org/10.1093/nar/gkv1344.

(176) Katoh, K.; Misawa, K.; Kuma, K.; Miyata, T. MAFFT a Novel Method for Rapid Multiple Sequence Aligment Based on Fast Fourier Transform. *Nucleic Acids Res.* **2002**, *30* (14), 3059–3066. https://doi.org/10.1093/nar/gkf436.

(177) Capella-Gutiérrez, S.; Silla-Martínez, J. M.; Gabaldón, T. TrimAl: A Tool for Automated Alignment Trimming in Large-Scale Phylogenetic Analyses. *Bioinformatics* **2009**, *25* (15), 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

(178) Price, M. N.; Dehal, P. S.; Arkin, A. P. Fasttree: Computing Large Minimum Evolution Trees with Profiles Instead of a Distance Matrix. *Mol. Biol. Evol.* **2009**, *26* (7), 1641–1650. https://doi.org/10.1093/molbev/msp077.

(179) Stamatakis, A. RAxML Version 8: A Tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* **2014**, *30* (9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

(180) Finn, R. D.; Mistry, J.; Tate, J.; Coggill, P.; Heger, A.; Pollington, J. E.; Gavin, O. L.; Gunasekaran, P.; Ceric, G.; Forslund, K.; et al. The Pfam Protein Families Database. *Nucleic Acids Res.* **2010**, *38* (Database issue), D211-22. https://doi.org/10.1093/nar/gkp985.

(181) Mitchell, A. L.; Attwood, T. K.; Babbitt, P. C.; Blum, M.; Bork, P.; Bridge, A.; Brown, S. D.; Chang, H.-Y.; El-Gebali, S.; Fraser, M. I.; et al. InterPro in 2019: Improving Coverage, Classification and Access to Protein Sequence Annotations. *Nucleic Acids Res.* **2018**, 1–10. https://doi.org/10.1093/nar/gky1100.

(182) Marchler-Bauer, A.; Bryant, S. H. CD-Search: Protein Domain Annotations on the Fly. *Nucleic Acids Res.* **2004**, *32* (WEB SERVER ISS.), 327–331. https://doi.org/10.1093/nar/gkh454.

(183) Hoff, K. J.; Lange, S.; Lomsadze, A.; Borodovsky, M.; Stanke, M. BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and

AUGUSTUS. *Bioinformatics* **2015**, *32* (5), 767–769.

https://doi.org/10.1093/bioinformatics/btv661.

(184) Gotz, S.; Garcia-Gomez, J. M.; Terol, J.; Williams, T. D.; Nagaraj, S. H.; Nueda, M.

J.; Robles, M.; Talon, M.; Dopazo, J.; Conesa, A. High-Throughput Functional

Annotation and Data Mining with the Blast2GO Suite. *Nucleic Acids Res.* **2008**, *36*

(10), 3420–3435. https://doi.org/10.1093/nar/gkn176.

(185) Mondo, S. J.; Dannebaum, R. O.; Kuo, R. C.; Louie, K. B.; Bewick, A. J.; LaButti,

K.; Haridas, S.; Kuo, A.; Salamov, A.; Ahrendt, S. R.; et al. Widespread Adenine N6-

Methylation of Active Genes in Fungi. *Nat. Genet.* **2017**, *49* (6), 964–968.

https://doi.org/10.1038/ng.3859.

(186) ES, N.; CD, N.; AC, S.; KE, B.-J.; Y-M, K.; JE, K.; MM, M.; AK, S.; RK, C.; AA,

S.; et al. MPLEx : A Robust and Universal Protocol. *mSystems* **2016**, *1* (3), 1–14.

https://doi.org/10.1128/mSystems.00043-16.Editor.

(187) Kelly, R. T.; Page, J. S.; Luo, Q.; Moore, R. J.; Orton, D. J.; Tang, K.; Smith, R. D.

Chemically Etched Open Tubular and Monolithic Emitters for Nanoelectrospray

Ionization Mass Spectrometry. *Anal. Chem.* **2006**, *78* (22), 7796–7801.

https://doi.org/10.1021/ac061133r.

(188) Katajamaa, M.; Miettinen, J.; Orešič, M. MZmine: Toolbox for Processing and

Visualization of Mass Spectrometry Based Molecular Profile Data. *Bioinformatics*

2006, *22* (5), 634–636. https://doi.org/10.1093/bioinformatics/btk039.

(189) Bowen, B. P.; Northen, T. R. Dealing with the Unknown: Metabolomics and

Metabolite Atlases. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (9), 1471–1476.

https://doi.org/10.1016/j.jasms.2010.04.003.

(190) Wang, Y.; Kora, G.; Bowen, B. P.; Pan, C. MIDAS: A Database-Searching Algorithm for Metabolite Identification in Metabolomics. *Anal. Chem.* **2014**, *86* (19), 9496–9503. https://doi.org/10.1021/ac5014783.

(191) Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **1999**, *27* (1), 29–34. https://doi.org/10.1093/nar/27.1.29.

(192) Mohanta, T. K. Fungi Contain Genes Associated with Flavonoid Biosynthesis Pathway. *J. Funct. Foods* **2020**, *68* (March), 103910. https://doi.org/10.1016/j.jff.2020.103910.

(193) Frank, A. M.; Monroe, M. E.; Shah, A. R.; Carver, J. J.; Bandeira, N.; Moore, R. J.; Anderson, G. A.; Smith, R. D.; Pevzner, P. A. Spectral Archives: Extending Spectral Libraries to Analyze Both Identified and Unidentified Spectra. *Nat. Methods* **2011**, *8* (7), 587–594. https://doi.org/10.1038/nmeth.1609.

(194) Orpin, C. G. The Rumen Flagellate Piromonas Communis: Its Life-History and Invasion of Plant Material in the Rumen. *J. Gen. Microbiol.* **1977**, *99*, 107–117.

(195) Theodorou, M. K.; Mennim, G.; Davies, D. R.; Zhu, W.-Y.; Trinci, A. P. J.; Brookman, J. L. Anaerobic Fungi in the Digestive Tract of Mammalian Herbivores and Their Potential for Exploitation. *Proc. Nutr. Soc.* **1996**, *55* (03), 913–926. https://doi.org/10.1079/PNS19960088.

(196) Youssef, N. H.; Couger, M. B.; Struchtemeyer, C. G.; Liggenstoffer, A. S.; Prade, R. A.; Najar, F. Z.; Atiyeh, H. K.; Wilkins, M. R.; Elshahed, M. S. The Genome of the Anaerobic Fungus Orpinomyces Sp. Strain C1A Reveals the Unique Evolutionary History of a Remarkable Plant Biomass Degrader. *Appl. Environ. Microbiol.* **2013**, *79*

(15), 4620–4634. https://doi.org/10.1128/AEM.00821-13.

(197) Akin, D. E.; Rigsby, L. L. Mixed Fungal Populations and Lignocellulosic Tissue Degradation in the Bovine Rumen. *Appl. Environ. Microbiol.* **1987**, *53* (9), 1987–1995. https://doi.org/10.1128/AEM.70.7.4402.

(198) Akin, D. E.; Borneman, W. S.; Lyon, C. E. Degradation of Leaf Blades and Stems by Monocentric and Polycentric Isolates of Ruminal Fungi. *Anim. Feed Sci. Technol.* **1990**, *31* (3–4), 205–221. https://doi.org/10.1016/0377-8401(90)90125-R.

(199) Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P. M.; Henrissat, B. The Carbohydrate-Active Enzymes Database (CAZy) in 2013. *Nucleic Acids Res.* **2014**, *42* (D1), 490–495. https://doi.org/10.1093/nar/gkt1178.

(200) Bauchop, T.; Mountfort, D. O. Cellulose Fermentation by a Rumen Anaerobic Fungus in Both the Absence and the Presence of Rumen Methanogens. *Appl. Envir. Microbiol.* **1981**, *42* (6), 1103–1110.

(201) Mountfort, D. O.; Asher, R. A.; Bauchop, T.; Mountfort, D. O. Fermentation of Cellulose to Methane and Carbon Dioxide by a Rumen Anaerobic Fungus in a Triculture with Fermentation of Cellulose to Methane and Carbon Dioxide by a Rumen Anaerobic Fungus in a Triculture with Methanobrevibacter Sp . Strain RAl and Methan. *Appl. Environ. Microbiol.* **1982**, *44* (1), 128–134.

(202) Joblin, K. N.; Naylor, G. E.; Williams, A. G. Effect of Methanobrevibacter Smithii on Xylanolytic Activity of Anaerobic Ruminal Fungi. *Appl. Environ. Microbiol.* **1990**, *56* (8), 2287–2295.

(203) Joblin, K. N.; Williams, A. G. Effect of Cocultivation of Ruminal Chytrid Fungi with Methanobrevibacter Smithii on Lucerne Stem Degradation and Extracellular Fungal

Enzyme Activities. *Lett. Appl. Microbiol.* **1991**, *12* (4), 121–124.

https://doi.org/10.1111/j.1472-765X.1991.tb00520.x.

(204) Li, Y.; Jin, W.; Mu, C.; Cheng, Y.; Zhu, W. Indigenously Associated Methanogens

Intensified the Metabolism in Hydrogenosomes of Anaerobic Fungi with Xylose as

Substrate. *J. Basic Microbiol.* **2017**, *57* (11), 933–940.

https://doi.org/10.1002/jobm.201700132.

(205) Holmes, D. E.; Smith, J. A. Biologically Produced Methane as a Renewable Energy

Source. *Adv. Appl. Microbiol.* **2016**, *97*, 1–61.

https://doi.org/10.1016/bs.aambs.2016.09.001.

(206) Nagy, T.; Tunnicliffe, R. B.; Higgins, L. D.; Walters, C.; Gilbert, H. J.; Williamson,

M. P. Characterization of a Double Dockerin from the Cellulosome of the Anaerobic

Fungus Piromyces Equi. *J. Mol. Biol.* **2007**, *373* (3), 612–622.

https://doi.org/10.1016/j.jmb.2007.08.007.

(207) Haitjema, C. H.; Solomon, K. V; Henske, J. K.; Theodorou, M. K.; O'Malley, M. A.

Anaerobic Gut Fungi: Advances in Isolation, Culture, and Cellulolytic Enzyme

Discovery for Biofuel Production. *Biotechnol. Bioeng.* **2014**, *111* (8), 1471–1482.

https://doi.org/10.1002/bit.25264.

(208) Wei, Y. Q.; Long, R. J.; Yang, H.; Yang, H. J.; Shen, X. H.; Shi, R. F.; Wang, Z. Y.;

Du, J. G.; Qi, X. J.; Ye, Q. H. Fiber Degradation Potential of Natural Co-Cultures of

Neocallimastix Frontalis and Methanobrevibacter Ruminantium Isolated from Yaks

(Bos Grunniens) Grazing on the Qinghai Tibetan Plateau. *Anaerobe* **2016**, *39*, 158–

164. https://doi.org/10.1016/j.anaerobe.2016.03.005.

(209) Jin, W.; Cheng, Y. F.; Mao, S. Y.; Zhu, W. Y. Isolation of Natural Cultures of

Anaerobic Fungi and Indigenously Associated Methanogens from Herbivores and Their Bioconversion of Lignocellulosic Materials to Methane. *Bioresour. Technol.* **2011**, *102* (17), 7925–7931. https://doi.org/10.1016/j.biortech.2011.06.026.

(210) Teunissen, M. J.; Op den Camp, H. J. M.; Orpin, C. G.; Huis in 't Veld, J. H. J.; Vogels, G. D. Comparison of Growth Characteristics of Anaerobic Fungi Isolated from Ruminant and Non-Ruminant Herbivores during Cultivation in a Defined Medium. *J. Gen. Microbiol.* **1991**, *137* (6), 1401–1408. https://doi.org/10.1099/00221287-137-6-1401.

(211) Henske, J. K.; Gilmore, S. P.; Haitjema, C. H.; Solomon, K. V; Malley, M. A. O. Biomass-Degrading Enzymes Are Catabolite Repressed in Anaerobic Gut Fungi. *AIChE J.* **2018**, *64* (12). https://doi.org/10.1002/aic.16395.

(212) Koonin, E.; Fedorova, N.; Jackson, J.; Jacobs, A.; Krylov, D.; Makarova, K.; Mazumder, R.; Mekhedov, S.; Nikolskaya, A.; Rao, B.; et al. A Comprehensive Evolutionary Classification of Proteins Encoded in Complete Eukaryotic Genomes. *Genome Biol.* **2004**, *5* (2), R7. https://doi.org/10.1186/gb-2004-5-2-r7.

(213) Perlin, M. H.; Andrews, J.; San Toh, S. Essential Letters in the Fungal Alphabet: ABC and MFS Transporters and Their Roles in Survival and Pathogenicity. *Adv. Genet.* **2014**, *85*, 201–253. https://doi.org/10.1016/B978-0-12-800271-1.00004-4.

(214) ter Beek, J.; Guskov, A.; Slotboom, D. J. Structural Diversity of ABC Transporters. *J. Gen. Physiol.* **2014**, *143* (4), 419–435. https://doi.org/10.1085/jgp.201411164.

(215) Zhang, W.; Kou, Y.; Xu, J.; Cao, Y.; Zhao, G.; Shao, J.; Wang, H.; Wang, Z.; Bao, X.; Chen, G.; et al. Two Major Facilitator Superfamily Sugar Transporters from Trichoderma Reesei and Their Roles in Induction of Cellulase Biosynthesis. *J. Biol.*

Chem. **2013**, *288* (46), 32861–32872. https://doi.org/10.1074/jbc.M113.505826.

(216) Seppala, S.; Solomon, K. V.; Gilmore, S. P.; Henske, J. K.; O'Malley, M. A. Mapping

the Membrane Proteome of Anaerobic Gut Fungi Identifies a Wealth of Carbohydrate

Binding Proteins and Transporters. *Microb. Cell Fact.* **2016**, *Accepted f*, 1–14.

https://doi.org/10.1186/s12934-016-0611-7.

(217) Theodorou, M. K.; Davies, D. R.; Nielsen, B. B.; Lawrence, M. I. G.; Trinci, A. P. J.

Determination of Growth of Anaerobic Fungi on Soluble and Cellulosic Substrates

Using a Pressure Transducer. *Microbiology* **1995**, *141* (3), 671–678.

https://doi.org/10.1099/13500872-141-3-671.

(218) Schink, B. Energetics of Syntrophic Cooperation in Methanogenic Degradation.

Microbiol. Mol. Biol. Rev. **1997**, *61* (2), 262–280. https://doi.org/1092-

2172/97/$04.0010/$04.00+0.

(219) Balch, W. E.; Fox, G. E.; Magrum, L. J.; Woese, C. R.; Wolfe, A. R. S. Methanogens:

Reevaluation of a Unique Biological Group. *Microbiol. Rev.* **1979**, *43* (2), 260–296.

(220) Kim, D.; Langmead, B.; Salzberg, S. L. HISAT: A Fast Spliced Aligner with Low

Memory Requirements. *Nat. Methods* **2015**, *12* (4), 357–360.

https://doi.org/10.1038/nmeth.3317.

(221) Liao, Y.; Smyth, G. K.; Shi, W. FeatureCounts: An Efficient General Purpose

Program for Assigning Sequence Reads to Genomic Features. *Bioinformatics* **2014**,

30 (7), 923–930. https://doi.org/10.1093/bioinformatics/btt656.

(222) Love, M. I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and

Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* **2014**, *15* (12), 550.

https://doi.org/10.1186/s13059-014-0550-8.

(223) Mi, H.; Muruganujan, A.; Casagrande, J. T.; Thomas, P. D. Large-Scale Gene Function Analysis with the PANTHER Classification System. *Nat. Protoc.* **2013**, *8* (8), 1551–1566. https://doi.org/10.1038/nprot.2013.092.

(224) Bernard, T.; Cantarel, B. L.; Henrissat, B.; Lombard, V.; Coutinho, P. M.; Rancurel, C. The Carbohydrate-Active EnZymes Database (CAZy): An Expert Resource for Glycogenomics. *Nucleic Acids Res.* **2008**, *37* (Database), D233–D238. https://doi.org/10.1093/nar/gkn663.

(225) Krogh, A.; Larsson, B.; von Heijne, G.; Sonnhammer, E. L. . Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes. *J. Mol. Biol.* **2001**, *305* (3), 567–580. https://doi.org/10.1006/jmbi.2000.4315.

(226) Stephen F. Altschul, Warren Gish, Webb Miller, Eugene W. Myers, D. J. L. Basic Local Alignment Search Tool. *J. Mol. Biol.* **1990**, *215*, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

(227) Di Francesco, A.; Martini, C.; Mari, M. Biological Control of Postharvest Diseases by Microbial Antagonists: How Many Mechanisms of Action? *Eur. J. Plant Pathol.* **2016**, *145* (4), 711–717. https://doi.org/10.1007/s10658-016-0867-0.

(228) Dukare, A. S.; Paul, S.; Nambi, V. E.; Gupta, R. K.; Singh, R.; Sharma, K.; Vishwakarma, R. K. Exploitation of Microbial Antagonists for the Control of Postharvest Diseases of Fruits: A Review. *Crit. Rev. Food Sci. Nutr.* **2019**, *59* (9), 1498–1513. https://doi.org/10.1080/10408398.2017.1417235.

(229) Carmona-Hernandez, S.; Reyes-Pérez, J. J.; Chiquito-Contreras, R. G.; Rincon-Enriquez, G.; Cerdan-Cabrera, C. R.; Hernandez-Montiel, L. G. Biocontrol of

Postharvest Fruit Fungal Diseases by Bacterial Antagonists: A Review. *Agronomy* **2019**, *9* (3). https://doi.org/10.3390/agronomy9030121.

(230) García-Bayona, L.; Comstock, L. E. Bacterial Antagonism in Host-Associated Microbial Communities. *Science (80-.).* **2018**, *361* (6408). https://doi.org/10.1126/science.aat2456.

(231) Kerr, B.; Riley, M. A.; Feldman, M. W.; Bohannan, B. J. M. Local Dispersal Promotes Biodiversity in a Real-Life Game of Rock-Paper-Scissors. *Nature* **2002**, *418* (6894), 171–174. https://doi.org/10.1038/nature00823.

(232) Kirkup, B. C.; Riley, M. A. Antibiotic-Mediated Antagonism Leads to a Bacterial Game of Rock-Paper-Scissors in Vivo. *Nature* **2004**, *428* (6981), 412–414. https://doi.org/10.1038/nature02429.

(233) Sassone-Corsi, M.; Nuccio, S. P.; Liu, H.; Hernandez, D.; Vu, C. T.; Takahashi, A. A.; Edwards, R. A.; Raffatellu, M. Microcins Mediate Competition among Enterobacteriaceae in the Inflamed Gut. *Nature* **2016**, *540* (7632), 280–283. https://doi.org/10.1038/nature20557.

(234) Borgeaud, S.; Metzger, L. C.; Scrignari, T.; Blokesch, M. The Type VI Secretion System of Vibrio Cholerae Fosters Horizontal Gene Transfer. *Science (80-.).* **2015**, *347* (6217), 63–67. https://doi.org/10.1126/science.1260064.

(235) Coyte, K. Z.; Schluter, J.; Foster, K. R. The Ecology of the Microbiome: Networks, Competition, and Stability. *Science (80-.).* **2015**, *350* (6261), 663–666. https://doi.org/10.1126/science.aad2602.

(236) Orpin, C. G.; Ho, Y. W. Ecology and Function of the Anaerobic Rumen Fungi. Serdang, Malaysia, Malaysian Society of Animal Production 1991.

(237) Jung, B.; Park, J.; Kim, N.; Li, T.; Kim, S.; Bartley, L. E.; Kim, J.; Kim, I.; Kang, Y.; Yun, K.; et al. Cooperative Interactions between Seed-Borne Bacterial and Air-Borne Fungal Pathogens on Rice. *Nat. Commun.* **2018**, *9* (1). https://doi.org/10.1038/s41467-017-02430-2.

(238) Swift, C. L.; Brown, J. L.; Seppälä, S.; O'Malley, M. A. Co-Cultivation of the Anaerobic Fungus Anaeromyces Robustus with Methanobacterium Bryantii Enhances Transcription of Carbohydrate Active Enzymes. *J. Ind. Microbiol. Biotechnol.* **2019**, *46*, 1427–1433. https://doi.org/10.1007/s10295-019-02188-0.

(239) Khan, N.; Maezato, Y.; McClure, R. S.; Brislawn, C. J.; Mobberley, J. M.; Isern, N.; Chrisler, W. B.; Markillie, L. M.; Barney, B. M.; Song, H. S.; et al. Phenotypic Responses to Interspecies Competition and Commensalism in a Naturally-Derived Microbial Co-Culture. *Sci. Rep.* **2018**, *8* (1), 1–9. https://doi.org/10.1038/s41598-017-18630-1.

(240) Daly, P.; van Munster, J. M.; Kokolski, M.; Sang, F.; Blythe, M. J.; Malla, S.; Velasco de Castro Oliveira, J.; Goldman, G. H.; Archer, D. B. Transcriptomic Responses of Mixed Cultures of Ascomycete Fungi to Lignocellulose Using Dual RNA-Seq Reveal Inter-Species Antagonism and Limited Beneficial Effects on CAZyme Expression. *Fungal Genet. Biol.* **2017**, *102*, 4–21. https://doi.org/10.1016/j.fgb.2016.04.005.

(241) Neumann, A. P.; Suen, G. The Phylogenomic Diversity of Herbivore-Associated Fibrobacter Spp. Is Correlated to Lignocellulose-Degrading Potential. *mSphere* **2018**, *3* (6), 1–18. https://doi.org/10.1128/mSphere.00593-18.

(242) Stachelhaus, T.; Mootz, H. D.; Bergendahl, V.; Marahiel, M. A. Peptide Bond

Formation in Nonribosomal Peptide Biosynthesis. *J. Biol. Chem.* **1998**, *273* (35), 22773–22781. https://doi.org/10.1074/jbc.273.35.22773.

(243) Karimi-Aghcheh, R.; Bok, J. W.; Phatale, P. A.; Smith, K. M.; Baker, S. E.; Lichius, A.; Omann, M.; Zeilinger, S.; Seiboth, B.; Rhee, C.; et al. Functional Analyses of Trichoderma Reesei LAE1 Reveal Conserved and Contrasting Roles of This Regulator . *G3 &#58; Genes|Genomes|Genetics* **2013**, *3* (2), 369–378. https://doi.org/10.1534/g3.112.005140.

(244) Shwab, E. K.; Jin, W. B.; Tribus, M.; Galehr, J.; Graessle, S.; Keller, N. P. Historie Deacetylase Activity Regulates Chemical Diversity in Aspergillus. *Eukaryot. Cell* **2007**, *6* (9), 1656–1664. https://doi.org/10.1128/EC.00186-07.

(245) Bok, J. W.; Chiang, Y. M.; Szewczyk, E.; Reyes-Domingez, Y.; Davidson, A. D.; Sanchez, J. F.; Lo, H. C.; Watanabe, K.; Strauss, J.; Oakley, B. R.; et al. Chromatin-Level Regulation of Biosynthetic Gene Clusters. *Nat. Chem. Biol.* **2009**, *5* (7), 462–464. https://doi.org/10.1038/nchembio.177.

(246) Greer, E. L.; Shi, Y. Histone Methylation: A Dynamic Mark in Health, Disease and Inheritance. *Nat. Rev. Genet.* **2012**, *13* (5), 343–357. https://doi.org/10.1038/nrg3173.

(247) Bozhenok, L.; Wade, P. A.; Varga-Weisz, P. WSTF-ISWI Chromatin Remodeling Complex Targets Heterochromatic Replication Foci. *EMBO J.* **2002**, *21* (9), 2231–2241. https://doi.org/10.1093/emboj/21.9.2231.

(248) Chen, I. M. A.; Chu, K.; Palaniappan, K.; Pillay, M.; Ratner, A.; Huang, J.; Huntemann, M.; Varghese, N.; White, J. R.; Seshadri, R.; et al. IMG/M v.5.0: An Integrated Data Management and Comparative Analysis System for Microbial Genomes and Microbiomes. *Nucleic Acids Res.* **2019**, *47* (D1), D666–D677.

https://doi.org/10.1093/nar/gky901.

(249) Sharff, A.; Fanutti, C.; Shi, J.; Calladine, C.; Luisi, B. The Role of the TolC Family in Protein Transport and Multidrug Efflux from Stereochemical Certainty to Mechanistic Hypothesis. *Eur. J. Biochem.* **2001**, *268* (19), 5011–5026. https://doi.org/10.1046/j.0014-2956.2001.02442.x.

(250) Weston, N.; Sharma, P.; Ricci, V.; Piddock, L. J. V. Regulation of the AcrAB-TolC Efflux Pump in Enterobacteriaceae. *Res. Microbiol.* **2018**, *169* (7–8), 425–431. https://doi.org/10.1016/j.resmic.2017.10.005.

(251) Nikaido, H.; Takatsuka, Y. Mechanisms of RND Multidrug Efflux Pumps. *Biochim. Biophys. Acta - Proteins Proteomics* **2009**, *1794* (5), 769–781. https://doi.org/10.1016/j.bbapap.2008.10.004.

(252) Makarova, K. S.; Grishin, N. V.; Koonin, E. V. The HicAB Cassette, a Putative Novel, RNA-Targeting Toxin-Antitoxin System in Archaea and Bacteria. *Bioinformatics* **2006**, *22* (21), 2581–2584. https://doi.org/10.1093/bioinformatics/btl418.

(253) Fineran, P. C.; Blower, T. R.; Foulds, I. J.; Humphreys, D. P.; Lilley, K. S.; Salmond, G. P. C. The Phage Abortive Infection System, ToxIN, Functions as a Protein-RNA Toxin-Antitoxin Pair. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106* (3), 894–899. https://doi.org/10.1073/pnas.0808832106.

(254) Buchanan, S. K.; Smith, B. S.; Venkatramani, L.; Xia, D.; Esser, L.; Palnitkar, M.; Chakraborty, R.; Van Der Helm, D.; Deisenhofer, J. Crystal Structure of the Outer Membrane Active Transporter FepA from Escherichia Coli. *Nat. Struct. Biol.* **1999**, *6* (1), 56–63. https://doi.org/10.1038/4931.

(255) Shultis, D. D.; Purdy, M. D.; Banchs, C. N.; Wiener, M. C. Outer Membrane Active Transport: Structure of the BtuB:TonB Complex. *Science (80-.).* **2006**, *312* (5778), 1396–1399. https://doi.org/10.1126/science.1127694.

(256) Braun, V.; Hantke, K. Recent Insights into Iron Import by Bacteria. *Curr. Opin. Chem. Biol.* **2011**, *15* (2), 328–334. https://doi.org/10.1016/j.cbpa.2011.01.005.

(257) Xia, J.; Wishart, D. S. Web-Based Inference of Biological Patterns , Functions and Pathways from Metabolomic Data Using MetaboAnalyst. *Nat. Protoc.* **2011**, *6* (6), 743–760. https://doi.org/10.1038/nprot.2011.319.

(258) Xia, J.; Sinelnikov, I. V; Han, B.; Wishart, D. S. MetaboAnalyst 3 . 0 — Making Metabolomics More Meaningful. *Nucleic Acids Res.* **2015**, *43* (April), 251–257. https://doi.org/10.1093/nar/gkv380.

(259) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13* (11), 2498–2504. https://doi.org/10.1101/gr.1239303.

(260) Baran, R.; Robert, M.; Suematsu, M.; Soga, T.; Tomita, M. Visualization of Three-Way Comparisons of Omics Data. *BMC Bioinformatics* **2007**, *8*, 1–8. https://doi.org/10.1186/1471-2105-8-72.

(261) Zheng, L.; Lin, Y.; Lu, S.; Zhang, J.; Bogdanov, M. Biogenesis, Transport and Remodeling of Lysophospholipids in Gram-Negative Bacteria. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **2017**, *1862* (11), 1404–1413. https://doi.org/10.1016/j.bbalip.2016.11.015.

(262) Davies, D. R.; Theodorou, M. K.; Lawrence, M. I.; Trinci, a P. Distribution of

Anaerobic Fungi in the Digestive Tract of Cattle and Their Survival in Faeces. *J. Gen. Microbiol.* **1993**, *139 Pt 6*, 1395–1400. https://doi.org/10.1099/00221287-139-6-1395.

(263) Ghali, I.; Sofyan, A.; Ohmori, H.; Shinkai, T.; Mitsumori, M. Diauxic Growth of Fibrobacter Succinogenes S85 on Cellobiose and Lactose. *FEMS Microbiol. Lett.* **2017**, *364* (15), 1–9. https://doi.org/10.1093/femsle/fnx150.

(264) BBMap short read aligner, and other bioinformatic tools https://sourceforge.net/projects/bbmap/ (accessed Feb 17, 2020).

(265) Nothias, L. F.; Petras, D.; Schmid, R.; Dührkop, K.; Rainer, J.; Sarvepalli, A.; Protsyuk, I.; Ernst, M.; Tsugawa, H. Feature-Based Molecular Networking in the GNPS Analysis Environment. *biorXiv* **2019**.

(266) Myers, O. D.; Sumner, S. J.; Li, S.; Barnes, S.; Du, X. One Step Forward for Reducing False Positive and False Negative Compound Identifications from Mass Spectrometry Metabolomics Data: New Algorithms for Constructing Extracted Ion Chromatograms and Detecting Chromatographic Peaks. *Anal. Chem.* **2017**, *89* (17), 8696–8703. https://doi.org/10.1021/acs.analchem.7b00947.

(267) Mohimani, H.; Gurevich, A.; Shlemov, A.; Mikheenko, A.; Korobeynikov, A.; Cao, L.; Shcherbin, E.; Nothias, L. F.; Dorrestein, P. C.; Pevzner, P. A. Dereplication of Microbial Metabolites through Database Search of Mass Spectra. *Nat. Commun.* **2018**, *9* (1), 1–12. https://doi.org/10.1038/s41467-018-06082-8.

(268) Chong, J.; Wishart, D. S.; Xia, J. Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. *Curr. Protoc. Bioinforma.* **2019**, *68* (1). https://doi.org/10.1002/cpbi.86.

(269) Katajamaa, M.; Orešič, M. Processing Methods for Differential Analysis of LC/MS Profile Data. *BMC Bioinformatics* **2005**, *6*. https://doi.org/10.1186/1471-2105-6-179.

(270) R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing https://www.r-project.org/ (accessed Feb 20, 2020).

(271) Glover, J. R.; Lindquist, S. Hsp104, Hsp70, and Hsp40: A Novel Chaperone System That Rescues Previously Aggregated Proteins. *Cell* **1998**, *94* (1), 73–82. https://doi.org/10.1016/S0092-8674(00)81223-4.

(272) Boreham, D. R.; Mitchel, R. E. J. Regulation of Heat and Radiation Stress Responses in Yeast by Hsp-104. *Radiat. Res.* **1994**, *137* (2), 190. https://doi.org/10.2307/3578811.

(273) Borchsenius, A. S.; Wegrzyn, R. D.; Newnam, G. P.; Inge-Vechtomov, S. G.; Chernoff, Y. O. Yeast Prion Protein Derivative Defective in Aggregate Shearing and Production of New "Seeds." *EMBO J.* **2001**, *20* (23), 6683–6691. https://doi.org/10.1093/emboj/20.23.6683.

(274) Fernandes, P. M. B.; Domitrovic, T.; Kao, C. M.; Kurtenbach, E. Genomic Expression Pattern in Saccharomyces Cerevisiae Cells in Response to High Hydrostatic Pressure. *FEBS Lett.* **2004**, *556* (1–3), 153–160. https://doi.org/10.1016/S0014-5793(03)01396-6.

(275) Estruch, F.; Valencia, U. De. Hsf1p and Msn2/4p Cooperate in the Expression of Saccharomyces Cerevisiae Genes HSP26 and HSP104 in a Gene- and Stress Type-Dependent Manner. *Mol. Microbiol.* **2001**, *39* (6), 1523–1532.

(276) Burnie, J.; Matthews, R. Therapeutic Composition, April 11, 2006.

(277) Gasch, A. P. Comparative Genomics of the Environmental Stress Response in Ascomycete Fungi. *Yeast* **2007**, *24*, 961–976. https://doi.org/10.1002/yea.

(278) Gasch, A. P.; Werner-Washburne, M. The Genomics of Yeast Responses to Environmental Stress and Starvation. *Funct. Integr. Genomics* **2002**, *2* (4–5), 181–192. https://doi.org/10.1007/s10142-002-0058-2.

(279) Gasch, A. P.; Spellman, P. T.; Kao, C. M.; Carmel-Harel, O.; Eisen, M. B.; Storz, G.; Botstein, D.; Brown, P. O. Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes. *Mol. Biol. Cell* **2000**, *11* (12), 4241–4257. https://doi.org/10.1091/MBC.11.12.4241.

(280) Causton, H. C.; Ren, B.; Sang Seok Koh; Harbison, C. T.; Kanin, E.; Jennings, E. G.; Tong Ihn Lee; True, H. L.; Lander, E. S.; Young, R. A. Remodeling of Yeast Genome Expression in Response to Environmental Changes. *Mol. Biol. Cell* **2001**, *12* (2), 323–337. https://doi.org/10.1091/mbc.12.2.323.

(281) O'Malley, M. A.; Mancini, J. D.; Young, C. L.; McCusker, E. C.; Raden, D.; Robinson, A. S. Progress toward Heterologous Expression of Active G-Protein-Coupled Receptors in Saccharomyces Cerevisiae: Linking Cellular Stress Response with Translocation and Trafficking. *Protein Sci.* **2009**, *18* (11), 2356–2370. https://doi.org/10.1002/pro.246.

(282) Guillemette, T.; van Peij, N. N.; Goosen, T.; Lanthaler, K.; Robson, G. D.; van den Hondel, C. A.; Stam, H.; Archer, D. B. Genomic Analysis of the Secretion Stress Response in the Enzyme-Producing Cell Factory Aspergillus Niger. *BMC Genomics* **2007**, *8* (1), 158. https://doi.org/10.1186/1471-2164-8-158.

(283) Mattanovich, D.; Gasser, B.; Hohenblum, H.; Sauer, M. Stress in Recombinant

Protein Producing Yeasts. *J. Biotechnol.* **2004**, *113* (1–3), 121–135. https://doi.org/10.1016/j.jbiotec.2004.04.035.

(284) Liu, Y.; Chang, A. Heat Shock Response Relieves ER Stress. *EMBO J.* **2008**, *27* (7), 1049–1059. https://doi.org/10.1038/emboj.2008.42.

(285) Hollien, J. Evolution of the Unfolded Protein Response. *Biochim. Biophys. Acta* **2013**, *1833* (11), 2458–2463. https://doi.org/10.1016/j.bbamcr.2013.01.016.

(286) Hernández-Elvira, M.; Torres-Quiroz, F.; Escamilla-Ayala, A.; Domínguez-Martin, E.; Escalante, R.; Kawasaki, L.; Ongay-Larios, L.; Coria, R. The Unfolded Protein Response Pathway in the Yeast Kluyveromyces Lactis. A Comparative View among Yeast Species. *Cells* **2018**, *7* (8), 106. https://doi.org/10.3390/cells7080106.

(287) Travers, K. J.; Patil, C. K.; Wodicka, L.; Lockhart, D. J.; Weissman, J. S.; Walter, P. Functional and Genomic Analyses Reveal an Essential Coordination between the Unfolded Protein Response and ER-Associated Degradation. *Cell* **2000**, *101* (3), 249–258. https://doi.org/10.1016/S0092-8674(00)80835-1.

(288) Jung, K. W.; So, Y. S.; Bahn, Y. S. Unique Roles of the Unfolded Protein Response Pathway in Fungal Development and Differentiation. *Sci. Rep.* **2016**, *6* (September), 1–14. https://doi.org/10.1038/srep33413.

(289) Cox, J. S.; Walter, P. A Novel Mechanism for Regulating Activity of a Transcription Factor That Controls the Unfolded Protein Response. *Cell* **1996**, *87*, 391–404.

(290) Mori, K.; Kawahara, T.; Yoshida, H.; Yanagi, H.; Yura, T. Signalling from Endoplasmic Reticulum to Nucleus: Transcription Factor with a Basic-Leucine Zipper Motif Is Required for the Unfolded Protein-Response Pathway. *Genes to Cells* **1996**, *1* (9), 803–817. https://doi.org/10.1046/j.1365-2443.1996.d01-274.x.

(291) Cox, J. S.; Shamu, C. E.; Walter, P. Transcriptional Induction of Genes Encoding Endoplasmic Reticulum Resident Proteins Requires a Transmembrane Protein Kinase. *Cell* **1993**, *73* (6), 1197–1206. https://doi.org/10.1016/0092-8674(93)90648-A.

(292) Morl, K.; Ma, W.; Gething, M. J.; Sambrook, J. A Transmembrane Protein with a Cdc2+ CDC28-Related Kinase Activity Is Required for Signaling from the ER to the Nucleus. *Cell* **1993**, *74* (4), 743–756. https://doi.org/10.1016/0092-8674(93)90521-Q.

(293) Gardner, B. M.; Pincus, D.; Gotthardt, K.; Gallagher, C. M.; Walter, P. Endoplasmic Reticulum Stress Sensing in the Unfolded Protein Response. *Cold Spring Harb Perspect Biol* **2013**, *5* (3), a013169. https://doi.org/10.1101/cshperspect.a013169.

(294) Dever, T. E. Gene-Specific Regulation by General Translation Factors. **2002**, *108*, 545–556.

(295) Wek, R. C.; Cavener, D. R. Translational Control and the Unfolded Protein Response. *Antioxid. Redox Signal.* **2007**, *9* (12), 2357–2372. https://doi.org/10.1089/ars.2007.1764.

(296) Wek, R. C.; Jiang, H. Y.; Anthony, T. G. Coping with Stress: EIF2 Kinases and Translational Control. *Biochem. Soc. Trans.* **2006**, *34* (1), 7–11. https://doi.org/10.1042/BST0340007.

(297) Hernández, G.; Jagus, R. *Evolution of the Protein Synthesis Machinery and Its Regulation*; 2016.

(298) Heimel, K. Unfolded Protein Response in Filamentous Fungi—Implications in Biotechnology. *Appl. Microbiol. Biotechnol.* **2014**, *99* (1), 121–132. https://doi.org/10.1007/s00253-014-6192-7.

(299) Gasser, B.; Maurer, M.; Rautio, J.; Sauer, M.; Bhattacharyya, A.; Saloheimo, M.; Penttilä, M.; Mattanovich, D. Monitoring of Transcriptional Regulation in Pichia Pastoris under Protein Production Conditions. *BMC Genomics* **2007**, *8*, 1–18. https://doi.org/10.1186/1471-2164-8-179.

(300) Valkonen, M.; Penttilä, M.; Saloheimo, M. Effects of Inactivation and Constitutive Expression of the Unfolded-Protein Response Pathway on Protein Production in the Yeast Saccharomyces Cerevisiae. *Appl. Environ. Microbiol.* **2003**, *69* (4), 2065–2072. https://doi.org/10.1128/AEM.69.4.2065-2072.2003.

(301) Valkonen, M.; Ward, M.; Wang, H.; Penttilä, M.; Saloheimo, M. Improvement of Foreign-Protein Production in Aspergillus Niger Var. Awamori by Constitutive Induction of the Unfolded-Protein Response. *Appl. Environ. Microbiol.* **2003**, *69* (12), 6979–6986. https://doi.org/10.1128/AEM.69.12.6979-6986.2003.

(302) Seppälä, S.; Wilken, S. E.; Knop, D.; Solomon, K. V.; O'Malley, M. A. The Importance of Sourcing Enzymes from Non-Conventional Fungi for Metabolic Engineering and Biomass Breakdown. *Metab. Eng.* **2017**, *44* (June), 45–59. https://doi.org/10.1016/j.ymben.2017.09.008.

(303) Henske, J. K.; Wilken, S. E.; Solomon, K. V.; Smallwood, C. R.; Shutthanandan, V.; Evans, J. E.; Theodorou, M. K.; O'Malley, M. A. Metabolic Characterization of Anaerobic Fungi Provides a Path Forward for Bioprocessing of Crude Lignocellulose. *Biotechnol. Bioeng.* **2018**, *115* (4), 874–884. https://doi.org/10.1002/bit.26515.

(304) Solomon, K. V.; Henske, J. K.; Theodorou, M. K.; O'Malley, M. A. Robust and Effective Methodologies for Cryopreservation and DNA Extraction from Anaerobic Gut Fungi. *Anaerobe* **2016**, *38*, 39–46.

https://doi.org/10.1016/j.anaerobe.2015.11.008.

(305) Spatafora, J. W.; Chang, Y.; Benny, G. L.; Lazarus, K.; Smith, M. E.; Berbee, M. L.; Bonito, G.; Corradi, N.; Grigoriev, I.; Gryganskyi, A.; et al. A Phylum-Level Phylogenetic Classification of Zygomycete Fungi Based on Genome-Scale Data. *Mycologia* **2016**, *108* (5), 1028–1046. https://doi.org/10.3852/16-042.

(306) Mager, W. H.; Ferreira, P. M. Stress Response of Yeast. *Biochem. J.* **1993**, *290* (Pt 1), 1–13.

(307) Parlati, F.; Dominguez, M.; Bergeron, J. J. M.; Thomas, D. Y. Saccharomyces Cerevisiae CNE1 Encodes an Endoplasmic Reticulum (ER) Membrane Protein with Sequence Similarity to Calnexin and Calreticulin and Functions as a Constituent of the ER Quality Control Apparatus. *Journal of Biological Chemistry*. 1995, pp 244–253. https://doi.org/10.1074/jbc.270.1.244.

(308) Wang, P.; Li, J.; Tao, J.; Sha, B. The Luminal Domain of the ER Stress Sensor Protein PERK Binds Misfolded Proteins and Thereby Triggers PERK Oligomerization. *J. Biol. Chem.* **2018**, *293* (11), 4110–4121. https://doi.org/10.1074/jbc.RA117.001294.

(309) Narasimhan, J.; Joyce, B. R.; Naguleswaran, A.; Smith, A. T.; Livingston, M. R.; Dixon, S. E.; Coppens, I.; Wek, R. C.; Sullivan, W. J. Translation Regulation by Eukaryotic Initiation Factor-2 Kinases in the Development of Latent Cysts in Toxoplasma Gondii. *J. Biol. Chem.* **2008**, *283* (24), 16591–16601. https://doi.org/10.1074/jbc.M800681200.

(310) Amezcua, C. A.; Harper, S. M.; Rutter, J.; Gardner, K. H. Structure and Interactions of PAS Kinase N-Terminal PAS Domain: Model for Intramolecular Kinase

Regulation. *Structure* **2002**, *10* (10), 1349–1361. https://doi.org/10.1016/S0969-2126(02)00857-2.

(311) Taylor, B. L.; Zhulin, I. B. PAS Domains: Internal Sensors of Oxygen, Redox Potential, and Light. *Microbiol. Mol. Biol. Rev.* **1999**, *63* (2), 479–506. https://doi.org/10.1128/mmbr.63.2.479-506.1999.

(312) Almagro Armenteros, J. J.; Tsirigos, K. D.; Sønderby, C. K.; Petersen, T. N.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 5.0 Improves Signal Peptide Predictions Using Deep Neural Networks. *Nat. Biotechnol.* **2019**, *37* (4), 420–423. https://doi.org/10.1038/s41587-019-0036-z.

(313) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7* (1), 539. https://doi.org/10.1038/msb.2011.75.

(314) Waterhouse, A. M.; Procter, J. B.; Martin, D. M. A.; Clamp, M.; Barton, G. J. Jalview Version 2-A Multiple Sequence Alignment Editor and Analysis Workbench. *Bioinformatics* **2009**, *25* (9), 1189–1191. https://doi.org/10.1093/bioinformatics/btp033.

(315) Marchler-Bauer, A.; Derbyshire, M. K.; Gonzales, N. R.; Lu, S.; Chitsaz, F.; Geer, L. Y.; Geer, R. C.; He, J.; Gwadz, M.; Hurwitz, D. I.; et al. CDD: NCBI's Conserved Domain Database. *Nucleic Acids Res.* **2014**, *43* (D1), D222-226. https://doi.org/10.1093/nar/gku1221.

(316) Chow, C.; Cloutier, S.; Dumas, C.; Chou, M. N.; Papadopoulou, B. Promastigote to Amastigote Differentiation of Leishmania Is Markedly Delayed in the Absence of

PERK EIF2alpha Kinase-Dependent EIF2alpha Phosphorylation. *Cell. Microbiol.* **2011**, *13* (7), 1059–1077. https://doi.org/10.1111/j.1462-5822.2011.01602.x.

(317) Cloutier, S.; Laverdière, M.; Chou, M. N.; Boilard, N.; Chow, C.; Papadopoulou, B. Translational Control through EIF2alpha Phosphorylation during the Leishmania Differentiation Process. *PLoS One* **2012**, *7* (5). https://doi.org/10.1371/journal.pone.0035085.

(318) da Silva Augusto, L.; Moretti, N. S.; Ramos, T. C. P.; de Jesus, T. C. L.; Zhang, M.; Castilho, B. A.; Schenkman, S. A Membrane-Bound EIF2 Alpha Kinase Located in Endosomes Is Regulated by Heme and Controls Differentiation and ROS Levels in Trypanosoma Cruzi. *PLoS Pathog.* **2015**, *11* (2), 1–27. https://doi.org/10.1371/journal.ppat.1004618.

(319) Hope, R.; Ben-Mayor, E.; Friedman, N.; Voloshin, K.; Biswas, D.; Matas, D.; Drori, Y.; Günzl, A.; Michaeli, S. Phosphorylation of the TATA-Binding Protein Activates the Spliced Leader Silencing Pathway in Trypanosoma Brucei. *Sci. Signal.* **2014**, *7* (341), 1–9. https://doi.org/10.1126/scisignal.2005234.

(320) Brookman, J. L.; Ozkose, E.; Rogers, S.; Trinci, A. P. J.; Theodorou, M. K. Identification of Spores in the Polycentric Anaerobic Gut Fungi Which Enhance Their Ability to Survive. *FEMS Microbiol. Ecol.* **2000**, *31* (3), 261–267. https://doi.org/10.1111/j.1574-6941.2000.tb00692.x.

(321) Orpin, C. G. Studies on the Rumen Flagellate Neocallimastix Frontalis. *J. Gen. Microbiol.* **1975**, No. 1975, 249–262.

(322) Orpin, C. G. Studies on the Rumen Flagellate Sphaeromonas Communis. *J. Gen. Microbiol.* **1976**, *94* (2), 270–280. https://doi.org/10.1099/00221287-94-2-270.

(323) Lee, Y. J.; Kim, K. J.; Kang, H. Y.; Kim, H. R.; Maeng, P. J. Involvement of GDH3-Encoded NADP+-Dependent Glutamate Dehydrogenase in Yeast Cell Resistance to Stress-Induced Apoptosis in Stationary Phase Cells. *J. Biol. Chem.* **2012**, *287* (53), 44221–44233. https://doi.org/10.1074/jbc.M112.375360.

(324) Kaufman, R. J.; Scheuner, D.; Schröder, M.; Shen, X.; Lee, K.; Liu, C. Y.; Arnold, S. M. The Unfolded Protein Response in Nutrient Sensing and Differentiation. *Nat. Rev. Mol. Cell Biol.* **2002**, *3* (6), 411–421. https://doi.org/10.1038/nrm829.

(325) Brewster, J. L.; de Valoir, T.; Dwyer, N. D.; Winter, E.; Gustin, M. C. An Osmosensing Signal Transduction Pathway in Yeast. *Science* **1993**, *259* (5102), 1760–1763. https://doi.org/10.1126/science.7681220.

(326) Lindquist, S.; Craig, E. A. The Heat -Shock Proteins. *Annu. Rev. Genet* **1988**, *22*, 631–677.

(327) Wu, J.; Wang, M.; Zhou, L.; Yu, D. Small Heat Shock Proteins, Phylogeny in Filamentous Fungi and Expression Analyses in Aspergillus Nidulans. *Gene* **2016**, *575* (2), 675–679. https://doi.org/10.1016/j.gene.2015.09.044.

(328) Schmuths, H.; Meister, A.; Horres, R.; Bachmann, K. Genome Size Variation among Accessions of Arabidopsis Thaliana. *Ann. Bot.* **2004**, *93* (3), 317–321. https://doi.org/10.1093/aob/mch037.

(329) Itoh, T.; Tanaka, T.; Barrero, R. A.; Yamasaki, C.; Fujii, Y.; Hilton, P. B.; Antonio, B. A.; Aono, H.; Apweiler, R.; Bruskiewich, R.; et al. Curated Genome Annotation of Oryza Sativa Ssp. Japonica and Comparative Genome Analysis with Arabidopsis Thaliana: The Rice Annotation Project. *Genome Res.* **2007**, *17* (2), 175–183. https://doi.org/10.1101/gr.5509507.

(330) Padamsee, M.; Kumar, A. T. K.; Riley, R.; Binder, M.; Boyd, A.; Calvo, A. M.; Furukawa, K.; Hesse, C.; Hohmann, S.; James, T. Y.; et al. The Genome of the Xerotolerant Mold Wallemia Sebi Reveals Adaptations to Osmotic Stress and Suggests Cryptic Sexual Reproduction. *Fungal Genet. Biol.* **2012**, *49* (3), 217–226. https://doi.org/10.1016/j.fgb.2012.01.007.

(331) Zajc, J.; Liu, Y.; Dai, W.; Yang, Z.; Hu, J.; Gostinčar, C.; Gunde-Cimerman, N. Genome and Transcriptome Sequencing of the Halophilic Fungus Wallemia Ichthyophaga: Haloadaptations Present and Absent. *BMC Genomics* **2013**, *14* (1). https://doi.org/10.1186/1471-2164-14-617.

(332) Wilken, S. E.; Seppälä, S.; Lankiewicz, T. S.; Saxena, M.; Henske, J. K.; Salamov, A. A.; Grigoriev, I. V.; O'Malley, M. A. Genomic and Proteomic Biases Inform Metabolic Engineering Strategies for Anaerobic Fungi. *Metab. Eng. Commun.* **2020**, *10* (November 2019). https://doi.org/10.1016/j.mec.2019.e00107.

(333) Muralidharan, V.; Oksman, A.; Pal, P.; Lindquist, S.; Goldberg, D. E. Plasmodium Falciparum Heat Shock Protein 110 Stabilizes the Asparagine Repeat-Rich Parasite Proteome during Malarial Fevers. *Nat. Commun.* **2012**, *3*, 1310. https://doi.org/10.1038/ncomms2306.

(334) Muralidharan, V.; Goldberg, D. E. Asparagine Repeats in Plasmodium Falciparum Proteins: Good for Nothing? *PLoS Pathog.* **2013**, *9* (8), 8–11. https://doi.org/10.1371/journal.ppat.1003488.

(335) Quevillon, E.; Silventoinen, V.; Pillai, S.; Harte, N.; Mulder, N.; Apweiler, R.; Lopez, R. InterProScan: Protein Domains Identifier. *Nucleic Acids Res.* **2005**, *33* (SUPPL. 2), 116–120. https://doi.org/10.1093/nar/gki442.

(336) Sievers, F.; Higgins, D. G. Clustal Omega for Making Accurate Alignments of Many Protein Sequences. *Protein Sci.* **2018**, *27* (1), 135–145. https://doi.org/10.1002/pro.3290.

(337) Kumar, S.; Tamura, K.; Nei, M. MEGA: Molecular Evolutionary Genetics Analysis Software for Microcomputers. *Bioinformatics* **1994**, *10* (2), 189–191. https://doi.org/10.1093/bioinformatics/10.2.189.

(338) Steinegger, M.; Söding, J. MMseqs2 Enables Sensitive Protein Sequence Searching for the Analysis of Massive Data Sets. *Nat. Biotechnol.* **2017**, *35* (11), 1026–1028. https://doi.org/10.1038/nbt.3988.

(339) Seedorf, H.; Fricke, W. F.; Veith, B.; Brüggemann, H.; Liesegang, H.; Strittmatter, A.; Miethke, M.; Buckel, W.; Hinderberger, J.; Li, F.; et al. The Genome of Clostridium Kluyveri, a Strict Anaerobe with Unique Metabolic Features. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (6), 2128–2133. https://doi.org/10.1073/pnas.0711093105.

(340) Ant, T.; Magalh, S.; Mendes, D. O.; Santanta, M. F.; Huws, S. A.; Creevey, C. J.; Mantovani, C. Genomic and Gene Expression Evidence of Nonribosomal Peptide and Polyketide Production among Ruminal Bacteria : A Potential Role in Niche Colonization ? **2020**, No. April 2019, 1–13. https://doi.org/10.1093/femsec/fiz198.

(341) Wallace, R. J.; McEwan, N. R.; McIntosh, F. M.; Teferedegne, B.; Newbold, C. J. Natural Products as Manipulators of Rumen Fermentation. *Asian-Australasian J. Anim. Sci.* **2002**, *15* (10), 1458–1468. https://doi.org/10.5713/ajas.2002.1458.

(342) Egan, K.; Field, D.; Rea, M. C.; Ross, R. P.; Hill, C.; Cotter, P. D. Bacteriocins: Novel Solutions to Age Old Spore-Related Problems? *Front. Microbiol.* **2016**, *7*, 461.

https://doi.org/10.3389/fmicb.2016.00461.

(343) Bahar, A.; Ren, D. Antimicrobial Peptides. *Pharmaceuticals* **2013**, *6* (12), 1543–
1575. https://doi.org/10.3390/ph6121543.

(344) Elliott, K. A.; Kenny, C.; Madan, J. A Global Treaty to Reduce Antimicrobial Use in
Livestock. *Cent. Glob. Dev.* **2017**, *102* (January), 27.

(345) Aarestrup, F. M. The Livestock Reservoir for Antimicrobial Resistance: A Personal
View on Changing Patterns of Risks, Effects of Interventions and the Way Forward.
Phil. Trans. R. Soc. B **2015**, *370*, 20140085. https://doi.org/10.1098/rstb.2014.0085.

(346) Hitch, T. C. A.; Thomas, B. J.; Friedersdorff, J. C. A.; Ougham, H.; Creevey, C. J.
Deep Sequence Analysis Reveals the Ovine Rumen as a Reservoir of Antibiotic
Resistance Genes. *Environ. Pollut.* **2018**, *235*, 571–575.
https://doi.org/10.1016/j.envpol.2017.12.067.

(347) Lawson, C. E.; Harcombe, W. R.; Hatzenpichler, R.; Lindemann, S. R.; Löffler, F. E.;
O'Malley, M. A.; García Martín, H.; Pfleger, B. F.; Raskin, L.; Venturelli, O. S.; et al.
Common Principles and Best Practices for Engineering Microbiomes. *Nat. Rev.
Microbiol.* **2019**, *17* (12), 725–741. https://doi.org/10.1038/s41579-019-0255-9.

(348) Tang, X.; Li, J.; Millán-Aguiñaga, N.; Zhang, J. J.; O'Neill, E. C.; Ugalde, J. A.;
Jensen, P. R.; Mantovani, S. M.; Moore, B. S. Identification of Thiotetronic Acid
Antibiotic Biosynthetic Pathways by Target-Directed Genome Mining. *ACS Chem.
Biol.* **2015**, *10* (12), 2841–2849. https://doi.org/10.1021/acschembio.5b00658.

(349) O'Neill, E. C.; Schorn, M.; Larson, C. B.; Millán-Aguiñaga, N. Targeted Antibiotic
Discovery through Biosynthesis-Associated Resistance Determinants: Target Directed
Genome Mining. *Crit. Rev. Microbiol.* **2019**, *45* (3), 255–277.

https://doi.org/10.1080/1040841X.2019.1590307.

(350) Almabruk, K. H.; Dinh, L. K.; Philmus, B. Self-Resistance of Natural Product Producers: Past, Present, and Future Focusing on Self-Resistant Protein Variants. *ACS Chem. Biol.* **2018**, *13* (6), 1426–1437. https://doi.org/10.1021/acschembio.8b00173.

(351) Bugg, T. D. H.; Wright, G. D.; Walsh, C. T.; Dutka-Malen, S.; Arthur, M.; Courvalin, P. Molecular Basis for Vancomycin Resistance in Enterococcus Faecium BM4147: Biosynthesis of a Depsipeptide Peptidoglycan Precursor by Vancomycin Resistance Proteins VanH and VanA. *Biochemistry* **1991**, *30* (43), 10408–10415. https://doi.org/10.1021/bi00107a007.

(352) Daniele Billot-Klein, Didier Blanot, L. G. and J. V. H. Association Constants for the Binding of Vancomycin and Teicoplanin to N-Acetyl-D-Alanyl-D-Alanine and N-Acetyl-D-Alanyl-D-Serine. *Biochem. J.* **1994**, *304*, 1021–1022.

(353) Mungan, M. D.; Alanjary, M.; Blin, K.; Weber, T.; Medema, M. H.; Ziemert, N. ARTS 2.0: Feature Updates and Expansion of the Antibiotic Resistant Target Seeker for Comparative Genome Mining. *Nucleic Acids Res.* **2020**, *48* (W1), W546–W552. https://doi.org/10.1093/nar/gkaa374.

(354) Stewart, R. D.; Auffret, M. D.; Warr, A.; Wiser, A. H.; Press, M. O.; Langford, K. W.; Liachko, I.; Snelling, T. J.; Dewhurst, R. J.; Walker, A. W.; et al. Assembly of 913 Microbial Genomes from Metagenomic Sequencing of the Cow Rumen. *Nat. Commun.* **2018**, *9* (1), 1–11. https://doi.org/10.1038/s41467-018-03317-6.

(355) Svartström, O.; Alneberg, J.; Terrapon, N.; Lombard, V.; De Bruijn, I.; Malmsten, J.; Dalin, A. M.; El Muller, E.; Shah, P.; Wilmes, P.; et al. Ninety-Nine de Novo Assembled Genomes from the Moose (Alces Alces) Rumen Microbiome Provide

New Insights into Microbial Plant Biomass Degradation. *ISME J.* **2017**, *11* (11), 2538–2551. https://doi.org/10.1038/ismej.2017.108.

(356) 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.; et al. Interspecies Cross-Feeding Orchestrates Carbon Degradation in the Rumen Ecosystem. *Nat. Microbiol.* **2018**, *3* (11), 1274–1284. https://doi.org/10.1038/s41564-018-0225-4.

(357) 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.; et al. Cultivation and Sequencing of Rumen Microbiome Members from the Hungate1000 Collection. *Nat. Biotechnol.* **2018**, 1–7. https://doi.org/10.1038/nbt.4110.

(358) Alexander Crits-Christoph, Spencer Diamond, Christina N. Butterfield, Brian C. Thomas, J. F. B. Novel Soil Bacteria Possess Diverse Genes for Secondary Metabolite Biosynthesis. *Crit. Rev. Biotechnol.* **2018**. https://doi.org/10.1038/s41586-018-0207-y.

(359) 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.; et al. Cultivation and Sequencing of Rumen Microbiome Members from the Hungate1000 Collection. *Nat. Biotechnol.* **2018**, No. August 2017. https://doi.org/10.1038/nbt.4110.

(360) Clark, K.; Karsch-Mizrachi, I.; Lipman, D. J.; Ostell, J.; Sayers, E. W. GenBank. *Nucleic Acids Res.* **2016**, *44* (D1), D67–D72. https://doi.org/10.1093/nar/gkv1276.

(361) Kautsar, S. A.; Blin, K.; Shaw, S.; Navarro-Muñoz, J. C.; Terlouw, B. R.; van der Hooft, J. J. J.; van Santen, J. A.; Tracanna, V.; Suarez Duran, H. G.; Pascal Andreu, V.; et al. MIBiG 2.0: A Repository for Biosynthetic Gene Clusters of Known

Function. *Nucleic Acids Res.* **2020**, *48* (D1), D454–D458.

https://doi.org/10.1093/nar/gkz882.

(362) Schöner, T. A.; Gassel, S.; Osawa, A.; Tobias, N. J.; Okuno, Y.; Sakakibara, Y.;

Shindo, K.; Sandmann, G.; Bode, H. B. Aryl Polyenes, a Highly Abundant Class of

Bacterial Natural Products, Are Functionally Related to Antioxidative Carotenoids.

ChemBioChem **2016**, *17* (3), 247–253. https://doi.org/10.1002/cbic.201500474.

(363) Rateb, M. E.; Hallyburton, I.; Houssen, W. E.; Bull, A. T.; Goodfellow, M.;

Santhanam, R.; Jaspars, M.; Ebel, R. Induction of Diverse Secondary Metabolites in

Aspergillus Fumigatus by Microbial Co-Culture. *RSC Adv.* **2013**, *3* (34), 14444–

14450. https://doi.org/10.1039/c3ra42378f.

(364) Cui, L.; Morris, A.; Ghedin, E. The Human Mycobiome in Health and Disease.

Genome Med. **2013**, *5* (7), 1–12. https://doi.org/10.1186/gm467.

(365) Huttenhower, C.; Gevers, D.; Knight, R.; Abubucker, S.; Badger, J. H.; Chinwalla, A.

T.; Creasy, H. H.; Earl, A. M.; FitzGerald, M. G.; Fulton, R. S.; et al. Structure,

Function and Diversity of the Healthy Human Microbiome. *Nature* **2012**, *486* (7402),

207–214. https://doi.org/10.1038/nature11234.

(366) Den Besten, G.; Van Eunen, K.; Groen, A. K.; Venema, K.; Reijngoud, D. J.; Bakker,

B. M. The Role of Short-Chain Fatty Acids in the Interplay between Diet, Gut

Microbiota, and Host Energy Metabolism. *J. Lipid Res.* **2013**, *54* (9), 2325–2340.

https://doi.org/10.1194/jlr.R036012.

(367) Ishiuchi, K.; Nakazawa, T.; Ookuma, T.; Sugimoto, S.; Sato, M.; Tsunematsu, Y.;

Ishikawa, N.; Noguchi, H.; Hotta, K.; Moriya, H.; et al. Establishing a New

Methodology for Genome Mining and Biosynthesis of Polyketides and Peptides

through Yeast Molecular Genetics. *ChemBioChem* **2012**, *13* (6), 846–854. https://doi.org/10.1002/cbic.201100798.

(368) Hühner, E.; Backhaus, K.; Kraut, R.; Li, S. M. Production of α-Keto Carboxylic Acid Dimers in Yeast by Overexpression of NRPS-like Genes from Aspergillus Terreus. *Appl. Microbiol. Biotechnol.* **2018**, *102* (4), 1663–1672. https://doi.org/10.1007/s00253-017-8719-1.

(369) Gruenewald, S.; Mootz, H. D.; Stehmeier, P.; Stachelhaus, T. In Vivo Production of Artificial Nonribosomal Peptide Products in the Heterologous Host Escherichia Coli. *Appl. Environ. Microbiol.* **2004**, *70* (6), 3282–3291. https://doi.org/10.1128/AEM.70.6.3282-3291.2004.

(370) Alberti, F.; Foster, G. D.; Bailey, A. M. Natural Products from Filamentous Fungi and Production by Heterologous Expression. *Appl. Microbiol. Biotechnol.* **2017**, *101* (2), 493–500. https://doi.org/10.1007/s00253-016-8034-2.

(371) Rodriguez, E.; Menzella, H. G.; Gramajo, H. Heterologous Production of Polyketides in Bacteria. *Methods Enzymol.* **2009**, *459*, 339–365. https://doi.org/10.1016/S0076-6879(09)04615-1.

(372) Calkins, S. S.; Elledge, N. C.; Mueller, K. E.; Marek, S. M.; Couger, M. B.; Elshahed, M. S.; Youssef, N. H. Development of an RNA Interference (RNAi) Gene Knockdown Protocol in the Anaerobic Gut Fungus Pecoramyces Ruminantium Strain C1A. *PeerJ* **2018**, *2018* (1), 1–25. https://doi.org/10.7717/peerj.4276.

(373) Kjærbølling, I.; Vesth, T. C.; Frisvad, J. C.; Nybo, J. L.; Theobald, S.; Kuo, A.; Bowyer, P.; Matsuda, Y.; Mondo, S.; Lyhne, E. K.; et al. Linking Secondary Metabolites to Gene Clusters through Genome Sequencing of Six Diverse Aspergillus

Species. *Proc. Natl. Acad. Sci.* **2018**, 201715954.

https://doi.org/10.1073/pnas.1715954115.

(374) Henrikson, J. C.; Hoover, A. R.; Joyner, P. M.; Cichewicz, R. H. A Chemical

Epigenetics Approach for Engineering the in Situ Biosynthesis of a Cryptic Natural

Product from Aspergillus Niger. *Org. Biomol. Chem.* **2009**, *7* (3), 435–438.

https://doi.org/10.1039/b819208a.

(375) Fisch, K. M.; Gillaspy, A. F.; Gipson, M.; Henrikson, J. C.; Hoover, A. R.; Jackson,

L.; Najar, F. Z.; Wägele, H.; Cichewicz, R. H. Chemical Induction of Silent

Biosynthetic Pathway Transcription in Aspergillus Niger. *J. Ind. Microbiol.*

Biotechnol. **2009**, *36* (9), 1199–1213. https://doi.org/10.1007/s10295-009-0601-4.