DEVELOPING ANAEROBIC FUNGI AS A PLATFORM FOR EFFICIENT LIGNOCELLULOSE HYDROLYSIS

by

Casey Hooker

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Department of Agricultural and Biological Engineering

West Lafayette, Indiana

December 2018

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Kevin V. Solomon, Chair

    Department of Agricultural and Biological Engineering

Dr. Michael R. Ladisch

    Department of Agricultural and Biological Engineering

Dr. Nathan S. Mosier

    Department of Agricultural and Biological Engineering

Dr. Eduardo A. Ximenes

    Department of Agricultural and Biological Engineering

Approved by:

    Dr. Bernard Engel

        Head of the Graduate Program

For my parents

ACKNOWLEDGMENTS

I thank Dr. Kevin Solomon for his guidance and support of me as a graduate student. I am grateful for his patience when science and writing occurred slower than we anticipated. I am honored to have him as my advisor, and I hope that I can continue to learn from him in the future.

I am grateful for the many people who have provided feedstocks, equipment, or expertise for this work: Dr. Cliff Weil, Dr. Richard Meilan, Dr. Nathan Mosier, Dr. Bernard Engel, Dr. Klein Illeleji, C3Bio, Xingya Liu, LORRE, and the Indianapolis Zoo. Additionally, I owe many thanks to Blaine Brown and the Boiler Butcher Block for the opportunity to harvest rumen fluid so that we could culture these unique microorganisms.

Lastly, I am thankful for the farmers that inspired me to strive to make a difference in the agricultural community. Clifford Selby, Gary Wiltfang, and especially my father all have provided a great deal of insight on the challenges facing those who farm.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

Bovine serum albumin: BSA

Carbohydrate active enzyme: CAZyme

Carbohydrate binding protein: CBP

Carbon catabolite repression: CCR

Carboxy methyl cellulose: CMC

Dimethyl sulfoxide: DMSO

DNA methyltransferase: DNMT

Guiacyl lignin: G lignin

Histone acetyltransferase: HAT

Histone deacetylase: HDAC

Internal transcribed spacer: ITS

Neocallimastix sp. Gf-ma: Gf-ma

Neocallimastix sp. WI3B: WI3B

Para-hydroxy phenyl lignin: H Lignin

Piromyces sp. UH3-1: UH3-1

Suberoylanilide hydroxamic acid: SAHA

Syringyl lignin: S lignin

4' 6 diamidino-2-phenylindole: DAPI

28s rRNA large ribosomal subunit: LSU

ABSTRACT

Author: Hooker, Casey MS
Institution: Purdue University
Degree Received: December 2018
Title: Developing Anaerobic Gut Fungi as a Platform for Efficient Lignocellulose Hydrolysis
Major Professor: Dr. Kevin Solomon

Lignocellulose is an ubiquitous source of fixed carbon that is presently underexploited for renewable energy technologies. Currently, producing enzyme cocktails that robustly degrade these feedstocks is a significant economic bottleneck. Anaerobic gut fungi native to the digestive tracts of ruminants and hindgut fermenters are widely understudied despite their inherent ability to degrade a significant portion (~50%) of the lignocellulose in herbivorous animals. Challenges in cultivation due to their strict oxygen sensitivity, and the lack of a central repository to maintain axenic stocks substantially impede the progress with anaerobic fungi. Yet, these microbes have evolved elegant strategies and may harbor novel biomass degrading enzymes that could be used to more efficiently hydrolyze lignocellulose. Developing these organisms through characterization and genome engineering will yield significant contributions to the bioenergy community by improving hydrolysis technologies.

In this work, we report the isolation of four novel species of anaerobic gut fungi. A more complete characterization of one of our four fungal isolates is investigated, whereby the effects of substrate composition and the corresponding fungal growth rates are compared. I also explore the growth of one of our fungal isolates on transgenic poplar to understand how fungal growth and enzyme secretion adapt to variable lignin composition. Notably, no significant reductions in growth were observed highlighting the ability of anaerobic fungi to degrade diverse feedstocks regardless of lignin composition. I have additionally included preliminary work intended to identify what epigenetic regulational strategies exist for anaerobic fungi, and how they relate to carbohydrate active enzyme expression. We hope to leverage this knowledge to engineer base enzyme cocktails that release significant portions of the fermentable sugars in untreated or mildly treated plant biomass as a means to make bioenergy technologies more efficient.

1. LIGNOCELLULOSIC BIOMASS AND MICROBIAL SYSTEMS FOR DEPOLYMIZING PLANT CELL WALL POLYMERS

1.1 Motivation and problem statement

Lignocellulose represents an ubiquitous source of fixed carbon that is widely underused (<2%) despite the enormous annual production, which has been estimated to be between 150-170 $* 10^9$ tons [1]. As plant biomass contains fermentable sugars and fixed carbon, readily hydrolyzing these feedstocks into the corresponding monomers would provide a renewable and inexpensive substrate for bioenergy and biosynthetic production. Yet producing enzyme cocktails that efficiently depolymerize this material is a formidable challenge and a significant bottleneck in current technologies. Thus there is a critical need to develop enzyme systems that more efficiently release the monomers of lignocellulose so that this feedstock can be better exploited for biochemical production.

One approach to developing improved enzyme platforms is to look at natural sources for biomass degrading enzymes. Anaerobic gut fungi in particular are well suited to hydrolyze untreated plant material as these microorganisms are native to the digestive tracts of ruminant and hindgut fermenting animals. Gut fungi harbor significantly more lignocellulolytic enzymes than the fungi that are currently used to produce industrial enzyme preparations. Yet anaerobic fungi are heavily understudied given the equipment requirements and the need for a robust genome engineering toolkit. Hence, developing anaerobic fungi as a platform to breakdown diverse untreated lignocellulose may help to valorize this feedstock and provide an inexpensive source for renewable energy and chemical production.

1.2 The structure and complexity of lignocellulose

Lignocellulose is critical for plants and trees, as it provides structural support and rigidity that allows for vertical growth and accumulation of higher biomass yields [2]. Consequently many flowering plants have evolved strategies to hinder deconstruction of their cell walls (e.g. diverse monomeric components and unique bonds between them) as a form of protection from saprobic microorganisms. Therefore, efficiently hydrolyzing plant matter requires significant numbers of carbohydrate active enzymes that readily release the monomeric components in these diverse cell

wall polymers. A great amount of enzymatic research has focused on cellulolytic enzymes as cellulose is a core component of plant biomass which contains both crystalline and amorphous polymers of $\beta1{\rightarrow}4$ linked glucose (Figure 1.1) [3-6]. Thus, efficiently hydrolyzing this polymer releases sugars that are immediately integrated into central metabolism of the fermenting organism. Yet other polymers besides cellulose also contain monomers that may be metabolized for bioenergy production [7]. For example, hemicellulose contains many sugars, yet the structure and composition of this polymer vary significantly. Hemicellulose contains multiple monomers such as xylose, arabinose, galactose, glucose, and mannose that are bonded in variable sequences [8]. This variability leads to the formation of distinct polymers (e.g. xyloglucan, glucuronoxylan, glucuronoarabinoxylan, galactomannan, glucomannan, and β-(1${\rightarrow}$3 or 1${\rightarrow}$4)-glucans) which requires an array of enzymes to deconstruct (Figure 1.1) [8]. Importantly, these diverse hemicellulosic polymers have highly variable abundance, both in location of the plant cell wall (primary and secondary), and also vary between plant species (e.g. dicots, grasses, conifer plants) [8]. Taken together, the fermentable sugars of renewable plant biomass form complex polymers that partly contribute to the recalcitrance of lignocellulose by requiring significant numbers of carbohydrate active enzymes to efficiently deconstruct.

While understanding the structure of the plant cell wall components containing fermentable sugars, and the corresponding enzymes that hydrolyze them is key, accessing them for depolymerization is extraordinarily challenging due to lignin [9]. Lignin is composed of heterophenolic polymers that are covalently linked to hemicellulose, and increases the structural complexity of plant biomass [10]. These heterophenolic polymers are comprised of three unique monolignol units: para-coumeryl (H), coniferyl (G), and sinapyl (S) alcohols, which have different numbers of methoxy ring substitutions (Figure 1.1) [11]. Notably, these additional methoxy groups increase the number of crosslinks thus reinforcing the biomass making it more challenging to hydrolyze. While some organisms, such as white rot fungi secrete enzymes capable of partially hydrolyzing or modifying lignin, many organisms are strongly inhibited by it and are therefore unable to efficiently degrade untreated plant biomass [12]. Lignin is one of the most formidable components of lignocellulose that prevents the development of renewable plant biomass as an economical substrate for bioenergy production [13, 14]. In summary, lignin is a complex component of plant biomass that should be considered when designing platforms for hydrolyzing lignocellulose to access and release the fixed carbon for biofuel and biosynthetic platforms.
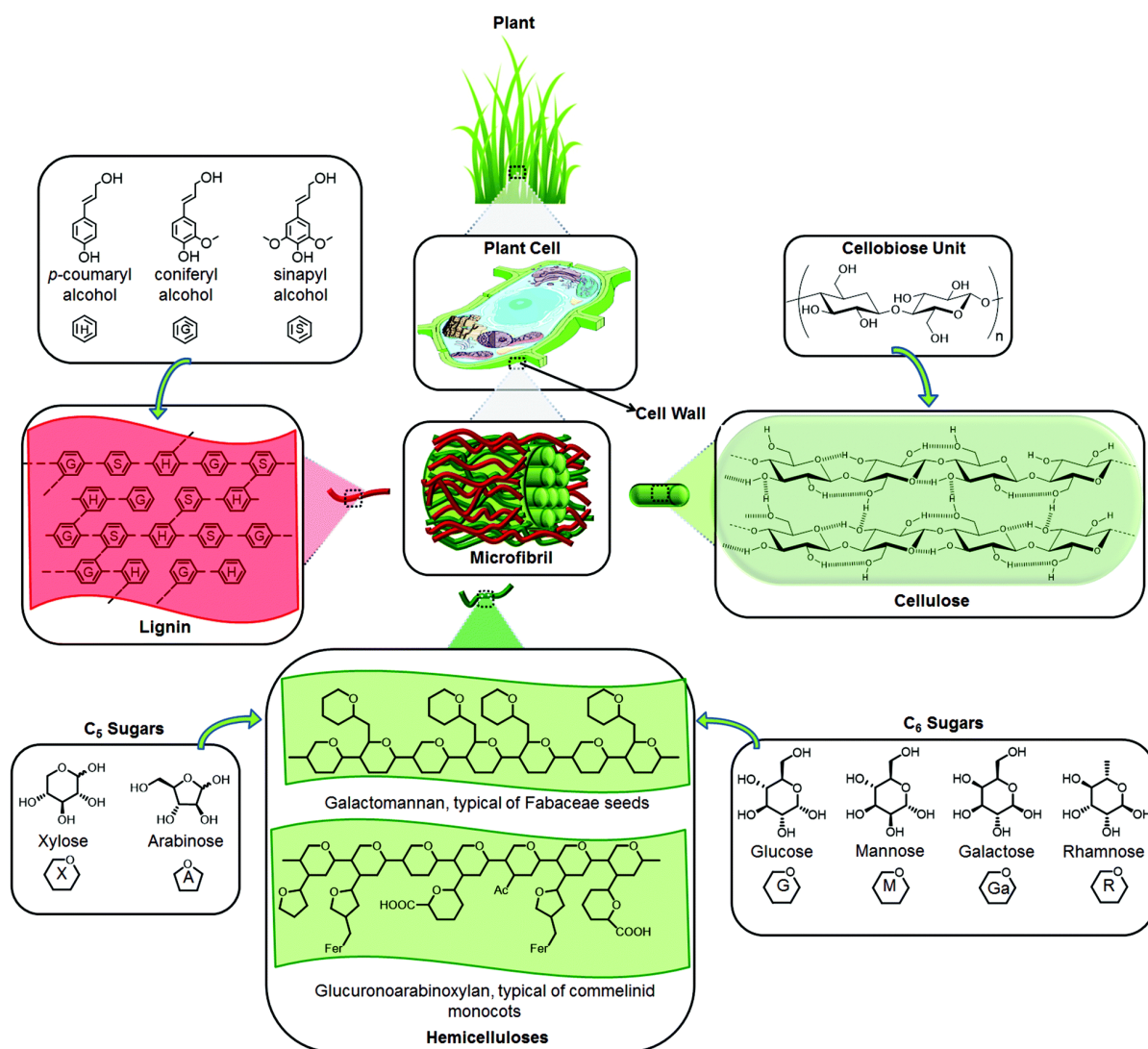
Figure 1.1: Composition of the plant cell wall [15]. Reproduced without alterations under the Creative Commons BY 3.0 license.

1.3 Lignocellulosic feedstock diversity

Many routes are pursued to overcome lignocellulose recalcitrance; however most are not suitable for all forms of plant biomass. The variable abundance and composition of lignocellulose significantly changes the feasibility and efficacy of a given technology for improving hydrolysis. One approach is to select certain varieties of crops that have desirable traits for bioenergy production, such as plants with relatively low lignin content and or high sugar content (e.g. switchgrass, miscanthus, sorghum, hybrid poplars) [16]. Some of these energy feedstocks are attractive for bioenergy production as they grow on both arable and marginal land, which is

generally regarded as unconducive for crop growth due to edaphic or climate limitations [17]. Similarly, many of these bioenergy crops have established genome engineering toolkits allowing for their composition to be further manipulated to enhance digestibility. For example, transgenic poplar containing low molar ratios of syringyl lignin improves digestibility by both white rot and brown rot fungi [18]. However, even if engineered strains of biomass are used, there still may be a need to use unengineered plant biomass. Changing climate and economic conditions will have significant impacts on the types of feedstocks that are available for bioenergy production which would be expected to fluctuate seasonally therefore contributing to substrate production instability (Figure 1.2) [19]. Yet, there are many issues that are associated with energy crops that are widely overlooked.
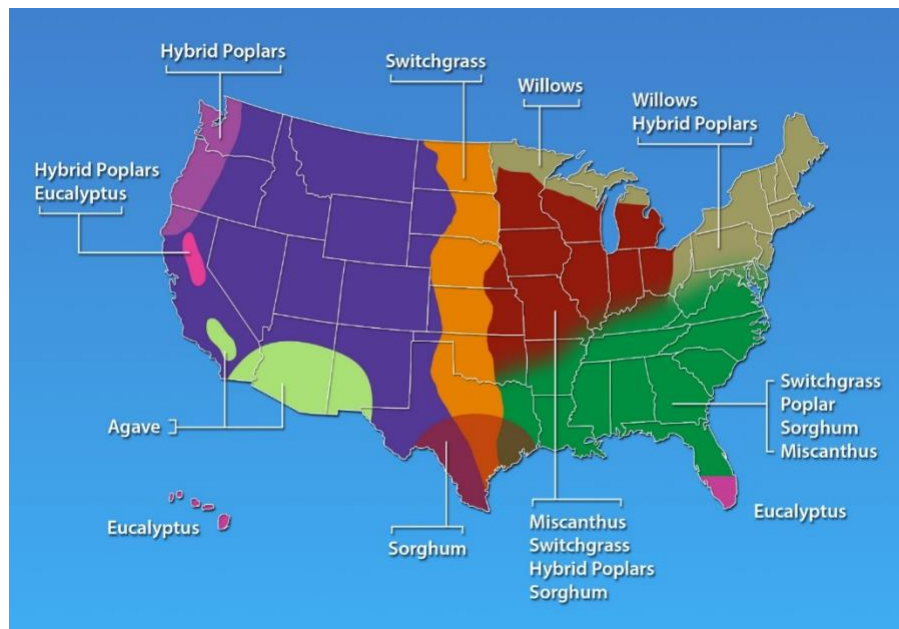
Figure 1.2: Bioenergy crop production potential in the United States as of 2015 [20].

These crops take multiple years to establish prior to being harvestable, weather and economic conditions may vary, additionally equipment requirements and return on investment significantly limits farmer involvement [21-23]. Thus, using energy crops, and or engineered lignocellulose may increase the feasibility for improving hydrolysis efficiencies, however changing market and climate conditions will significantly influence the extent to which these feedstocks are available. As a result, there is a critical need to incorporate diverse streams of agricultural wastes, forestry

residues, and possibly municipal solid wastes (e.g. paper) to mitigate issues substrate availability for these processes. In conclusion, developing mixed streams of lignocellulose as a substrate for bioenergy production requires an enzyme platform that efficiently degrades these complex and variable feedstocks without significant reductions in enzyme efficiency.

1.4   Depolymerizing lignocellulose

Efficiently hydrolyzing plant biomass into the component monomers remains as one of the key challenges facing the bioenergy community [24]. Given the substantial variation in cell wall structure (Figure 1.1), and the substrate diversity (Figure 1.2), a plethora of enzymes are needed for efficient hydrolysis [8, 9]. While bacteria harbor some of the required enzymes, fungi harbor significantly more (Table 1.1), and thus provide a more complete source for base enzyme preparations. Despite the large number of lignocellulolytic enzymes in many of the fungi used to produce these industrial preparations (e.g. *Trichoderma* and *Aspergillus*), they still have deficiencies that result in poor sugar release across highly diverse feedstocks. This phenomenon is partly due to poor CAZyme expression or the absence of a critical CAZyme in their genomes. For example, *Trichoderma reesei* QM6A expresses β- glucosidase, an enzyme essential to hydrolyzing cellobiose to glucose, at very low levels [25]. Similarly, the cellulases from *Aspergillus spp*. have high β-glucosidase activity, but low endoglucanase levels, and therefore also have limited efficiency in hydrolyzing lignocellulose [26]. Thus, enzymes from other organisms could be cloned into either of these fungi, or these fungal secretomes could be complemented with enzymes from other studied fungi and or bacteria to complete the cocktail. Identifying CAZymes from other organisms that supplement these deficiencies may be done by prospecting for novel enzymes having industrial characteristics (e.g. thermostability, pH stability, low product inhibition) [27]. This would not only lead to the identification of enzymes having increased activities, but may also lead to the discovery of novel enzymes that enhance biomass breakdown (i.e. accessory proteins or auxiliary enzymes). However, this is problematic as it leads to multi-species enzyme platforms that significantly increase the production cost [28]. Alternatively, these preparations may be refined by interfering with gene regulation; for example, promoter engineering could be used to increase gene expression [29].

Table 1.1: Fungi harbor diverse CAZymes capable of hydrolyzing plant polysaccharides.
Adapted from [12].

Enzyme Activity	Substrate	CAZyme family
α-L Arabinofuranosidase	Xyloglucan, xylan, pectin	GH51, 54
Endoarabinase	Pectin	GH43
Exoarabinase	Pectin	GH93
α-L Fucosidase	Xyloglucan	GH29,95
α-1,4-D Glucosidase	Starch	GH31
α-1,4-D Galactosidase	Xyloglucan, xylan, galactomannan	GH27,36
α-Glucuronidase	Xylan	GH67,115
α-Amylase	Starch	GH13
Acetyl xylan estease	Xylan	CE1
Arabinoxylan arabinofuranohydrolase	Xylan	GH62
α-D Xylosidase	Xyloglucan	GH31
β-1,4-D Glucosidase	Cellulose, xyloglucan	GH1,3,9
β-1,4-D Xylosidase	Xylan, pectin	GH31
Cellobiohydrolase	Celluose, xyloglucan	GH6,7,9
β-1,4-D Endoglucanase	Cellulose, xyloglucan	GH45,7,9,12,44,45
Glucuronyl esterase	Xylan	CE1
Glucoamylase	Starch	CH15
β-1,6 Endogalactanase	Pectin	GH5,30
Galactomannan acetyl esterase	Galactomannan	N/A
Endo-inulinase	Inulin	GH32
Exo-inulinase	Inulin	GH32
β-1,4 D Galactosidase	Xyloglucan, xylan, pectin, galactomannan	GH2,35,42
β-1,4 D Endomannanase	Galactomannan	GH5,26
β-1,4 D Mannosidase	Galactomannan	GH2
Pectin acetyl esterase	Pectin	N/A
Pectin lyase	Pectin	PL1
Rhamnogalacturonan lyase	Pectin	PL4,11
Endopolygalacturonase	Pectin	GH28
Exopolygalacturonase	Pectin	GH28
Pectate lyase	Pectin	PL1,3,9
Pectin methyl esterase	Pectin	CE8
Rhamnogalaturonan acetyl esterase	Pectin	CE12
Rhamnogalacturonan galaturonohydrolase/exorhamnogalacturonase	Pectin	GH28
α-Rhamnosidase/rhamnogalacturonan rhamnohydrolase	Pectin	GH78
Rhamnogalacturonan hydrolase endorhamnogalacturonase	Pectin	GH28
Invertase/fructofuranosidase	Inulin	GH32
d-4,5 Unsaturated-glucuronyl hydrolase	Pectin	GH105
Xyloglucan-active β-1,4-D endoglucanase	Xyloglucan	GH12,74
Xyloglucan acetyl esterase	Xyloglucan	N/A
Xylgalacturonan hydrolase	Pectin	GH28
β-1,4-D Endoxylanase	Xylan	GH10,11
β-1,3-D Exoxylanase	Xylan	GH43

While this is extremely useful, a deep knowledge of gene regulation networks is required (e.g. complete genome sequence, regulatory protein identification, transcription factor binding sites). Hence only a few fungi have really been pursued for their biomass degrading abilities given the need for such an enormous amount of knowledge on gene regulation [30-32]. Although, even when base cocktails from *Trichoderma* or *Aspergillus spp*. are used, their formulations are frequently optimized for specific substrates (e.g. poplar, corn stover, miscanthus), as the biomass composition significantly affects cocktail performance [33]. However, these enzymes compose a significant portion of the overall production cost for biofuels [28, 34, 35]. As feedstock supply would be expected to fluctuate with market conditions, lignocellulolytic cocktails may need to be refined based upon what is available and economical [36]. Thus, while manually optimizing enzyme

formulations for a given feedstock is helpful, an alternative solution would be to identify new enzymes or engineer more complete base enzyme cocktails that are highly active on diverse lignocellulose, thus mitigating the need to alter formulations based on feedstock availability.

Building enzyme preparations capable of degrading these diverse lignocellulose streams must also consider the need for pretreatment, and thus highlights the other fundamental problems in formulating cocktails on 1 or 2 fungal species. As *Trichoderma* and *Aspergillus* are not primary colonizers of plant biomass, their enzymes may not naturally hydrolyze large percentages of freshly harvested lignocellulose [37]. Consequently biomass pretreatment is employed to physically disrupt the lignocellulose structure and expose the fermentable sugars thus increasing conversion efficiencies [38]. In the absence of pretreatment, sugar conversion efficiencies can be as low as 20% while pretreating biomass can increase this number to as high as 90% [39]. However, these pretreatment technologies typically exploit high temperatures, high pressures, and increase the number of waste streams, all contributing to a higher production cost [38]. Similarly, as multiple types of pretreatment are available (e.g. ammonia fiber explosion, liquid hot water, dilute sulfuric acid) different components of the plant biomass are released and hydrolyzed [38, 40]. Thus the enzyme cocktails needed to hydrolyze the pretreated biomass may differ depending upon the pretreatment selected [41]. In conclusion, biomass pretreatment is an effective strategy to improve the sugar conversion efficiencies when current industrial enzyme preparations are used for hydrolysis although it increases production cost. An alternative strategy is to engineer new enzyme preparations which are adapted to hydrolyze untreated or mildly pretreated plant material.

An uncommon approach to circumventing these challenges is to develop other species of fungi as sources for base enzyme cocktails. As fungi are regarded for their ability to hydrolyze plant material given the range of CAZymes present in many of their genomes, characterizing understudied lineages provides a path forward to accomplish this task. However, since fungi have more complex genomes compared to bacteria, the number of robust genome engineering tools is low and have only been devised for some lineages of lignocellulolytic fungi [41, 42]. Thus many other fungi are not considered due to lack of facile tools and poor understanding of their expression systems. Yet some fungal lineages are particularly well suited to hydrolyze untreated lignocellulose such as fungi of the phylum Neocallimastigomycota, which are native to the digestive tracts of ruminants and hindgut fermenters [43]. Despite the knowledge that these organisms degrade as much as 50% of the untreated lignocellulose while only accounting for 5-7%

of the total microbiome of their respective hosts, they are still heavily understudied [44-46]. This is partly due to the challenges in culturing in a strict anaerobic environment and the absence of a central repository to maintain isolated species. Further, as of 2018 there are only 5 completely sequenced genomes [47]. Together, characterizing, developing, and engineering anaerobic fungi as a platform for a novel enzyme cocktail may provide a means by which diverse mildly treated lignocellulose can be more efficiently hydrolyzed to significantly improve current bioenergy technologies.

1.5 An overview of anaerobic gut fungi

Early diverging anaerobic fungi are primary colonizers of plant biomass and fibrous material ingested by herbivorous animals, and are thus poised to degrade lignocellulose without the need for significant pretreatments [48]. Further, these fungi have complex lifecycles that allow them to thrive in the highly competitive environments in herbivorous animals (Figure 1.3) [49]. Gut fungi reproduce asexually by releasing zoospores from mature sporangia, and these zoospores exhibit chemotaxis to soluble sugars and or phenolic acids which are presumably released by damaged plant tissues [50, 51]. Upon attaching to the plant material, the zoospores encyst and later germinate, upon which extensive rhizoids are developed. These rhizoidal networks penetrate the biomass allowing for significant hydrolysis of the fermentable sugars. Consequently, the complex lifecycles of anaerobic fungi provide an evolutionary adaptation to help degrade these recalcitrant feedstocks.
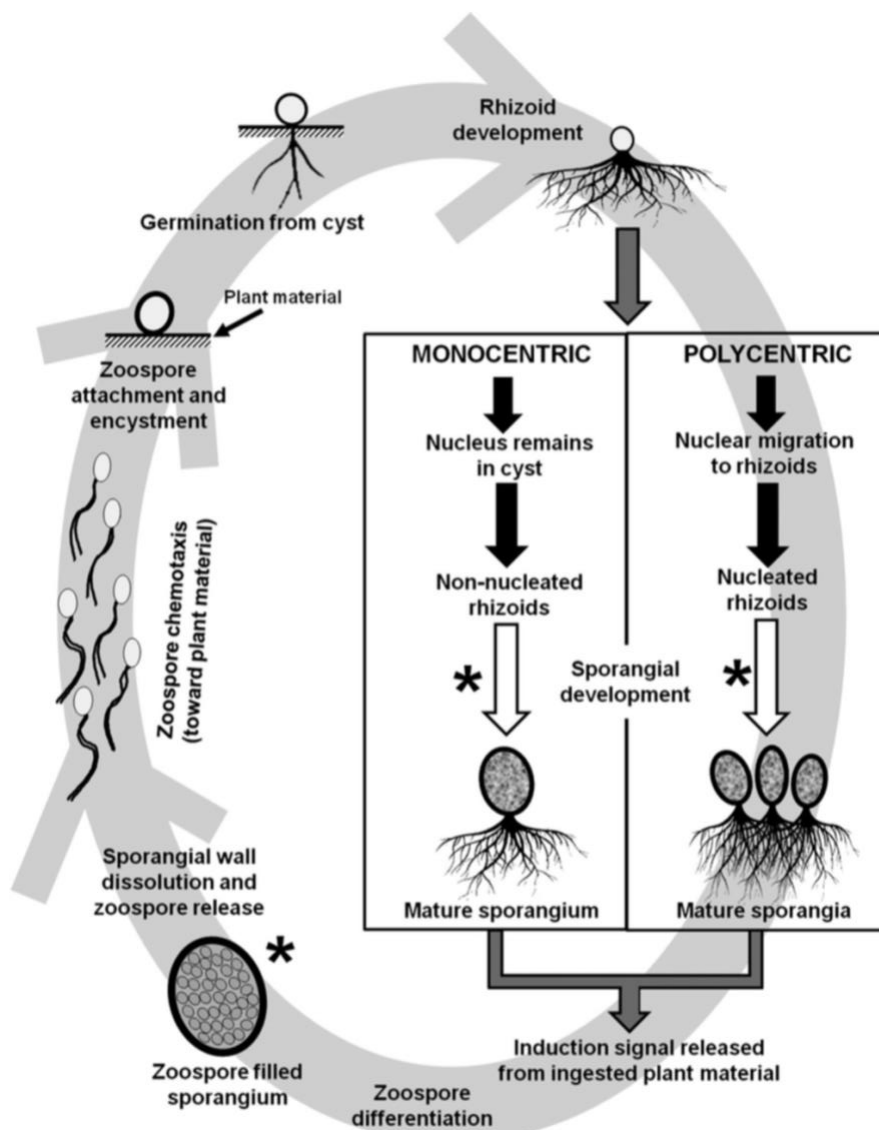
Figure 1.3: The lifecycle of anaerobic fungi. Asterisks denote where aero-tolerant structures have been reported [49]. Reprinted with permission from "Anaerobic fungi (phylum *Neocallimastigomycota*): advances in understanding their taxonomy, life cycle, ecology, role and biotechnological potential" by Robert Gruninger et al 2014. FEMS Microbiology Ecology, Volume 90, 1. 2014 by Oxford University Press.

While the knowledge of the lifecycles of these organisms is relatively thorough across nine genera of fungi within the division Neocallimastigomycota (Table 1.2), relatively few of these organisms have had their genomes completely sequenced. Further, the sequenced anaerobic fungal genomes are of relatively poor quality given the extremely high AT bias and the challenges

associated with harvesting fungal DNA [52]. Therefore, not only is the annotation quality extremely low, the ability to pursue functional genomics is substantially hindered. Yet even with these associated challenges, insight into the available genomes highlights the biomass degrading potential of anaerobic fungi [53]. Additionally, this leaves significant space to engineer these organisms for enhanced biomass breakdown. Together, there is a need to completely sequence more anaerobic fungi and increase genome annotation qualities.

Table 1.2: A brief overview of the known genera of anaerobic fungi as of 2018 (phylum Neocallimastigomycota). Adapted from [54].

Genus name	Zoospore type	Thallus type	Genome available	Citation
Anaeromyces	Uniflagellate	Polycentric	Yes	Novotna et al 2010 [55]
Buwchfawromyces	Monoflagellate	Monocentric	No	Callaghan et al 2015 [56]
Caecomyces	Uniflagellate	Monocentric	No	Orpin 1976 [57]
Cyllamyces	Uniflagellate	Polycentric	No	Ozkose et al 2001 [58]
Neocallimastix	Multiflagellate	Monocentric	Yes	Orpin 1975 [59]
Oontomyces	Uniflagellate	Monocentric	No	Dagar et al 2015 [60]
Orpinomyces	Multiflagellate	Polycentric	Yes	Borneman et al 1989 [61]
Pecoramyces	Uniflagellate	Monocentric	No	Hanafy et al 2017 [62]
Piromyces	Uniflagellate	Monocentric	Yes	Orpin 1981 [63]

Sequencing the genomes of these organisms has substantially increased knowledge of how gut fungi have adapted to breakdown plant material. The available data shows that the five sequenced isolates of anaerobic fungi encode substantially larger numbers of biomass degrading enzymes than the strains of *Trichoderma reesei* or *Aspergillus* that are used for current base platforms for industrial enzyme preparations (Figure 1.4) [47]. While this is partly due to unique evolutionary strategies for addressing the complexity of plant biomass (i.e. ascomycetes freely secrete their enzymes whereas anaerobic fungi form cellulosomes), this highlights the potential of anaerobic fungi to form a more complete enzyme cocktail that degrades complex and mildly treated lignocellulosic feedstocks [64]. Together, anaerobic fungi encode an enormous number of biomass degrading enzymes that may be capable of addressing the issues associated with current enzyme technologies, and allow for improved enzyme preparations to be engineered. To this end,

there is a need to understand the regulation of these enzymes, and how they interact to form cellulosomes for biomass breakdown.
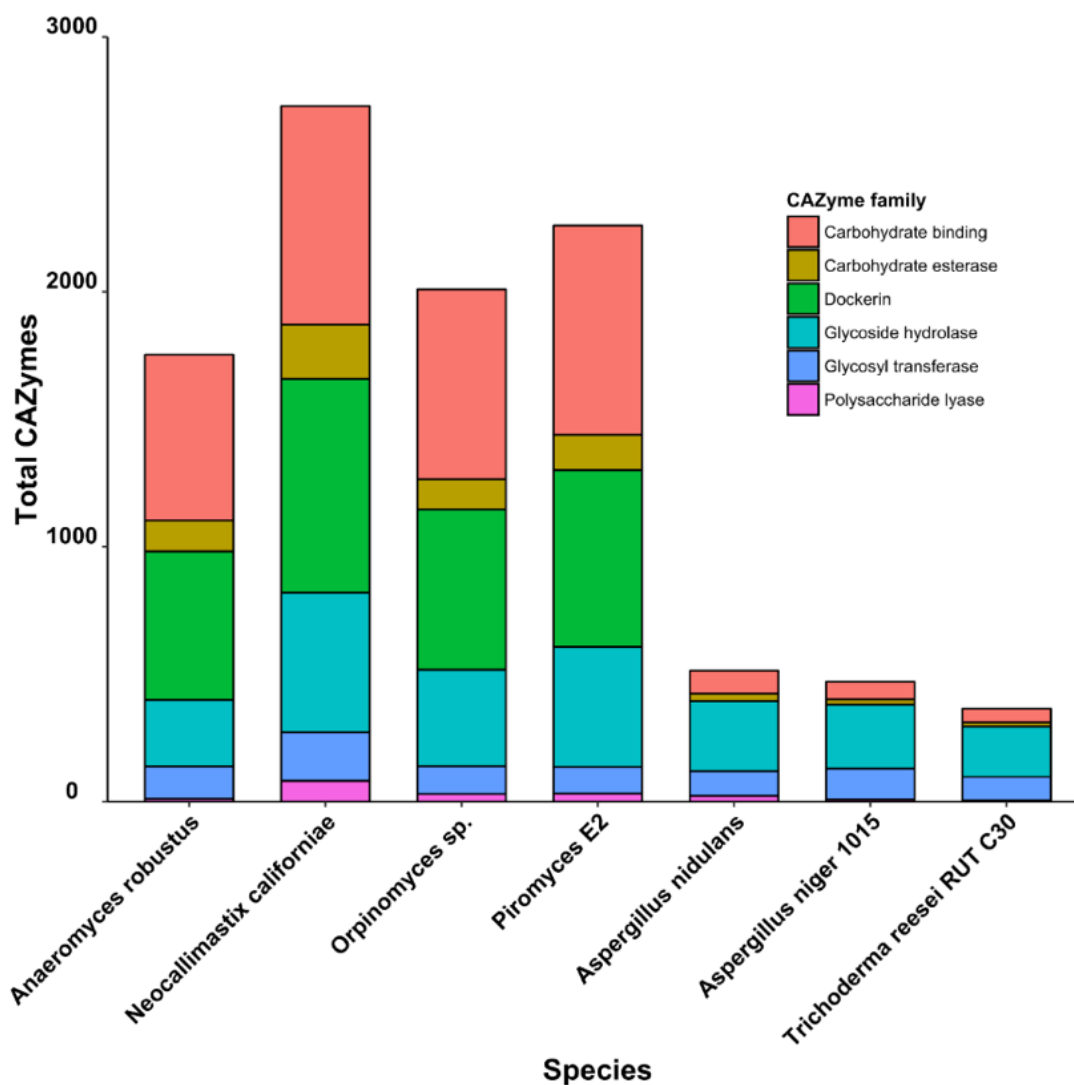
Figure 1.4: Anaerobic gut fungi harbor significantly more biomass degrading enzymes than the organisms that are currently used for industrial enzyme preparations [47].

The cellulosomes of gut fungi contain an array of proteins necessary for biomass hydrolysis (Figure 1.5). The component proteins have a range of different activities and functions, including carbohydrate binding proteins or domains (CBPs), which tether the fungal cells to the plant biomass [64]. Similarly, cellulosomes contain multiple unique CAZymes (e.g. glycosyl transferase,

polysaccharide lyase, glycoside hydrolase), which allow for concerted enzyme hydrolysis to improve degradation efficiencies. Additionally, scaffoldin domain proteins and cohesion-dockerin proteins (Figure 1.5) foster the necessary interactions that allow these superstructures to form and attach to the fungal cells [64]. In summary, anaerobic fungi encode significant numbers of biomass degrading enzymes, and exploit a different evolutionary approach to tackle lignocellulose complexity. By developing an enzyme cocktail from cellulosome forming fungi, it may be possible to use reduced enzyme loadings which would decrease production costs. In conclusion, to develop anaerobic fungi for diverse lignocellulose hydrolysis, there is a critical need to isolate, thoroughly characterize, and sequence anaerobic fungi to understand how these organisms may be engineered to more efficiently degrade plant biomass for bioenergy production.
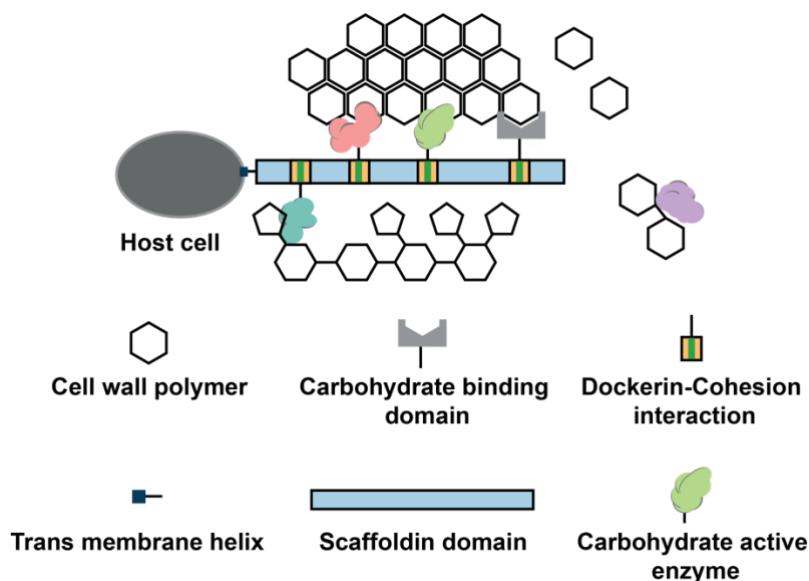
Figure 1.5: The structure of the cellulosomes of anaerobic fungi. Adapted from [64].

1.6 Thesis objectives and approaches

The goals of this work are to isolate and characterize novel species of anaerobic gut fungi (Table 1.3). Our approach is to identify the ways in which lignocellulolytic enzyme expression and hydrolysis adapt to mixed substrate compositions given the need to readily breakdown diverse plant biomass. We investigate fungal growth on a range of feedstocks, and show preliminary work

for the role of epigenetics in modulating the expression of biomass degrading enzymes. As there is no information on epigenetics in the fungi of Neocallimastigomycota, I hope to characterize the extent to which epigenetics is involved in biomass degradation. The aim of this work is to provide a foundation on which fungal genome engineering tools can be used to increase the efficiency and production of these biomass degrading enzymes for enhanced biomass breakdown.

Table 1.3: The isolated species of anaerobic gut fungi that are investigated in this work.

Isolated species of anaerobic fungi	Host organism	Abbreviated name	Date isolated	Place isolated
Piromyces sp. UH3-*1*	Horse	UH3-1	July 2016	Attica, Indiana
Neocallimastix sp. Gf-ma	Giraffe	Gf-ma	February 2017	Indianapolis, Indiana
Neocallimastix sp. WI3B	Wildebeest	WI3B	February 2017	Indianapolis, Indiana
Piromyces spp.	Rhinoceros	RS-3	February 2017	Indianapolis, Indiana

2. UNENGINEERED ANAEROBIC FUNGI DEGRADE DIVERSE UNTREATED LIGNOCELLULOSE WITH VARIABLE LIGNIN COMPOSITION

This chapter is from a paper by Casey A. Hooker, Ethan T. Hillman, Jonathan C. Overton, Adrian Ortiz-Velez, Makayla Schacht, Abigail Hunnicutt, Nathan S. Mosier, and Kevin V. Solomon. *Biotechnology for Biofuels* 11(1) (2018):293. [65].

2.1 Background

Lignocellulosic material is an inexpensive and abundant source of carbon that remains underexploited for biofuel production due to its complex heteropolymeric structure that hinders release of fermentable sugars by lignocellulolytic enzymes [66]. Available plant biomass for bioenergy is greatly dependent on geographic location and climate variability, leading to large differences in the types and compositions of the potential substrates [13]. More importantly, the biomass composition strongly affects the performance of a given enzyme cocktail [33]. As a result, the enzyme cocktails that are used to hydrolyze these feedstocks are optimized for individual substrates, and are not suitable for more economically-viable feedstock streams whose composition fluctuate greatly with market availability [67]. As enzyme cost is a significant bottleneck to the development of economical biofuels, enzyme systems that display superior performance on diverse feedstocks would advance the economic feasibility of bioenergy [28, 34, 35].

Current lignocellulolytic enzymes systems are based on well-known fungi such as *Trichoderma reesei* and *Aspergillus spp.* due to their oversecretion of many glycoside hydrolyases (CAZymes), which are active on the glycosidic bonds of lignocellulosic materials [25]. However, these species do not naturally express all of the enzymes needed to fully hydrolyze the sugars contained in plant biomass [30]. For example β-glucosidases in *T. reesei,* an enzyme essential to release the free glucose, form less than 1% of all secreted CAZymes [25]. Thus, enzyme cocktails based on *T. reesei* must be supplemented with enzymes from other species for sufficient activity [68]. The need for cocktail supplementation with enzymes from various species greatly increases enzyme production costs due to capital-intensive parallel enzyme production processes [28, 69]. Therefore, a single species enzyme platform would simplify enzyme production and reduce cost.

Degradation of untreated biomass is common in many underexplored environments that may harbor efficient microbial enzymes for biofuels. One example is the rumen and hindgut of large herbivores where grasses, shrubs, and other untreated fiber-rich plant biomass are processed daily by a consortium of microbes including early-divergent Neocallimastigomycota (anaerobic fungi) [70]. While anaerobic fungi are known to harness powerful biomass-degrading enzymes, the ability of these enzymes to hydrolyze diverse plant biomass remain poorly characterized [53]. To date, only five specimens in this phylum have been sequenced and studied in any detail [47]. The fungi of Neocallimastigomycota thrive under mild conditions (pH ≈ 7, 39 ºC), and possess large arrays of CAZymes that efficiently degrade untreated plant biomass [53, 64]. However, there is little data on the extent of the cellulosic and xylanosic degradation by these enzymes across a range of lignin compositions.

Given the potential for anaerobic fungi to reduce enzyme production costs, we sought to characterize their enzymatic performance as a function of substrate composition. Here, we report the isolation and taxonomic placement of a novel species of anaerobic gut fungi (*Piromyces* sp. UH3-1) in the Neocallimastigaceae family. We characterize the ability of *Piromyces* sp. UH3-1 to degrade and grow on an array of untreated substrates (e.g. corn stover, switchgrass, orange peel, and sorghum) under mild conditions. Additionally, we measure the free sugars released from untreated poplar across a range of lignin compositions to estimate fungal enzyme performance with feedstock composition. This work suggests that anaerobic fungal enzymes are robust for hydrolysis of diverse untreated lignocellulose and are promising new candidates for lignocellulosic enzyme production.

2.2   Methods

2.2.1   Isolating a novel species of anaerobic gut fungi

We suspended fresh donkey feces in Hungate tubes containing sterile anaerobic medium C supplemented with 15% clarified rumen fluid (150 ml: Bar Diamond Inc., Parma, ID, USA) under 100% CO_2 headspace [71]. Suspensions of donkey feces were serially diluted 1000-fold and used as a 10% inoculum in Hungate tubes containing 9 ml anaerobic medium C, supplemented with switchgrass as a carbon source (1% w/v) and chloramphenicol (25 µg/ml; Fisher Scientific, Waltham, MA, USA). After inoculation, the cultures were incubated at 39 ˚C for 72-96 hours.

To obtain axenic cultures, we inoculated roll tubes with liquid fungal culture and propagated individual colonies. Roll tubes were prepared by adding agar (2% w/v), glucose (0.45% w/v), and chloramphenicol (25 µg/ml) to anaerobic medium C under 100% CO_2 headspace [71]. We melted solid sterile media at 98 °C in a water bath and cooled the media to ~45-50 °C prior to the addition of chloramphenicol and 1 ml of inoculum from a liquid fungal culture in mid-exponential phase. Upon inoculation, the tubes were transferred to a benchtop and immediately rolled horizontally creating a uniform agar-inoculum completely coating the walls. The tubes were incubated at 39 °C until colonies were visible, typically between three and five days. Following incubation, we extracted individual colonies from the agar with a sterile needle while under CO_2 headspace and transferred them to new Hungate tubes containing 9 ml anaerobic medium C, switchgrass, and antibiotics (chloramphenicol [25 µg/ml in 40% ethanol], streptomycin [40 µg/ml], penicillin [50 µg/ml], and kanamycin [25 µg/ml]). After 72-96 hours, we used these cultures to inoculate new roll tubes. Colonies were passaged three times to obtain axenic cultures.

2.2.2   Substrate preparation

Lignocellulosic substrates were dried by placing them in a Fischer Scientific Isotemp convection oven at 45 °C until they reached approximately 10% moisture. Similarly, we collected the food waste (i.e. orange peel), washed it with deionized water, and dried it to approximately 10% moisture. We milled the dry substrates to 20 mesh (~ 0.85 mm) in a rotary mill. Milled substrates were loaded at 1% w/v prior to the addition of medium C [71]. For all soluble carbon sources, substrates were dissolved in anaerobic medium C at 0.5% w/v prior to being aliquoted into individual Hungate tubes under 100% CO_2 and autoclaved. Non-lignocellulosic substrates included arabinoxylan from beechwood (Megazyme, Bray, Ireland), xylan from beechwood (Crescent Chemical, Islandia, NY, USA), glucose, arabinose, xylose, cellobiose, filter paper, carboxy-methyl cellulose (Fisher Scientific, Waltham, MA, USA), Sigmacell Type 50, and Avicel pH 101 (Sigma Aldrich St. Louis, MO, USA). Genetically modified lines of poplar containing varying molar ratios of syringyl and guiacyl lignin were used to assess the response of *Piromyces* sp. UH3-1 to lignin composition [72, 73]. Poplar at approximately 10% moisture was milled to 40 mesh (~ 0.5 mm), and tubes were loaded with 1% w/v substrate. We tested eight different lines of debarked poplar. Two different lines of wild type poplar were used in this experiment; NM6, which is a global standard, and INRA 717 from which all of the modified lines were constructed [72, 74].

While autoclaved, all biomass in this study is effectively untreated; empirical calculations of the extent of pretreatment or severity factor are 4 orders of magnitude smaller than mild forms of pretreatment (Log R_0 – 2.10) [75]. Similarly, preliminary studies did not demonstrate significant increases in fungal growth rate or total fungal biomass accumulation when unautoclaved corn stover is used as the substrate (Appendix B: Figure B.11). The autoclaved biomass has not been washed to remove any potential fermentation inhibitors which hinder enzyme activity [76, 77].

2.2.3 Microscopy

All images of *Piromyces* sp. UH3-1 were collected via confocal microscopy (Nikon Eclipse Ti Microscope and A1-multiphoton imaging system). Mature fungal cultures containing lignocellulosic material were immobilized in 10% polyacrylamide prior to imaging with 4',6-diamidino-2-phenylindole (DAPI) (Thermofisher, Waltham, MA, USA). Zoospore images were collected using 3-day corn stover cultures, which were placed in Eppendorf tubes and fixed with formaldehyde to a 4% final concentration.

2.2.4 Species classification

An axenic stock culture (described in Isolating a novel species of anaerobic fungi) was used to inoculate 50 ml serum bottles containing medium C with glucose 0.45% w/v, and chloramphenicol (25 μg/ml in 40% ethanol) [71]. These serum bottles incubated at 39 ℃ for 3-4 days upon which the gDNA was harvested for species classification. Fungal genomic DNA was isolated with the MoBio PowerFecal kit (Carlsbad, CA, USA), yielding sufficient quality genomic DNA (260/280: 1.9 & 260/230: 1.5) at approximately 2 μg DNA per 50 ml culture. PCR (Phusion DNA polymerase, Thermoscientific, Waltham, MA, USA) was used to amplify the Internal Transcribed Spacer 1 (ITS1) and ITS2 regions of the isolated genomic DNA via JB206/205 primers (5' GGAAGTAAAAGTCGTAACAAGG 3' and 5' TCCTCCGCTTATTAATATGC 3') yielding an expected amplicon of approximately 700-750 base pairs [78]. We also amplified the D1/D2 portion of the 28S rRNA large subunit (LSU) gene with the NL1/NL4 primers (5' GCATATCAATAAGCGGAGGAAAAG 3' and 5' GGTCCGTGTTTCAAGACGG 3') [79]. DNA was amplified with the following PCR settings for 30 cycles: annealing at 56 ℃, elongating for 60 seconds at 72 ℃ and melting at 98 ℃. All of the same conditions were used for the LSU PCR reaction except the annealing temperature was changed to 67 ℃ [79]. DNA amplification was

checked on an agarose gel and imaged with a c600 Azure Biosystems imager. We concentrated these PCR products with the Zymogen DNA Clean & Concentrator (Zymo Research, Irvine, CA, USA) kit prior to sequence submission at GENEWIZ (South Plainfield, NJ, USA). We assembled the forward and reverse sequence reads of the ITS1 and ITS2 region into a single contig by trimming the ends of reads with poor base calls (>3 Ns in a 20 base window) and assembling reads with 85% overlap over at least 20 bps with the contig assembly feature in GeneStudio bioinformatics package (ver. 2.2.0.0, GeneStudio, Inc., Suwanee, GA, USA). ITS sequences were also validated by cloning PCR products into the pGEM-T Easy cloning vector (Promega, Madison, WI, USA) following the manufacturer's instructions and sequencing 3 resulting colonies. Phylogenetic reconstruction was performed using MEGA7 (v 7.0.14). Due to the lack of homogeneity in coverage across the ITS1 and ITS2 sequences in gut fungi, only ITS1 and LSU sequences were used [79]. ITS1 and LSU sequences were analyzed with the maximum likelihood method using a Tamura Nei nucleotide substitution model with 1000 bootstrap replications to estimate the confidence in node clustering.

2.2.5 Growth curve analyses for characterizing the substrate range of *Piromyces* sp. UH3-1

Fungal growth was tracked according to the method introduced by Thedorou et al. [80]. Briefly, Hungate tubes containing anaerobic medium C and untreated substrate were autoclaved prior to assessing growth (Appendix A: Tables A.1-A.5, Appendix B: Figures B.4-B.6) [71]. Every substrate was tested at least in triplicate for growth. Additionally, duplicate uninoculated tubes were used as negative controls for each substrate. Specific growth rates were determined by performing a linear regression of a semi-log plot of accumulated pressure (in psig) versus time (in hours). The Microsoft Excel LINEST function was used for each plot to calculate the slope and exponential phase. The data points in the exponential phase that were linearly increasing and had an R^2 of approximately 0.90 or higher (typically between 48 and 120 hours) were used to calculate the specific growth rates on each substrate. We prepared fresh media as described above. Lignocellulosic and insoluble substrates were loaded at 1% w/v while soluble substrates were loaded at 0.5% w/v to keep the total mass of fermentable sugars relatively constant. Tubes were inoculated in a random order to prevent systematic bias in inoculum quality. Pressure accumulation was measured with a pressure transducer (APG, Logan, Utah, USA), every eight hours for seven days. The growth of *Piromyces* sp. UH3-1 on wild type and genetically modified lines of poplar

was tested to evaluate the effect of lignin composition on fungal growth (Appendix A: Tables A.2-A.5) [72, 74]. For all analyses, individual growth rates and total accumulated pressures were calculated. For data normalization to glucose (Figure 2.3D), the average accumulated pressure (in psig) across culture (biological) replicates at 168 hours for each substrate was divided by the average accumulated pressure of glucose at 168 hours for all of the inoculated tubes. The error for these measurements was propagated accordingly. For data normalization to wild type poplar (Figure 2.5 A-B), the same procedures were followed for both growth rate and accumulated pressure.

2.2.6 Isolation of the carbohydrate active enzymes (CAZymes)

We used a pull-down purification protocol similar to the one by Solomon et al. to isolate and concentrate fungal CAZymes [53]. This procedure exploits the cellulose-binding domains of CAZymes to isolate lignocellulose degrading enzymes [53]. Cultures were centrifuged at 12,800g and the supernatant was transferred to a tube containing approximately 0.4% (w/v) Sigmacell type 50. These tubes were incubated overnight at 4 °C with gentle agitation. Tubes were then centrifuged at 12,800g and the supernatant was discarded. 0.1M pH 7.0 Tris-NaCl buffer was added to the Sigmacell to elute the cellulose-binding enzymes. The elutions were then stored at 4 °C for further analysis. Protein concentrations were determined by the method introduced by Bradford (Fisher Scientific, Waltham, MA, USA) [81].

2.2.7 SDS PAGE and zymography analyses for detailed enzyme characterization

Cellulose-binding proteins were separated and visualized on 10% acrylamide gels run for 70 minutes at 110V. Gels were then stained with Sypro Ruby Protein Stain (Fisher Scientific, Waltham, MA, USA). These proteins were also tested for activity via zymography with 0.2% w/v carboxy methyl cellulose (CMC) or 0.4 % w/v pectin added to the resolving portion of a 10% acrylamide gel under non-denaturing conditions. The SDS was removed from the gel with slight modification to the procedure of Tseng et al 2002 [82]. The gels were rinsed with ddH$_2$O and placed in 0.1M pH 7.0 Tris-NaCl (TN) buffer containing 25% (w/v) isopropanol (TNI) buffer. Zymogram gels incubated for 30 minutes at 4º C in TNI buffer with gentle agitation. The TNI buffer was then removed, and the gel was rinsed two more times with fresh TNI buffer. The zymograms were then washed with 0.1M pH 7.0 TN-buffer prior to incubating at 39 ºC for

substrate hydrolysis. CMC zymograms were incubated for one hour while pectin zymograms were incubated for 24 hours. Zymograms were then stained in 0.1% w/v Congo red stain (Fisher Scientific, Waltham, MA, USA), for 30 minutes, and de-stained with 1M NaCl (Fisher Scientific, Waltham, MA, USA) until the hydrolysis zones appeared relative to the red background. We fixed the zymograms with 0.1M acetic acid (Fisher Scientific, Waltham, MA, USA) prior to imaging.

2.2.8   Sugar reducing assay for xylanase activity

Piromyces sp. UH3-1 xylanase activity was measured after harvesting the cellulose binding proteins as discussed above. Briefly, we followed the 96 µl microplate procedure introduced by Xiao et al [83]. However, we used 0.05M potassium phosphate buffer (pH 7.0) in place of citrate, and a 2% solution of xylan from beechwood (Crescent Chemical, Islandia, NY, USA) as the substrate. Substrate hydrolysis proceeded for six hours at 50º C before the generated reducing sugars were measured at 540 nm on a Synergy Neo plate reader (Biotek, Winooski, VT, USA). All samples were measured in triplicate and normalized by total protein. To determine the extent of non-enzymatic xylan degradation, enzyme-free and protein (Bovine Serum Albumin [BSA], Fisher Scientific, Waltham, MA, USA) controls were tested.

2.2.9   Analyzing the composition of lignocellulosic material after fungal growth

To test the effect of syringyl lignin composition on sugar consumption by *Piromyces* sp. UH3-1, we grew the isolate in 100 ml serum bottles with 50 ml working volume and 1.4% (w/v) solids loading to generate sufficient spent biomass for analysis. Three different poplar constructs were used: 0998-45 (5% S-lignin), wild type INRA 717 (64% S-lignin), and F5H-64 (98% S-lignin) [72, 74]. After seven days, the spent lignocellulosic biomass and associated fungal residues were separated from the fermentation media by centrifuging at 5,000 RPM for 5 minutes. After centrifugation, the liquid phase was decanted and the solids were dried for five days at 45 ˚C. The sugar composition of the spent biomass was determined according to standard methods (Appendix A, Table A.6) [84-86]. Carbohydrates were determined using HPLC analysis (Waters 1525 Pump, Waters Corporation, Milford, MA, USA) equipped with an Aminex™ HPX-87H column (Bio-Rad, Hercules, CA, USA) maintained at 65 °C. The mobile phase was 5 mM aqueous H_2SO_4 at a flow rate of 0.6 mL/min. 50 µL of sample was injected, analyzed using a Waters 2414 Refractive Index detector (Waters Corporation, Milford, MA, USA) and quantified using Empower Pro

Software (Waters Corporation, Milford, MA, USA). The differences in glucan and xylan composition between the raw and spent biomass were calculated, and one-way ANOVA analyses were performed to evaluate the differences in composition.

2.3 Results

2.3.1 Isolation of a biomass degrading anaerobic gut fungus from a donkey

To identify more robust and efficient CAZymes and microbial systems that may be used for bioenergy applications, we isolated a previously uncharacterized microbe from the fecal samples of a donkey. Light microscopy revealed the presence of non-planktonic microorganisms that grew invasively into the plant substrates, reminiscent of a mature fungal sporangium (Figure 2.1A), after 3-4 days of growth with a simultaneous increase in headspace pressure. This isolated organism was cultured to axenic purity by repeated passage through roll tubes containing multiple antibiotics (see Methods, Figure 2.1B). Further microscopic analysis revealed that this organism produces zoospores with a single flagellum (~30 μm long) (Figure 2.1C), another key characteristic of the genus *Piromyces* of the fungal phylum Neocallimastigomycota. Additionally, this isolate exhibits endogenous zoosporangial development, where the zoosporangium retains its nuclei. The slow growth, zoospore presence, and well-differentiated stages of a life-cycle (Figure 2.1C-E) suggested a fungal specimen. DAPI staining of the nucleic acid in the developing fungal sporangium revealed that this isolate was monocentric (has nuclei only within the zoosporangium) (Figure 2.1E), which is also consistent with the morphology of the fungal genus *Piromyces*

Taxonomic classification of our novel fungal isolate was confirmed via phylogenetic analysis [79, 87]. Amplification of the 16s rRNA genes failed, while amplification of the ITS1, ITS2, and the 28s rRNA large ribosomal subunit (LSU) were all successful (Appendix B: Figure B.1) [88]. Therefore, the isolate was definitively fungal in origin, rather than bacterial or archaeal, which agrees with our morphological assessment. We aligned these amplicons against 51 Genbank deposited anaerobic fungal sequences (division Neocallimastigomycota) and confirmed that our isolate formed a distinct branch within the fungal *Piromyces* genus (Figure 2.2A-B) [79, 89]. The monocentric thallus and uniflagellated zoospore, both characteristic of *Piromyces* fungi, further support this placement (Figure 2.1E-F) [49].
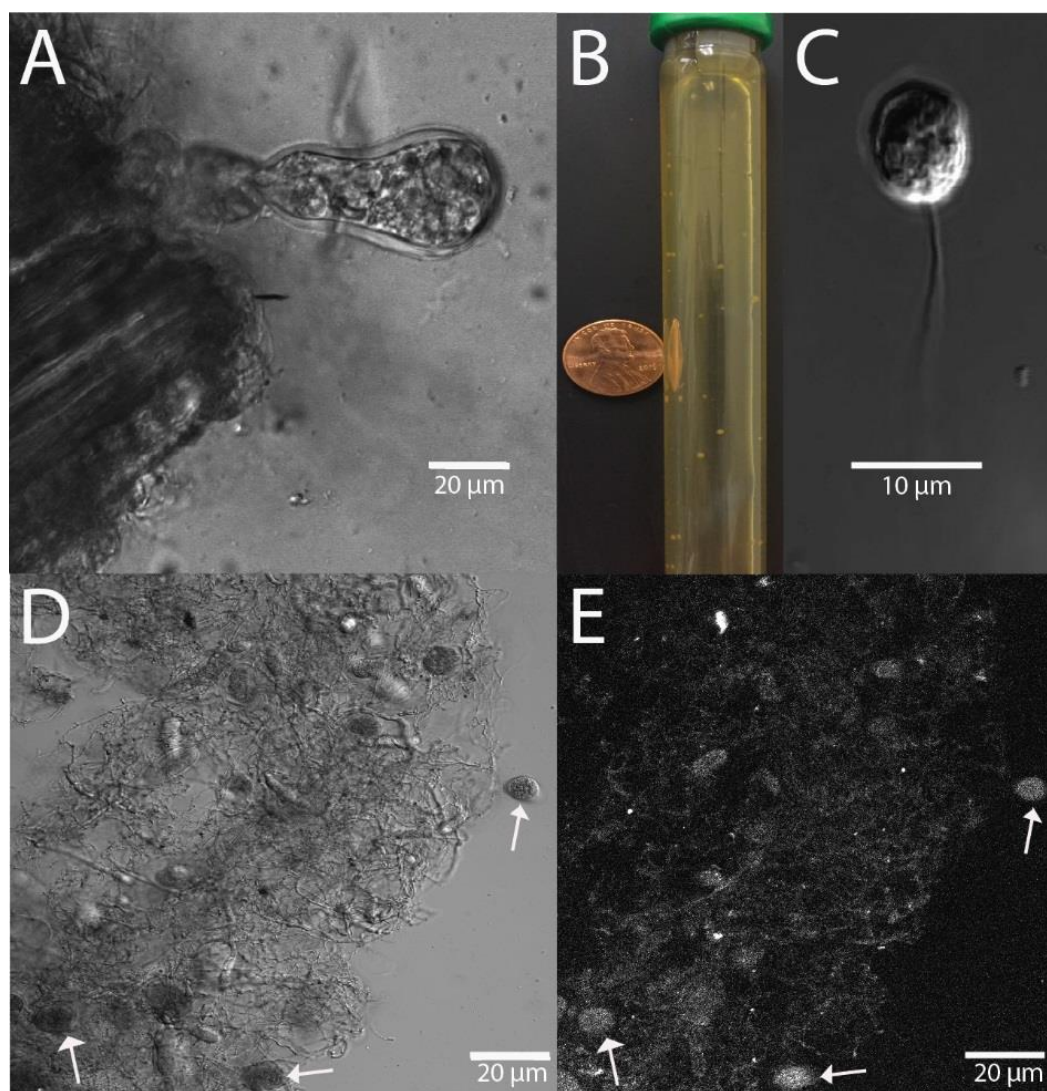
Figure 2.1: The host and life cycle of *Piromyces* sp. UH3-1: A) Individual mature sporangia on corn stover (left) displaying ovoid structure. B) Roll tube used to isolate individual axenic cultures of *Piromyces* sp. UH3-1. C) Uniflagellated zoospore of *Piromyces* sp. UH3-1 imaged after zoospore death D) Multiple sporangia, demonstrating the predominantly spherical to ovoid structure; arrows indicate individual sporangia in rhizomycelial network. E) DAPI stain indicating the monocentric nature as zoosporatic nuclei are contained with the sporangia [65].

This organism represents a novel cultured species as it has less than 90% BLAST similarity to known cultured species of anaerobic fungi. Therefore, we classify this organism as the species *Piromyces* sp. UH3-1 (NCBI Taxon ID: KY494854, JRMC: SF:012426, Index Fungorum: IF554555) and (Appendix B: Formal Species Description), in honor of the state in which it was isolated (Indiana, United States).
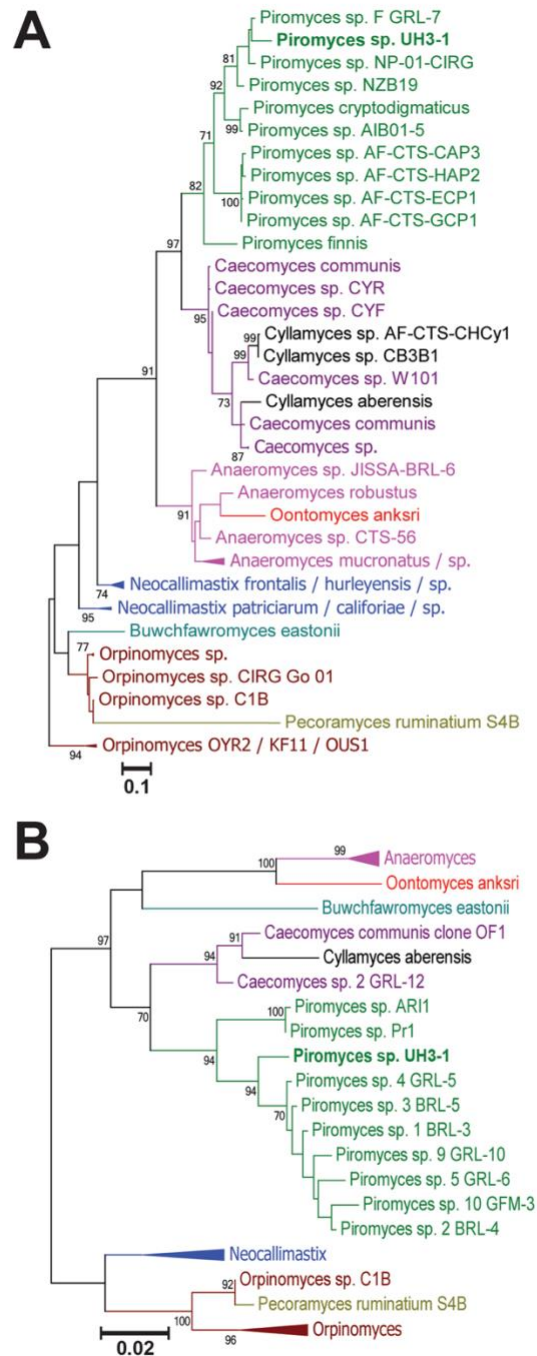
Figure 2.2: Phylogenetic analyses place our isolate within the genus *Piromyces*: A) Collapsed ITS1 phylogenetic tree and B) Collapsed LSU phylogenetic tree. Fully expanded phylogenetic trees displaying the Genbank accession numbers are in Appendix B: Figures B.2-B.3. Significant bootstrap values from 1000 iterations are indicated to the left of each branch [65].

2.3.2 Anaerobic fungi degrade complex substrates with efficiencies comparable to glucose

Untreated lignocellulosic substrates are rich in sugars that can sustain fungal growth; however, the degradation rate of these substrates into free sugars is frequently limiting for growth. Thus, to estimate hydrolysis efficiency, we assessed the ability of *Piromyces* sp. UH3-1 to grow on agricultural residues, bioenergy crops, food wastes, and forestry products that had not undergone pretreatment (Appendix A: Tables A.1-A.5). Anaerobic fungi secrete an array of CAZymes that break down diverse lignocellulosic material into fermentable sugars that the fungus metabolize to CO_2 and H_2, among other fermentation products such as lactate, formate, acetic acid, and ethanol[90]. Anaerobic gut fungi grow invasively into plant substrates forming plugs that trap the fermentation gasses leading to more buoyant floating cultures (Figure 2.3A). However, when grown on soluble substrates, the fungi grow into themselves to form a mat of biomass (Figure 2.3A). Gas accumulation is proportional to fungal biomass production and may be used as a convenient indicator of growth (Figure 2.3B) [80]. Both pressure accumulation, and visual analysis were used to assess the ability of *Piromyces* sp. UH3-1 to grow on these feedstocks. While the growth rates for these substrates varied significantly, the total pressure accumulations were comparable for lignocellulosic substrates (Figure 2.3C-D). Therefore, these results suggest that this organism secretes an array of CAZymes that liberate sufficient sugars, regardless of feedstock composition, to sustain fungal growth into stationary phase.
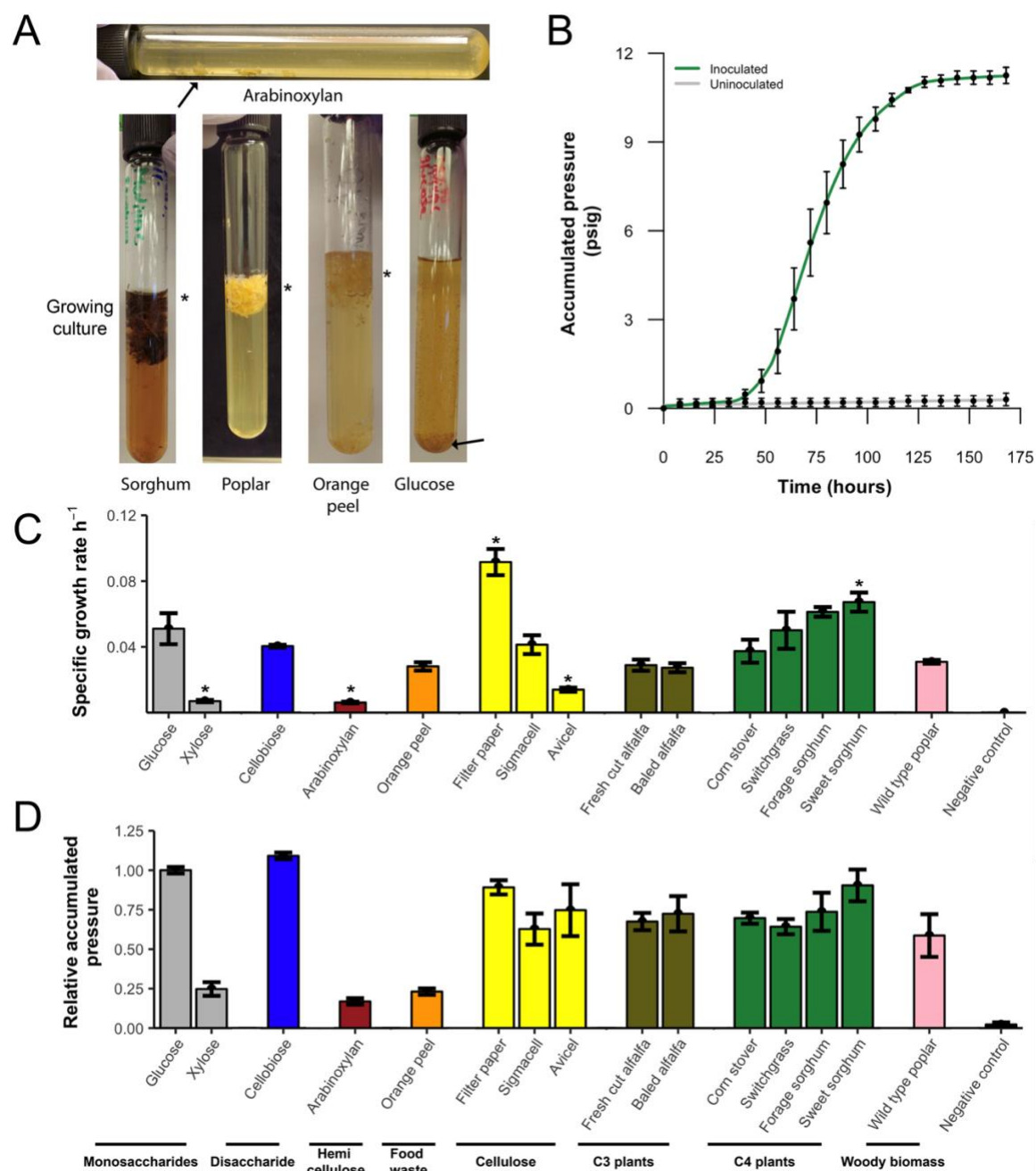
Figure 2.3: *Piromyces* sp. UH3-1 grows on diverse feedstocks: A) Growth of *Piromyces* sp. UH3-1 on soluble substrates leads to colony formation on the walls of the tubes (arrows indicating colony formation). Fungal cultures growing on lignocellulosic substrates float up during fermentation. B) A representative growth curve of *Piromyces* sp. UH3-1 on corn stover. C-D) *Piromyces* sp. UH3-1 degrade and proliferate on a wide array of untreated agricultural wastes, bioenergy feedstocks, and forestry wastes. All accumulated pressures are normalized to glucose. Asterisks denote statistically significant differences in specific growth rate relative to glucose ($p < 0.05$, unpaired t-test) [65].

To determine whether biomass hydrolysis was efficient or limiting for growth, we first established a baseline for growth on simple sugars. Glucose led to robust growth, (Figure 2.3, Appendix B: Figure B.4), and was used as a baseline to which all other substrates were compared. Similarly, the disaccharide cellobiose led to strong, robust fungal growth, suggesting that anaerobic fungi readily produce β-glucosidases that can cleave cellobiose to glucose at a rate in excess of glucose uptake and metabolism (Figure 2.3, Appendix B: B.4). In contrast, fungal growth on hemicellulosic components such as xylose and arabinose (Figure 2.3, Appendix B: Figure B.4, B.7-B.8) led to inconsistent pressure accumulation and a significantly reduced growth rate relative to glucose ($p = 0.0147$, unpaired t-test). Nonetheless, accumulation of fungal biomass on xylose was consistently observed (Appendix B: Figure B.8). Thus, xylose transport and incorporation into central metabolism likely occurs more slowly than six carbon sugars and may be limiting for growth. Taken together, these results suggest that this fungal isolate grows primarily on hexose sugars, and has robust β-glucosidase activity that is not a bottleneck for biomass hydrolysis, unlike *T. reesei* [91]. While fungal growth on hemicellulose components is poor, it must still remove hemicellulose and other carbohydrate polymers to access the glucose-rich cellulosic portions of lignocellulose. Arabinoxylan, a form of hemicellulose, contains fermentable arabinose and xylose sugars, and is highly abundant in the cell walls of cereals and grasses used as bioenergy crops [92]. Similarly, pectin is a complex and variable component in the middle lamella between the plant cell walls. As this surrounds the energy rich cellulosic and hemicellulosic polymers, pectin removal or deconstruction is advantageous for efficient lignocellulose hydrolysis [93-95]. The growth of *Piromyces* sp. UH3-1 on wheat arabinoxylan and pectin rich feedstocks such as orange peel, while consistent, was unlike typical microbial growth and non-sigmoidal in nature (Figure 2.3, Appendix B: Figure B.4). However, when the sugar monomers rhamnose and galacturonic acid were tested, pressure accumulation was irregular and suggested that these substrates could not be metabolized fast enough to support fungal growth (data not shown). Thus, the degradation products of pectin, and to a lesser extent arabinoxylan are unlikely to sustain robust growth. Given the poor growth on these polymeric substrates, we directly analyzed their hydrolysis by collecting the fungal secretome and testing for CAZyme activity.

By isolating the fungal enzymes we were able to test their activity via zymography, which exploits the ability of some stains to preferentially bind to polysaccharides (Figure 2.4) [96].
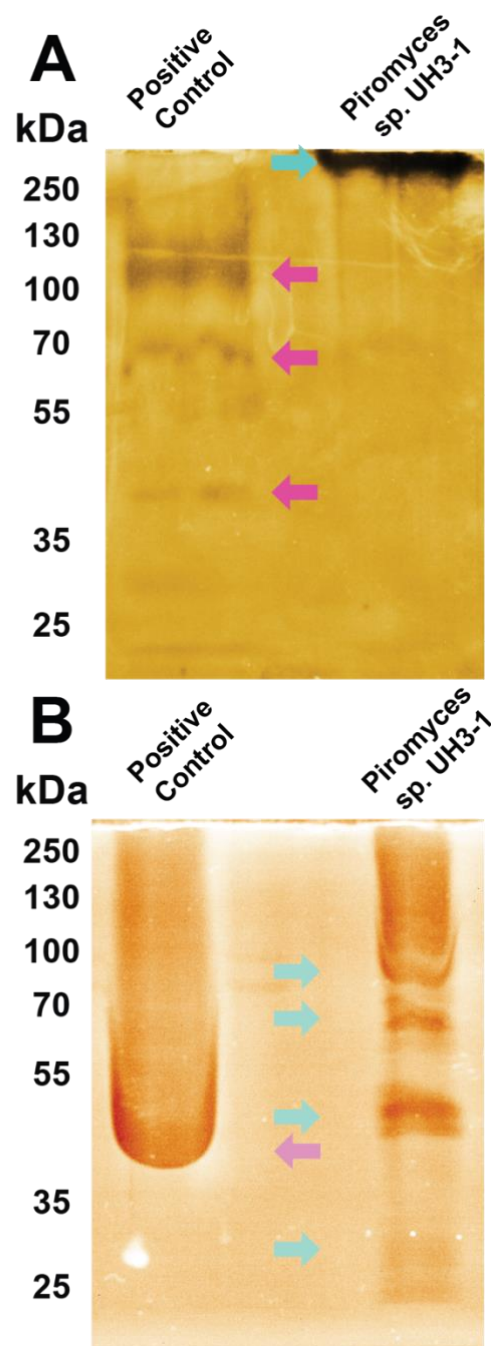
Figure 2.4: *Piromyces* sp. UH3-1 secretes diverse CAZymes for degrading the polymers of lignocellulose: A) A pectin zymogram shows strong pectinolytic activity for *Piromyces* sp. UH3-1 at the top of the gel (teal arrow), while *Aspergillus* shows multiple bands having pectinolytic activity (pink arrows). B) A carboxy methyl cellulose zymogram shows distinct cellulolytic activity for multiple proteins of *Piromyces* sp. UH3-1 (teal arrows), while *Aspergillus* (Viscozyme, positive control) shows high cellulolytic activity (Pink arrow). Controls and experimental samples were loaded with the same total protein mass as measured by a Bradford assay [65].

Differential staining around individual protein bands results from the consumption of substrate and is positive for hydrolytic activity. Pectin zymograms show a high molecular weight hydrolysis zone indicating that this fungal isolate can degrade this complex polymer (Figure 2.4). Similarly, reducing sugar assays reveal strong xylanolytic activity from anaerobic fungal secreted proteins (Appendix B: Figure B.9). Thus, while *Piromyces* sp. UH3-1 is unable to efficiently metabolize these substrates, it still expresses an array of CAZymes that break down the pectin and hemicellulose components of lignocellulose under mild conditions.

Readily degrading cellulose is critical to efficiently producing energy from renewable plant biomass [38, 97]. Given the variability in cellulose structure between plant sources and preprocessing before enzymatic hydrolysis occurs (e.g. degree of crystallinity, porosity, and specific surface area), we evaluated the efficiency of cellulose hydrolysis by testing three different substrates, which all yielded robust fungal growth (Figure 2.3, Appendix B: Figure B.4) [3, 98, 99]. Sigmacell from cotton linters and filter paper yielded growth rates that were equal to or in excess of growth on glucose suggesting that cellulase activity is not limiting for growth on lower crystallinity substrates. In contrast, growth on Avicel, a highly crystalline cellulose produced by acid hydrolysis of wood pulp, was reduced by 65% (p = 0.0268, unpaired t test), likely due to inhibition from the high crystallinity and reduced surface area caused by settling and packing of the substrate in these stationary fermentations [98]. Counterintuitively, growth on filter paper was faster than on glucose (p = 0.0023, unpaired t test). Despite these differences in growth rate, the total accumulated pressures were comparable, suggesting similar levels of carbon use, and thus sugar release, by the fungus independent of substrate crystallinity (Figure 2.3D). We sought to further characterize these cellulases by testing their activity through zymography (Figure 2.4B). Through this analysis, we identified multiple cellulose-binding proteins having cellulolytic activity. Taken together, these results suggest that this fungal isolate efficiently degrades cellulose by expressing multiple cellulases that have high activity in excess of glucose uptake and metabolism.

Piromyces sp. UH3-1 robustly grew on untreated lignocellulosic feedstocks, regardless of composition or photosynthetic type (Figure 2.3). Photosynthetic type (C3 or C4) leads to significant differences in cell wall structure, and thus the CAZymes needed to degrade the lignocellulose [100]. For C3 plants, we tested untreated alfalfa (*Medicago sativa)*, which resulted in strong fungal growth (Figure 2.3, Appendix B: Figure B.4). Commonly available C4 feedstocks for biofuel production such as corn stover (*Zea mays*), switchgrass (*Panicum virgatum*), and

sorghum (*Sorghum bicolor*) were consistently degraded by *Piromyces* sp. UH3-1 (Figure 2.3B-D, Appendix B: Figure B.5). Several varieties of sorghum, with differing cell wall compositions, were tested as they thrive in different climates and are planted in specific regions, unlike the other tested C4 feedstocks [16, 101, 102]. Notably, sweet sorghum was the only lignocellulosic substrate that yielded a significantly higher growth rate when compared to glucose (p = 0.0212, unpaired t test), possibly due to the excess free sugars common in sweet sorghum [16]. Thus, these results suggest that cell wall composition of untreated lignocellulose does not significantly reduce fungal growth rate, implying that the CAZymes of *Piromyces* sp. UH3-1 efficiently degrade these substrates.

2.3.3   Anaerobic fungal hydrolytic enzymes are robust to lignin composition

Woody biomass such as poplar has been proposed as a feedstock for second generation biofuel production as it is a fast-growing tree species capable of thriving in diverse geographic locations, has high biomass yields, and high glucan content (>40%) relative to other commonly used feedstocks (Appendix A: Tables A.2-A.5) [38, 103-106]. Furthermore, poplar can be grown on land that is marginally productive for most agricultural crops [107]. However, the lignin in poplar that has not undergone pretreatment is known to strongly affect cellulase and hemicellulase activity [108]. Despite this, *Piromyces* sp. UH3-1 still showed strong growth on wild type poplar (Figure 2.3, Appendix B: Figures B.4-B.6). This result is consistent with published data as anaerobic fungi are known to degrade untreated woody biomass [109]. However, as lignin composition may change for diverse feedstocks, we tested fungal growth on transgenic lines of poplar containing varying ratios (5%-98%) of Syringyl (S)-lignin (Figure 2.5, Appendix A: Tables A.2-A.5, Appendix B: Figures B.5-B.6). [72, 74]. S-lignin content is known to reduce the growth of some fungi by as much as 80% [18]
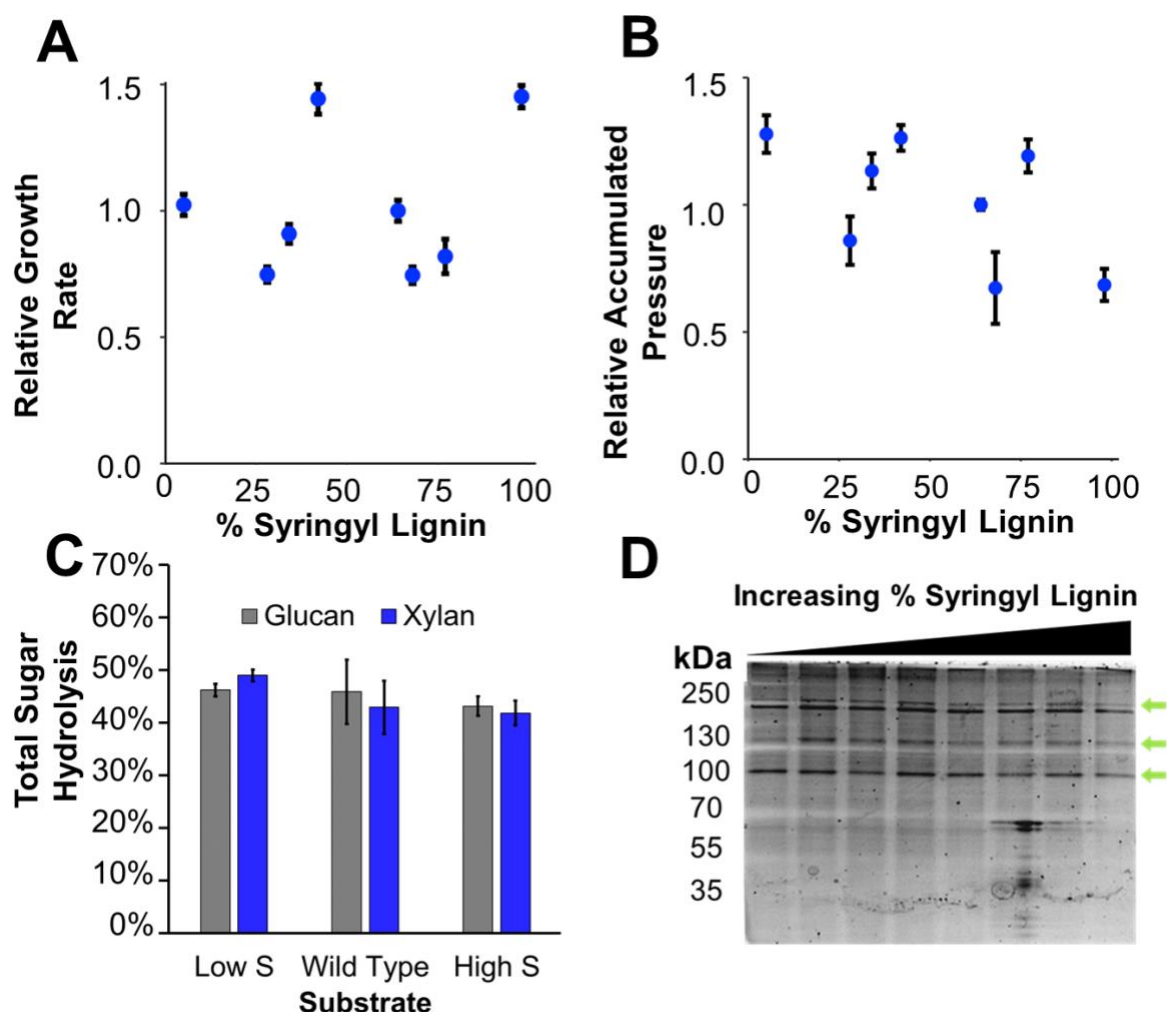
Figure 2.5: *Piromyces* sp. UH3-1 growth and sugar degradation is robust against lignin composition with optimal enzyme expression: A) Relative growth rates of *Piromyces* sp. UH3-1 on genetically modified lines of poplar relative to wild type INRA 717 (64% S-lignin), (p = 0.0317, R^2 = 0.1715). B) Relative fungal biomass accumulations of *Piromyces* sp. UH3-1 on genetically modified lines of poplar relative to wild type INRA 717 (p = 0.0011, R^2 = 0.2991). C) Minimum hydrolysis percentages on three of the lines of poplar [72, 74] D) The carbohydrate binding portion of the fungal secretome shows changes in response to S-lignin composition (green arrows) [65].

Our fungal isolate was insensitive to S-lignin content and degraded both S-lignin rich and poor substrates with high efficiency. Both growth rate and fungal biomass accumulation appeared to be independent of S-lignin content (Figure 2.5A-B). While an ANOVA analysis of this data yielded statistically significant trends for relative growth rate (p = 0.0317), and for relative fungal

biomass accumulation (p = 0.0011), this correlation was weak with R^2 values of 0.1715 and 0.2991, respectively. To further test these results, we repeated this experiment with another batch of poplar with S lignin molar percentages ranging from 20 to 93 percent (Appendix A: Tables A.4-A.5, Appendix B: Figure B.10). Further, this poplar was only milled to 20 mesh, while the above results were milled to 40 mesh (~0.420 mm). Notably, the total fungal biomass accumulation was insensitive to S lignin for these constructs despite having larger particle sizes (Appendix B: Figure B.10) R^2= 0.00817, p = 0.64723. To further evaluate the degradation of polymeric sugars in the presence of varying S-lignin compositions from the 2014 harvested poplar, we grew *Piromyces* sp. UH3-1 on three different poplar constructs and measured the hydrolysis of polymeric sugars to monomeric sugars (Figure 2.5C). As fungal rhizomycelia penetrate the plant material, it is currently not possible to distinguish fungal from plant biomass and accurately measure biomass loss, and thus total sugar consumption. However, an analysis of the glucan and xylan contents of spent and fresh poplar biomass, with the conservative assumption that total plant biomass is constant, suggests that *Piromyces* sp. UH3-1 metabolizes at least 43% of the glucose sugars, and 42% of the pentose sugars within 168 hours (Appendix A: Table A.6). These results are consistent with those reported for an isolate of *Neocallimastix*, a different genera of fungi also within *Neocallimastigomycota* [109]. Specifically this isolate released glucan and xylan at efficiencies of 47% and 34% respectively on untreated poplar after 11 days of growth [109]. Notably, glucan release was independent of S-lignin composition in the poplar constructs tested (wild type vs high S-lignin, p = 0.6499; wild type vs low S-lignin, p = 0.9951). There was also no significant difference in glucan release between low S-lignin and high S-lignin constructs (p = 0.5945). Similar trends were also observed for xylan release (wild type vs high S, p = 0.9105; wild-type vs low S-lignin, p = 0.1308; high S vs low S constructs, p = 0.0771). Taken together, these results suggest that these anaerobic fungal enzymes are robust against inhibitory syringyl lignin content and hydrolyze glucan and xylan in untreated lignocellulose with similar efficiency regardless of lignin composition. More importantly, while there are no known mechanisms by which anaerobic fungi can metabolize lignin constituents, our results suggest that fungal pathways may exist to recognize lignin composition to increase S lignin resistance. To analyze this, we collected the cellulose-binding portion of the fungal secretome after growth on modified poplar lines (Figure 2.5D). The relative concentration of several proteins changed non-linearly with S-lignin content suggesting a complex response to combat S-lignin recalcitrance. More importantly, while there are

no known mechanisms by which anaerobic fungi can metabolize lignin constituents, our results do suggest that fungal pathways exist to recognize lignin composition to regulate secretion of enzymes that perhaps trade hydrolytic activity for increased S-lignin resistance.

2.4 Discussion

Producing biofuels from lignocellulose that are competitive with current energy technologies requires more efficient use of existing biomass reserves in processes that incorporate multiple feedstocks of variable composition. Increasing the number of potential feedstocks will help to protect second generation platforms from changing production conditions that may result due to inconsistency in plant biomass yield, climate variability, and market volatility. One way to move toward this goal is by pretreating plant biomass, which has traditionally been used to overcome lignin inhibition. While pretreatment helps to mitigate these issues, new waste streams are introduced, toxic inhibitors are released that hinder the growth of the fermenting organism, and higher enzyme loadings are required [76, 97, 110, 111]. A more promising strategy is, thus, to identify enzyme platforms that readily degrade diverse untreated lignocellulose and are robust to variations in biomass composition. However, for this to be industrially economical, high fermentable sugar conversions are a necessity. Key challenges include product inhibition of the cellulases and the release of lignin among other contaminants that can inactivate the secreted enzymes. Despite these barriers unengineered anaerobic gut fungi such as *Piromyces* sp. UH3-1 show strong conversions on untreated plant biomass (Figure 2.5). When grown on milled (~0.5 mm), untreated corn stover, *Piromyces* sp. UH3-1 converts at least 58 and 28 percent of the available glucan and xylan, respectively (Appendix A, Table A.6). These values are comparable to current commercial enzyme cocktails which release 48 and 30 percent of the available glucan and xylan, respectively, on ball-milled (~100 micron) corn stover [112]. While hydrolytic rates were not measured as the secretome of *Piromyces* sp. UH3-1 is a crude preparation of lignocellulolytic enzymes and other unrelated proteins that are released over time, similar studies of anaerobic fungi suggest that our observed glucan conversion are not limiting and would improve with increased enzyme loading and or time [113]. In contrast, conversion with current cocktails saturate at these conversions unless supplemented with additional enzymes [114, 115]. Thus, the ability of *Piromyces* sp. UH3-1 to degrade these untreated feedstocks without additional supplementation at comparable efficiencies to commercial cocktails, supports further study of

these CAZymes, which may provide next generation solutions to critical issues with lignocellulose recalcitrance.

Engineering fungi for enhanced enzyme production has been a subject of considerable research [30, 116]. Both *Aspergillus* and *Trichoderma* are widely used to produce industrial enzyme cocktails, yet these organisms are not the strains that were originally isolated given their natural deficiencies. [117, 118]. For example, *Aspergillus* is known to express low amounts of endoglucanases, which is critical for efficiently degrading cellulose [26]. Additionally, the QM6A strain of *Trichoderma reesei* (previously *T. viride*) that was initially isolated went through multiple rounds of mutagenesis to obtain the hyper-producing, catabolite repression resistant strain Rut-C30 that is the basis for commercial enzyme production [117]. Similar to the original *Trichoderma* QM6A, fungi of *Neocallimastigomycota* are known to have catabolite repression that directly represses CAZyme expression [53, 117]. Despite the presence of catabolite repression, gut fungi still robustly degrade untreated lignocellulose. Manipulating anaerobic gut fungi, through mutagenesis or genome engineering would likely lead to improved conversions and make anaerobic fungal enzymes more competitive with current commercial formulations. Similarly, further analysis of how these fungi alter CAZyme expression for diverse untreated lignocellulose may identify new enzymes optimized for certain classes of feedstocks that could be exploited for efficient bioenergy production. However, full-scale industrial exploitation will also require the development of new technologies to cultivate anaerobic fungi at large scale, and may be energetically limited by the inherent anaerobic nature of such processes.

2.5   Conclusions

In this work, we present the isolation, taxonomic placement, and characterization of a novel species of anaerobic gut fungus. We tested fungal growth on diverse untreated feedstocks to estimate the full range of CAZyme activities, and their ability to degrade plant biomass at rates sustainable for fungal growth. *Piromyces* sp. UH3-1 thrives on an array of untreated agricultural residues and bioenergy crops by hydrolyzing and fermenting the cellulosic and hemicellulosic fractions of these substrates. Importantly, we show for the first time that anaerobic fungi, such as this isolate, grow and release sugars to similar efficiencies regardless of lignin composition. Thus, this study not only highlights the ability of unengineered gut fungi to degrade diverse untreated lignocellulose, but also suggests that novel adaptations to overcome compositional variability may exist.

Characterizing these adaptations and isolating the responsible enzymes may lead to more efficient enzyme cocktails that can more fully use available renewable biomass for lignocellulosic biofuel production.

3. GENOMIC AND EPIGENOMIC REGULATION IN ANAEROBIC FUNGI

3.1 Introduction

Fungi have complex life cycles and relatively small genomes compared to eukaryotes in other kingdoms [119]. Given the need to express genes under certain environmental conditions, and at defined life cycle stages, the accessibility of the chromatin is extremely important. Therefore, precisely manipulating the chromatin structure is one means by which gene expression is controlled. While some of these changes occur by altering transcription factor expression, others occur by modifying DNA bases and or histone N terminal tails. All of these modifications fall under the term epigenetic control, which is broadly defined as non-heritable changes in gene expression. Accordingly, characterizing epigenetics as a means to regulate gene expression is a tractable approach to improve fungal genome engineering technologies. However the strategies employed to manipulate chromatin structure, and the timescales on which these changes occur varies significantly between organisms [120]. Since certain lineages of fungi are primarily employed for genetic and epigenetic studies many of the beliefs and strategies for gene manipulation are based upon a few organisms (e.g. *S. cerevisiae, S. pombe, N. crassa, T. reesei*). This knowledge, while useful may not accurately represent the understudied lineages, and bias our perspective in how to efficiently engineer fungal genomes to enhance CAZyme expression. In summary there is a need to investigate the role of chromatin structure in understudied anaerobic fungi, as this may help to increase the expression of their biomass degrading enzymes.

Designing better fungal enzyme producers depends significantly on transcription. Transcription however is affected by epigenetics as these modifications can allow local chromatin structures to be more accessible to RNA polymerase and other DNA binding factors. As epigenetics has shown to be present in many other CAZyme producing fungi, for example *Trichoderma* [121, 122], *Leptosphoria* [123], and *Botrytis* [124], it provides an additional opportunity to engineer these organisms for enhanced CAZyme production. However to engineer anaerobic gut fungi for enhanced CAZyme expression requires significant knowledge of the role of histone N terminal modifications and DNA methylation in regulating gene expression. Together, understanding the types of epigenetic signals that are possible, and the cues that lead to epigenetic

changes would significantly advance our understanding of the roles of chromatin structure and epigenetics on regulating CAZyme expression in anaerobic fungi.

Until very recently, the role of DNA methylation in anaerobic fungi has been unknown. While methylation of cytosine bases (CpG) has received much attention in eukaryotic organisms, the role of adenine methylation (m6A) is likely of more relevance given the extremely high AT content (~80%) of anaerobic fungi [44, 64]. In these organisms, the AT methylation signatures are especially prevalent near the transcriptional start sites, to as far as +1500 bases downstream [125]. Despite the increased likelihood of expression in the presence of 6mA, the quantity of these marks does not have a clear relationship with the expressional levels [125]. Further, the location of these methylated adenines varies throughout the fungal genomes, and has been proposed to be related to gene function [125]. In summary, elucidating the role of DNA methylation and the mechanisms driving the modifications of these DNA bases will need to be further investigated for characterizing how DNA becomes methylated in anaerobic fungi, and what the implications are for metabolic engineering purposes.

While DNA methylation is a major theme in epigenetics, another equally prominent concept is post translational modification of histone N-terminal tails [126]. Specifically, a variety of different covalent modifications occur including methylation (mono, di, or tri), acetylation, phosphorylation, small ubiquitin like modifier or SUMOylation, and ADP-ribosylation [127]. It is the combination of these different histone modifications that affect the structure of the nucleosomes (i.e. constitutive heterochromatin, facultative heterochromatin, or euchromatin), which is more commonly known as the histone code [126, 128]. Yet the histone codes for eukaryotic species may differ significantly. For example, the histone lysine methyltransferase Set1 (*ScSet1*) and the corresponding homologs are unique to fungi [126]. Whether anaerobic fungi have any unique DNA and histone modifying proteins remains to be elucidated. Below is a summary of some of the known modifications that have been reported in fungi as of 2017 (Table 3.1) [126]. Much of the knowledge on fungal histone post translational modifications is the presence of methylation and or acetylation of N-terminal lysine residues. However, it is known that arginine residues can be methylated and acetylated [129]. In fungi, histone 3, lysine 4 and lysine 36 methylation (H3K4me2, H3K4me3, and H3K36me3) are associated with active transcription in euchromatin, while H3K9me3 is associated with constitutive heterochromatin, and H3K27me is associated with facultative heterochromatin [126].

Table 3.1: Histone H3 methylation in selected taxa of the fungal kingdom. Presence (check mark), absence (cross) through experimental validation. Species for which only genome sequencing based evidence is available for the presence (plus-sign), or absence (minus-sign) for members of Basidiomycota, Chytrids, and Zygomycetes. Adapted from [126].

Species	H3K4	H3K9	H3K27	H3K36
Neurospora crassa	✓	✓	✓	✓
Magnaporthe oryzae	✓	+	+	+
Fusarium graminearum	✓	✓	✓	✓
Trichoderma reesei	✓	✓	+	+
Botryotinia fuckeliana	+	+	+	+
Aspergillus nidulans	✓	✓	×	✓
Aspergillus fumigatus	✓	✓	—	✓
Histoplasma capsulatum	+	+	—	+
Zymoseptoria tritci	✓	✓	✓	+
Leptosphaeria maculans	✓	✓	+	+
Candida albicans	✓	×	×	+
Saccharomyces cerevisiae	✓	×	×	✓
Schizosaccharomyces pombe	✓	✓	×	✓
Coprinopsis cinerea	+	+	+	+
Rhizoctonia solani	+	+	+	+
Crytococcus neoformans	+	✓	✓	+
Ustilago maydis	+	+	—	+
Puccinia graminis	+	+	+	+
Mucor circinelloides	+	+	—	+
Rhizopus oryzae	+	+	—	+
Batrachochytrium dendrobatidis	+	+	—	+

Taken together, characterizing histone N terminal post translational modifications provides an untapped opportunity for understanding the dynamic changes in chromatin structure.

Characterizing the ways anaerobic fungi modulate gene expression and chromatin structure requires an understanding of both the genetic changes (e.g. changing expression of a master transcription factor), and the epigenetic changes (e.g. changing DNA methylation state or histone post translational modifications). These changes may be related through gene regulatory networks, and therefore, improving genome annotation in these organisms is also critical. Thus, we aim to isolate, sequence, and annotate genomes of anaerobic fungi to understand the regulatory networks and the role of chromatin structure on gene expression by using many of the techniques that have been successful for the more highly studied lineages of fungi.

3.1.1 Applying genome engineering strategies for enhancing cellulase expression

Fungi are exploited for their ability to express a diverse array of CAZymes. Yet significant amounts of research and resources have focused exclusively on *Trichoderma reesei* for the knowledge that has accumulated since its discovery about 70 years ago. While *T. reesei* and its subsequent mutagenized strains (e.g. RUT NG14 and RUTC30) are extremely useful for producing industrial enzyme preparations, many other fungi that are equally well suited to accomplish the same task are overlooked. Unfortunately, many of the understudied fungi are not characterized due to lack of genetic knowledge and absence of engineering tools [117]. With knowledge of gene regulation, and the development of a suite of facile genome engineering tools, we will be able to manipulate anaerobic fungi in similar ways to those done for *Trichoderma reesei*. A brief review of some of these technologies and strategies is discussed below.

In *T. reesei*, promoter engineering and transcription factor engineering are successful for enhancing cellulase expression. For example, the transcriptional factor xylanase regulator 1 (XYR1) is a major activator of cellulase and xylanase expression and is involved in carbon catabolite repression (CCR) [32]. In this form of repression the presence of a sugar product (e.g. glucose) prevents the further expression of biomass degrading enzymes. While this may be beneficial for the host organism as metabolic resources are not wasted, maximizing production of the lignocellulolytic enzymes will require CCR to be bypassed. Overexpression of the XYR1 activator improves hydrolysis of sugarcane bagasse in *T. reesei* by 25% [130]. Yet an overlooked aspect of this work is the nature of the chromatin during these studies. For promoter engineering to be successful at enhancing gene expression, the promoter needs to be accessible to the RNA polymerase, which may not occur at all stages of the life cycle or under all environmental stimuli. Further, expression of these regulators (i.e. XYR1), could be linked to the expression of an epigenetic modifying factor (e.g. histone deacetylase, DNA methyltransferase) through a gene regulatory network. Lastly, the type of the promoters (i.e. constitutive, inducible, and repressible) involved in these regulatory networks may have significant implications on the overall outcome. Taken together, there is a critical need to identify how transcriptional regulators interact with different classes of promoters, and what connections exist with epigenetic modifying factors for altering chromatin structure. Collectively this will allow for genome engineering to produce stable results without pleiotropic affects.

Knowledge of chromatin structure and epigenetics will provide critical information for how gene regulatory networks adapt to changes in environmental stimuli. Additionally, this may offer insight that would be useful for determining chromosome number in these organisms. Lastly, knowledge of the histone signatures and variants may be beneficial for identifying centromeres in anaerobic fungi. Specific histone variants (i.e. CenH3), interact with centromeric DNA sequences, however the cues calling for their deposition are unknown in many fungal lineages [131]. Together, uncovering the regulation induced changes in the chromatin will provide vital information necessary for accelerating genome engineering in anaerobic fungi. Given the ability of anaerobic fungi to readily degrade plant biomass by secreting an array of lignocellulose degrading enzymes, we are working on building genome engineering tools and characterizing the fungal genomes to develop this platform. While a range of methods exist for studying and improving gene expression, I have focused my current work on chromatin accessibility and epigenetics for streamlining future technologies. The work reported below intends to provide support for pursuing epigenetics, and more specifically histone modifications to understand the ways chromatin structure may be leveraged to enhance gene expression.

3.2   Methods

3.2.1   Preparing media

All tubes and serum bottles used for in this work were prepared in fungal medium B as previously described [132]. Briefly, tubes and bottles were loaded with substrate prior to adding the reduced media. For one liter of medium B the following was added: KCl 0.6g (Fisher Scientific, Waltham, MA, USA), NaCl 0.6 g (Fisher Scientific, Waltham MA, USA), $MgSO_4*6H_2O$ 0.5 g (Fisher Scientific, Waltham MA, USA), $CaCl_2*2H_2O$ (Fisher Scientific, Waltham MA, USA), 0.2 g NH_4Cl 0.54 g (Acros Organics, NJ, USA), trypticase peptone 1 g (BD Chemical Company Greenwood Village, CO, USA), 1,4-Piperazinediethanesulfonnic acid sesquisodium salt (PIPES) 1.5 g (Fisher Scientific, Waltham, MA, USA), β-mercaptoethanol 36 μl (Sigma Aldrich, St. Louis, MO, USA), fatty acid solution 10 ml, trace elements solution 10 ml, haemin solution 10 ml, 1000X resazurin (1 g/l 1000X) 1 ml (Acros Organics, NJ, USA), Na_2CO_3 4 g (Fisher Scientific, Waltham, MA, USA), KH_2PO_4 0.7 g (Fisher Scientific, Waltham, MA, USA), and yeast extract 0.5 g (Fisher Scientific, Waltham, MA, USA), to a one liter flask and the resulting solution was brought to 1

liter final volume by adding sterile ddH_2O. This solution was then microwaved for approximately 12 minutes prior to bubbling with CO_2 (Indiana Oxygen, West Lafayette, IN, USA) for 15 minutes to remove any oxygen present in the media. Prior to aliquoting into individual tubes and serum bottles, cysteine hydrochloride 1.25 g (Fisher Scientific, Waltham, MA, USA), was added for the final reduction of the media. Tubes and bottles were sealed and autoclaved for 30 minutes at 120 °C.

For 1 liter of the fatty acid solution, acetic acid 6.85 ml (Fisher Scientific, Waltham, MA, USA), propionic acid 3.0 ml (Acros Organics, NJ, USA), n-butyric acid 1.84 ml (Acros Organics, NJ, USA), 2-methylbutyric acid 0.55 ml (Acros Organics, NJ, USA), isobutyric acid 0.47 ml (Acros Organics, NJ, USA), valeric acid 0.55 ml (Acros Organics, NJ, USA), and isovaleric acid 0.55 ml (Acros Organics, NJ, USA) were dissolved in 700 ml of 0.2 M sodium hydroxide (Fisher Scientific, Waltham, MA, USA). The pH of the solution was adjusted to 7.5 with sodium hydroxide and the volume was adjusted to 1 liter with sterile H_2O.

For one liter of the trace elements solution, $MnCl*4H_2O$ 0.25 g (Fisher Scientific, Waltham, MA, USA), $NiCl_2*6H_2O$ 0.25 g (Acros Organics, NJ, USA), $NaMoO_4*2H_2O$ 0.25 g (Acros Organics, NJ, USA), H_3BO_3 0.25 g (Acros Organics, NJ, USA), $FeSO_4*7H_2O$ 0.20 g (Acros Organics, NJ, USA), $CoCl_2*6H_2O$ 0.05 g (Acros Organics, NJ, USA), SeO_2 0.05 g (Acros Organics, NJ, USA), $NaVO_3*4H_2O$ 0.05 g (Acros Organics, NJ, USA), $ZnCl_2$ 0.025 g (Acros Organics, NJ, USA), $CuCl_2*2H_2O$ 0.025 g (Acros Organics, NJ, USA) were all prepared in 0.2 M HCl (Fisher Scientific, Waltham, MA, USA).

For 1 liter of the haemin solution, haemin from porcine 0.1 g (Alfa Aesar, Haverhill, MA, USA), was dissolved in ethanol 10 ml, and the volume was adjusted to one liter with 0.05 M NaOH.

For 1 liter of the vitamin solution, 1-4 naphthoquinone 0.25 g (Acros Organics, NJ, USA), calcium D-pantothenate 0.2 g (Acros Organics, NJ, USA), nicotinamide 0.2 g (Acros Organics, NJ, USA), riboflavin 0.2 g (Acros Organics, NJ, USA), thiamine HCl 0.2 g (Fisher Scientific, Waltham, MA, USA), pyridoxine HCl 0.2 g (Acros Organics, NJ, USA), biotin 0.025 g (Fisher Scientific, Waltham, MA, USA), folic acid 0.025 g (Acros Organics, NJ, USA), cyanocobalamin 0.025 g (Acros Organics, NJ, USA), and para-aminobenzoic acid (PABA) 0.025 g (Acros Organics, NJ, USA) were prepared in 5 mM 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES) buffer (Acros Organics, NJ, USA).

Four different substrates prepared in medium B were tested: glucose (5 mg/ml) (Fisher Scientific, Waltham, MA, USA), xylan from beechwood (10 mg/ml) (Crescent Chemical, Islandia, New York, USA), corn stover (10 mg/ml) (Pioneer variety 1197 harvested in Rensselaer, Indiana, USA), and wild type poplar (10 mg/ml) (variety INRA 717) [72, 74]. Autoclaved media incubated at 39 °C for 1 hour, and then received chloramphenicol (Fisher Scientific, Waltham, MA, USA), 3.5 mg/ml dissolved in 60% EtOH (Fisher Scientific, Waltham, MA, USA), 40% H_2O, that had been sterile filtered, and purged with CO_2 to minimize the amount of O_2 added to the cultures. The final concentration of chloramphenicol in the media was 3.5 μg/ml. Media was then inoculated with 1 ml for Hungate tubes, while serum bottles were inoculated with 3 ml from an actively growing fungal culture in mid exponential phase. Tubes and bottles were zeroed by venting off any positive pressure and were stored at 39 °C until further use.

3.2.2 Culturing anaerobic fungi in the presence of epigenetic antagonists

To test the effects of interfering with DNA methylation in anaerobic fungi, 5'azacytidine (Sigma Aldrich, St. Louis, MO, USA), a potent DNA methyltransferase inhibitor was tested [133]. Similarly to test the effects of interfering with histone deacetylation, suberoylanilide hydroxamic acid (SAHA) (Sigma Aldrich, St. Louis, MO, USA), and n-butyric acid (Sigma Aldrich St. Louis, MO, USA) were tested for their effects on interfering with fungal growth and protein expression [134, 135]. 5'azacytidine and SAHA were dissolved in deionized water and dimethyl sulfoxide (Sigma Aldrich, St. Louis, MO, USA) respectively. These solutions were either added prior to inoculation of the fungal culture or during mid-exponential phase (~72-96 hours), where specified. All inhibitor solutions were purged with CO_2 (Indiana Oxygen, West Lafayette, IN, USA) to minimize the amount of oxygen added to the culture media.

3.2.3 Harvesting fungal biomass for western blot analyses

To determine some of the possible histone modifications in anaerobic fungi, we prepared 50 ml serum bottles containing medium B as described previously (Methods 3.2.1) with glucose as the substrate (5 mg/ml) [132]. Fungal cultures were grown at 39 °C until approximately mid-exponential phase, or when enough fungal biomass was present to harvest histone protein (approximately 3-5 days). Fungal cultures were vented daily to monitor growth. To isolate the fungal biomass, we transferred the serum bottles to an anaerobic chamber (PLAS labs, Lansing,

MI, USA), and quickly decanted the fungal biomass. We immediately transferred this fungal biomass to falcon tubes and submerged them in liquid nitrogen to flash freeze them. The falcon tubes containing the frozen fungal biomass were stored at -80 °C until further use. This frozen fungal biomass was used directly for SDS PAGE protein gels by adding SDS buffer (Fisher Scientific, Waltham, MA, USA) and denaturing the samples. Western blots were then performed on these gels with various antibodies for detecting histone modifications.

3.2.4   Isolation of the carbohydrate active enzymes (CAZymes)

We isolated fungal CAZymes to assay for hemicellulolytic activities. The total fugal secretome was harvested by separating the spent culture media (~50 ml) from the fungal biomass. This spent culture media was then used for activity assays. Additionally, we used half of the culture media (~25mls) to isolate and enrich for the cellulose binding enzymes [53]. This procedure exploits the cellulose binding domains of CAZymes to isolate carbohydrate degrading enzymes [53]. Cultures were centrifuged at 12,800 g and the supernatant was transferred to a tube containing approximately 0.4% (w/v) type 20 Sigmacell (Sigma Aldrich, St Louis, MO, USA). These tubes were incubated two hours at 4 °C with gentle agitation. Tubes were then centrifuged at 12,800 g at 4 °C and the supernatant was discarded. Potassium phosphate mono and dibasic salts (Fisher Scientific, Waltham, MA, USA) were dissolved to make phosphate buffer 0.1M pH 7.0. This phosphate buffer was added to the Sigmacell, and the tubes incubated at room temperature with gentle agitation for one hour to elute the cellulose binding enzymes. The elutions were then stored at 4 °C for further analysis. Protein concentrations were determined by the bicinchoninic assay (Fisher Scientific, Waltham, MA, USA) [136].

3.2.5   SDS PAGE and zymography analyses for detailed enzyme characterization

For the SDS-PAGE analysis of our cellulose binding proteins, we casted 10% acrylamide gels and ran them at 110 volts for sufficient protein separation. Gels were then stained with Sypro Ruby protein stain (Fisher Scientific, Waltham, MA, USA). Zymography gels were prepared by adding 0.4% w/v carboxy methyl cellulose (CMC) (Acros Organics, NJ, USA) to the resolving portion of a 10% acrylamide gel. These samples were run under non-denaturing conditions with a lower voltage (55 volts) at 4 ℃. The SDS was removed with slight modification to the procedure of Tseng et al 2002 [82]. Samples were rinsed with ddH$_2$O and placed in 0.1M Tris-NaCl buffer

pH 7.0 containing 25% (w/v) isopropanol (TNI buffer). Zymogram gels incubated for 30 minutes at 4 ℃ in TNI buffer with gentle agitation. The TNI buffer was then removed, and fresh TNI buffer was added a total of three times. The zymograms were then washed with Tris-NaCl buffer (TN-buffer) 0.1M pH 7.0 prior to incubating at 39 ℃ for substrate hydrolysis. We incubated the CMC zymograms for approximately 18 hours to allow for substrate hydrolysis. Zymograms were then soaked in 0.1% w/v Congo red stain (Fisher Scientific, Waltham, MA, USA), for approximately 30 minutes, and de-stained with 1M NaCl until the hydrolysis zones appeared relative to the red background. We fixed the zymograms with 0.1M acetic acid (Fisher Scientific, Waltham, MA, USA) prior to imaging.

3.2.6   Determination of fungal xylanolytic activities

Piromyces sp. UH3-1 xylanase activity was measured on the entire fungal secretome and also on the purified cellulose binding proteins as discussed above. Briefly, we followed the 96 µl microplate procedure introduced by Xiao et al [83]. However, we used 0.05M sodium phosphate buffer (pH 6.5), and a 2% solution of xylan from beechwood (Crescent Chemical, Islandia, NY, US). Substrate hydrolysis occurred for one hour at 50 ℃. All samples were run in triplicate, all protein loading values were normalized, and a control to determine the extent of non-enzymatic xylan degradation was included. A glucose standard curve was run to determine sugar concentrations under these reaction conditions. Reducing sugars were calculated by measuring the absorbance at 540 nm on a Synergy Neo plate reader (Biotek, Winooski, VT, USA).

3.3   Results and discussion

As there may be subtle links between expression of epigenetic factors or chromatin modifying proteins with carbohydrate active enzyme expression, we first analyzed published anaerobic fungal RNAseq data. While anaerobic fungi transcribe an array of biomass degrading enzymes, the expression of any of their DNA methyltransferases (DNMT), histone deacetylates (HDAC), and histone acetyltransferases (HAT) remains uncharacterized [53]. To identify the trends between these two groups of genes, we searched for CAZyme families having both positive and negative correlations with these epigenetic factors (Figure 3.1). Notably, anaerobic fungi have multiple CAZyme families that are coregulated with DNA and histone modifying genes. Further, some of the fungal CAZyme families show correlations between all three of the epigenetic

modifying factors that we have selected (Figure 3.1). This suggests that the degree to which a given CAZyme family is epigenetically regulated depends partly on the activity of the enzyme family. Therefore, transcription of selected histone and DNA modifying proteins are correlated with transcription of multiple CAZyme families in anaerobic fungi. Given the presence of these relationships, we then sought ways to interfere with epigenetic regulation in anaerobic fungi to determine the potential effects it may have on growth and enzyme activity.
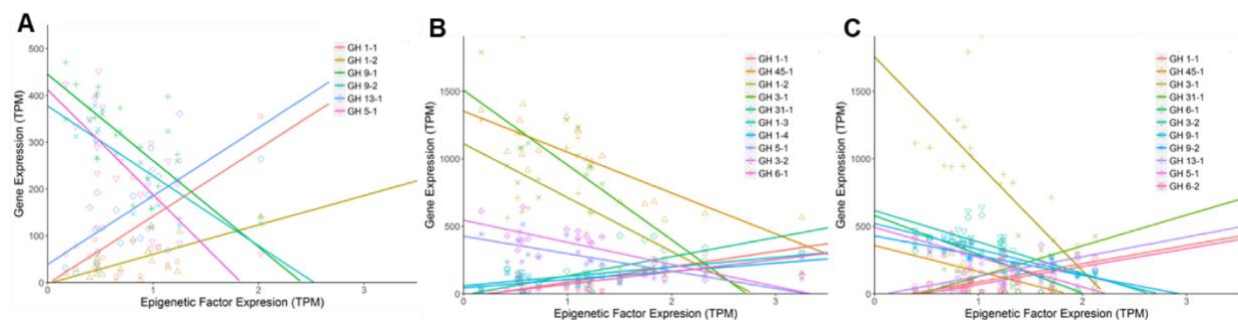
Figure 3.1: Anaerobic fungal CAZyme expression is correlated with histone and DNA modifying protein expression: Expression of A) histone deacetylase; B) H3K56 acetyltransferase; C) DNA methyltransferase versus regulated CAZymes in transcripts per million (TPM) in *Neocallimastix californiae*. ANOVA p-value ≤ 0.05, $R^2 \geq$ to 0.4. [53]

To investigate the role of HDAC and DNMT activities in anaerobic fungi, we acquired known small molecule epigenetic antagonists, and observed what effects they had on fungal growth and CAZyme activity. The compounds selected inhibit the activity of DNA methyltransferases (5' azacytidine) or the activity of histone deacetylases (suberoylanilide hydroxamic acid (SAHA), and butyric acid (Figure 3.2) [133, 137, 138]. To ensure that the observed changes in gene expression were not due to inhibitor toxicity, we tested fungal growth in the presence of these chemicals at a range of concentrations (Methods 3.2.2). Even at the micromolar range of inhibitor, changes to fungal biomass morphology and pressure accumulation were observed (Appendix C: Figures C.1-C.4). Initial rounds of culturing indicated that at 100 μM, both SAHA and 5' azacytidine were lethal with glucose as the substrate. Notably, when cultured on a lignocellulosic substrate (i.e. poplar), the effects of 5' azacytidine appeared less pronounced for *Neocallimastix sp*. GF-ma (Appendix C Figures C.3-C.4).

5' Azacytidine **SAHA** **Butyric Acid**
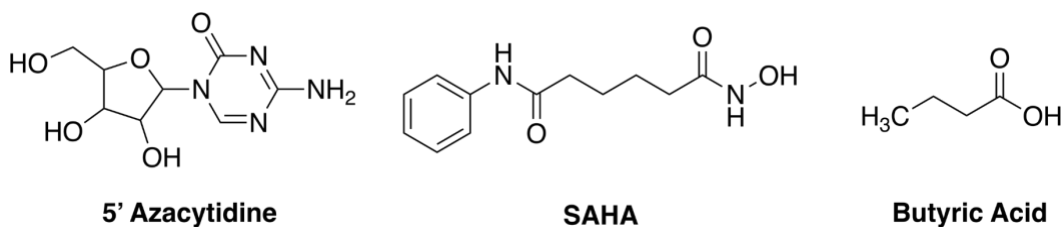
Figure 3.2: Structures of the small molecule inhibitors being explored for their ability to modify fungal epigenetic status

As the effects of SAHA appeared to produce equal results (i.e. equally toxic at high concentrations across a range of feedstocks), we focused additional experiments exclusively on that inhibitor. Together, the effects of SAHA on fungal growth suggested that reduced inhibitor concentration ranges would be required to detect any changes in CAZyme expression and activity. As the SAHA concentrations tested (2.5 μM and 100 μM) were at the low and high extremes when added prior to inoculation, we wanted to determine if more pronounced changes in fungal gene expression would be observed when adding the inhibitor during mid exponential phase rather than prior to inoculation.

In adding the histone deacetylase inhibitors prior to inoculation, it may be possible that the anaerobic fungi have an increased susceptibility to SAHA given the more compact structure, and reduced need to express an array of CAZymes. In the zoospore state (Figure 2.1C), anaerobic fungi would not be expected to express many of the genes necessary for biomass hydrolysis, as this is prior to encystment and zoospore germination. Thus many genes may be repressed under these conditions. Therefore, by adding in high concentrations of SAHA, a global derepression of transcription might be expected to occur, and be lethal due to a high metabolic burden. In summary, this could be addressed one of two ways, either by adding reduced concentrations of SAHA at inoculation, or change the time at which SAHA is added.

To observe how SAHA treatment impacts fungal growth and CAZyme activity, and if reductions in histone deacetylation lead to a more temporary response in anaerobic fungi, we focused on treating these fungi during mid-exponential phase. By treating actively growing fungal cultures in exponential phase (~90 hours) with SAHA for an exposure period of 6 hours, we would

be able to immediately collect the fungal secretomes of the cultures and assay for endocellulase activity. While our zymographic results (Methods 3.2.2 and 3.2.6) did not show any significant changes for endocellulase activity for either *Piromyces* sp. UH3-1 or *Neocallimastix* sp. Gf-ma at all SAHA treated levels (Figure 3.3), none of the cellulases appeared to have negative correlations with increasing inhibitor concentrations. Together, adding HDAC inhibitors during mid-exponential phase, and an exposure time of 6 hours does not inhibit the expression or activity of endocellulases from two genera of anaerobic fungi. (Figure 3.3).

While no changes in cellulase activity were observed for the two isolates of anaerobic fungi treated with SAHA during mid-exponential phase, it is possible that other families of CAZymes may be more strongly regulated through epigenetic control (Figure 3.1). Further, as the types of histone modifications may vary significantly (Table 3.1), understanding the relationship between histone post translational modification and enzyme activity would provide a clear route forward for engineering purposes. Therefore, identifying some of the types of histone modifications and mapping these to changes in CAZyme activity would accomplish this goal. To do this, anaerobic fungi cultured on simple substrates such as glucose were selected, given the ease of harvesting the fungal biomass from the fermentation media (Figure 2.3A). Further, soluble substrates such as this would also inform if carbon catabolite repression has any epigenetic levels of regulation. In other words, by treating fungal cultures with SAHA, any increases in CAZyme expression would support a role of histone acetylation for increasing transcription of biomass degrading enzymes. Additionally, by testing the fungal secretomes for CAZyme activity, and the associated fungal biomass for specific histone signatures, we would be able to see which signatures are associated with more active transcription.
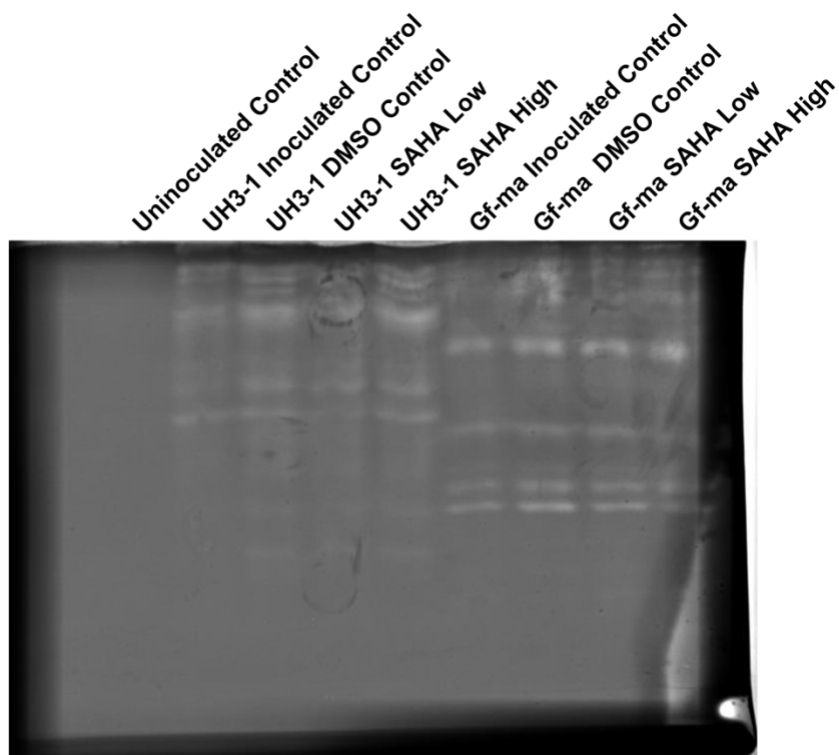
Figure 3.3: CMC zymogram testing the fungal secretomes for changes in endocellulase activity. The secretomes are collected from fungal isolates *Neocallimastix sp. GF-Ma* and *Piromyces* sp. UH3-1 cultured in the presence of SAHA with glucose as the substrate. The inoculated control contains no DMSO or inhibitor, the DMSO control contains 10% DMSO without SAHA, and the cultures with low SAHA (10 μM) and high SAHA (50 μM) are labeled accordingly. No ladder can be included as the CMC changes the migration patterns, and reverse staining is used, which binds to the substrate in the zymogram instead of the proteins.

In culturing one fungal isolate on glucose in the presence or absence of SAHA, the role this inhibitor has on CAZyme activity and histone poste translational modifications could be tested. As the literature supports strong xylanolytic activities for anaerobic fungi, we monitored xylanase activity [139]. Culture supernatants from fungal isolate *Neocallimastix* sp. WI3B showed significant changes in xylanase activity when treated with SAHA (Figure 3.4). Notably, this change was observed both on the full secretome, and the enriched cellulose binding enzymes (Figure 3.4). In summary, this data supports a role for histone acetylation in promoting xylanase activity in one genus of anaerobic fungi. Given the encouraging results for histone acetylation in

modulating xylanase activity, understanding which histone N-terminal modifications are associated with SAHA treated cultures would help bridge this knowledge gap.
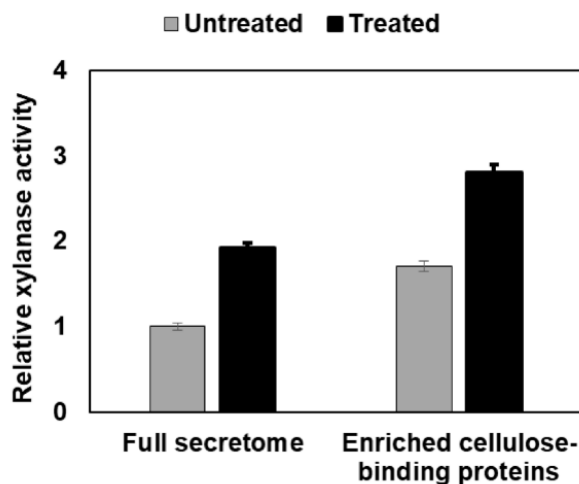
Figure 3.4: Fungal xylanase activity in the presence of SAHA: SAHA treatment at 17 μM upregulates xylanase activity in *Neocallimastix sp*. WI-3. Xylanase activity of culture supernatant (secretome), and enriched cellulose binding portion at 50 °C.

While the *Neocallimastix* WI3B fungal biomass from the activity analysis (Figure 3.4) was harvested for Western blotting, the pipeline for preparing and running the Western blots was still being refined (results not shown). After a few more attempts at running the western blots on SAHA treated fungal cultures, changes in histone post translational modification were observed. In testing the fungal biomass of the *Piromyces* sp. UH3-1 isolate, increases in H3K4 trimethylation, and a concomitant decrease in methylation of H3K27 were observed for SAHA treated cultures (Figure 3.5). Notably, as *S. cerevisiae*, *S. pombe*, *T. reesei*, and *N. crassa*, are used frequently for epigenetic analyses, these results show that signatures which are absent in *S. cerevisiae* are present in anaerobic fungi (Table 3.1), thus supporting the hypothesis that using small numbers of fungal lineages may misrepresent the range of histone post translational modifications in the fungal kingdom. Taken together, SAHA treatment not only leads to enhanced xylanase activity, but it also changes the types of histone modifications observed in anaerobic fungi. Further, the histone

modifications tested are absent in *S. cerevisiae*, highlighting the need to investigate other fungal lineages for epigenetic analyses
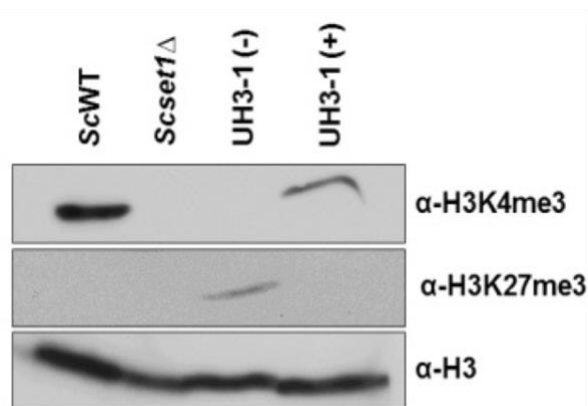
Figure 3.5: Western blotting shows the presence of covalent histone modifications in anaerobic fungi: Histone modification can be controlled in anaerobic fungi with chemical supplementation. Whole cell extracts from *Piromyces* sp. UH3-1 treated with (+) or without (-) histone deacetylase inhibitor SAHA. Histone site and methyl specific antibodies were used to detect H3K4 trimethylation. *Saccharomyces cerevisiae* wild type *Sc*WT and *set1Δ* (*Scset1Δ*) strains were used as controls for H3K4 trimethylation. Set 1 is the sole histone H3K4 methyltransferase in *S. cerevisiae*. *S. cerevisiae* lacks H3K27 methylation. Histone H3 was used as a loading control.

3.4   Conclusions

While the role of epigenetics in anaerobic fungi has been investigated in this work, the extent to which epigenetics is used as a form of control in anaerobic fungi remains untested. In the current work, some of the preliminary data suggests a role of histone deacetylases for modulating fungal xylanase activity and histone N terminal modifications. This work highlights the role of histone methylation and acetylation in modulating fungal CAZyme expression. Published RNAseq data for *N. californiae* show multiple relationships between histone modifying proteins and various CAZyme families. Additionally, this work shows that small molecule antagonists such as SAHA can be used at non-toxic concentrations to interfere with fungal CAZyme activity and may provide a means by which anaerobic fungi modulate xylanase activity. These small molecules can also be used to detect different histone post translational modifications, which to date has never been

investigated in anaerobic fungi. Thus this work shows that epigenetics may be a strategy by which anaerobic fungi regulate expression of their biomass degrading enzymes.

4. CONCLUSIONS AND FUTURE WORK

The need to more efficiently degrade lignocellulosic materials has been met with some success over the past few years. Through identification of new enzymes, for example lytic polysaccharide monooxygenases, significant improvements to enzyme preparations has resulted [140]. However, the recalcitrance of lignocellulose is still a formidable task that prevents this feedstock from being an economical source for renewable energy, and enzyme production costs are a significant bottleneck [28]. Hence engineering anaerobic fungi such as the species we have explored, which are naturally suited to degrade plant biomass provides a possible solution to some of these challenges. Anaerobic fungi harbor significant numbers of biomass degrading enzymes, far more than what are present in *Trichoderma* or *Aspergillus*, and exploit different strategies to hydrolyze the plant material, which may more efficiently release the sugars of lignocellulose [47]. While there are some basic challenges in cultivating these organisms, simply using them as a base platform for novel industrial cocktails may provide a more efficient path forward than current industrial enzyme technologies.

In order to engineer anaerobic fungi as a platform for robust lignocellulose hydrolysis, we have isolated and characterized previously uncultivated strains of anaerobic fungi. We show that these isolates robustly degrade diverse plant biomass of mixed lignin composition under mild conditions. I also show that expression of their carbohydrate active enzymes is partly regulated via epigenetic mechanisms. By thoroughly characterizing the role of epigenetics in modulating CAZyme expression, we may be able to leverage this as a strategy for overexpressing fungal biomass degrading enzymes.

Future work will involve identifying more of the histone modifications in anaerobic fungi as well as key transcriptional regulators that modulate expression of carbohydrate active enzymes. Similarly, I plan to identify the core genetic parts required for genome engineering purposes (i.e. centromeres, autonomously replicating sequences, and promoters), as well as build genome engineering tools. I hope to use the knowledge of epigenetics and our future engineering tools to develop anaerobic fungi as a production platform for robust lignocellulolytic enzymes that may more efficiently release the sugars in untreated or mildly pretreated plant biomass.

APPENDIX A. COMPOSITIONAL ANALYSIS OF THE PLANT BIOMASS USED IN THIS STUDY

Table A.1: NREL compositional analysis of the renewable plant biomass used in this study

Component	Corn Stover	Switchgrass	Sweet Sorghum	Grain Sorghum	Forage Sorghum	Orange Peel	Fresh Cut Alfalfa	Dry Bale Alfalfa
Glucan (%)	37.7 ± 1.1	34.8 ± 0.28	50.8 ± 0.40	44.7 ± 1.53	39.5 ± 1.37	22.8 ± 5.5	15.2 ± 4.4	15.6 ± 3.1
Xylan (%)	22.1 ± 0.42	22.1 ± 0.42	11.2 ± 0.04	14.0 ± 0.88	20.5 ± 0.61	14.7 ± 0.3	5.1 ± 1.8	6.3 ± 1.2
Arabinan (%)	3.4 ± 0.09	2.7 ± 0.01	2.2 ± 0	3.3 ± 0.25	2.8 ± 0.09	6.1 ± 1.2	2.1 ± 0.1	2.3 ± 0.6
Acetyl (%)	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Total Lignin (%)	17.7 ± 1.5	23.3 (NR)	16.1 ± 0.34	14.8 0.01	20.5 ± 0.38	ND	ND	ND
Water Extractives (%)	11 (NR)	6.8 (NR)	22.3 (NR)	34.0 (NR)	27.5 (NR)	0.556 (NR)	30.4 (NR)	35.7 (NR)
Ethanol Extractives (%)	4.5 (NR)	2.4 (NR)	7.1 (NR)	5.5 (NR)	3.8 (NR)	0.024 (NR)	6.0 (NR)	4.4 (NR)
NR- No Replicates, ND- No data								

Table A.2: NREL compositional analysis of 2014 harvested poplar constructs used.

Component	INRA 717	NM 6	0998-45	1036-73	1020-44	1035-41	F5H 37	F5H 64
Glucan	0.43 ± 0.00	0.41 ± 0.00	0.43 ± 0.00	0.45 ± 0.00	0.43 ± 0.00	0.45 ± 0.00	0.43 ± 0.00	0.44 ± 0.00
Xylan	0.21 ± 0.01	0.20 ± 0.00	0.22 ± 0.00	0.22 ± 0.00	0.22 ± 0.00	0.21 ± 0.00	0.22 ± 0.00	0.21 ± 0.01
Arabinan	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
Acetyl	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00	0.02 ± 0.00
Acid Insoluble Residue	0.15 ± 0.00	0.17 ± 0.00	0.14 ± 0.00	0.17 ± 0.00	0.15 ± 0.00	0.17 ± 0.00	0.14 ± 0.00	0.17 ± 0.00
Acid Soluble Lignin	0.06 ± 0.00	0.06 ± 0.00	0.04 ± 0.00	0.05 ± 0.00	0.08 ± 0.00	0.05 ± 0.00	0.09 ± 0.00	0.06 ± 0.01
Ash	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.00 ± 0.00
Glucose	0.01 ± N/A	0.01 ± N/A	0.01 ± N/A	0.01 ± N/A	0.00 ± N/A	0.01 ± N/A	0.01 ± N/A	0.00 ± N/A
Xylose	0.01 ± N/A	0.00 ± N/A	0.01 ± N/A	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A
Arabinose	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A
Acetic Acid	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A	0.00 ± N/A
Other Water Extractives	0.05 ± N/A	0.05 ± N/A	0.04 ± N/A	0.02 ± N/A	0.00 ± N/A	0.04 ± N/A	0.02 ± N/A	0.00 ± N/A
Ethanol Extractives	0.02 ± N/A	0.05 ± N/A	0.05 ± N/A	0.04 ± N/A	0.05 ± N/A	0.04 ± N/A	0.04 ± N/A	0.04 ± N/A
NA- No replicates								

Table A.3: Syringyl lignin content of the 2014 poplar constructs used in this study.

Construct	Syringyl lignin molar percentage
0998-45	5
1036-73	28
1065-41	34
1020-44	42
INRA 717	64
NM-6	68
F5H-37	77
F5H 64	98

Table A.4: NREL compositional analysis of the 2017 harvested poplar

	Control Lines				Low Syringyl Lignin/ Hydroxy Guaicyl Lignin				Low Syringyl Lignin/ High Guaicyl Lignin				High Syringyl Lignin			
	INRA 717		NM 6		0998-45		1036-73		1020-44		1035-41		F5H 37		F5H 64	
Component	% dry mass	Std. Dev.	% dry mass	Std. Dev.	% dry mass	Std. Dev.	% dry mass	Std. Dev.	% dry mass	Std. Dev.	% dry mass	Std. Dev.	% dry mass	Std. Dev.	% dry mass	Std. Dev.
Glucan	43.40%	0.30%	41.46%	0.41%	43.49%	0.41%	44.96%	0.17%	43.26%	0.21%	44.54%	0.21%	43.22%	0.25%	44.49%	0.34%
Xylan	21.20%	0.66%	19.52%	0.20%	22.14%	0.26%	21.76%	0.11%	22.14%	0.32%	20.97%	0.26%	22.04%	0.18%	21.48%	0.56%
Arabinan	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Acetyl	1.80%	0.06%	1.62%	0.03%	1.83%	0.03%	1.79%	0%	1.83%	0%	1.79%	0.14%	1.82%	0%	1.87%	0.03%
Acid Insoluble Residue	15.46%	0.18%	16.91%	0.18%	14.25%	0.06%	16.51%	0.28%	15.01%	0.18%	16.78%	0.39%	13.97%	0.12%	16.67%	0.07%
Acid Soluble Lignin	5.90%	0.28%	5.96%	0.30%	4.08%	0.16%	4.66%	0.21%	7.99%	0.26%	4.64%	0.26%	9.05%	0.28%	5.62%	0.54%
Ash	0.70%	0.05%	0.81%	0.01%	0.69%	0.18%	0.61%	0.11%	0.61%	0.05%	0.53%	0.06%	0.58%	0.02%	0.49%	0.07%
Glucose	0.99%	-	0.74%	-	1.37%	-	0.68%	-	0%	-	0.69%	-	0.70%	-	0%	-
Xylose	0.79%	-	0%	-	0.53%	-	0.40%	-	0%	-	0%	-	0%	-	0%	-
Arabinose	0%	-	0%	-	0%	-	0%	-	0%	-	0%	-	0%	-	0%	-
Acetic Acid	0%	-	0.12%	-	0%	-	0%	-	0%	-	0%	-	0%	-	0%	-
Other Water Extractives	5.41%	-	4.86%	-	3.76%	-	1.95%	-	0%	-	4.29%	-	1.57%	-	0.31%	-
Ethanol Extractives	2.43%	-	4.73%	-	4.55%	-	3.70%	-	4.55%	-	3.64%	-	4.29%	-	3.94%	-
Mass Balance	98.07%	-	96.74%	-	96.68%	-	97.03%	-	95.39%	-	98.24%	-	97.24%	-	94.88%	-

Table A.5: Lignin analysis of the 2017 harvested poplar

	Control Lines				Low Syringyl Lignin/ Hydroxy Guaicyl Lignin				Low Syringyl Lignin/ High Guaicyl Lignin				High Syringyl Lignin			
	INRA 717		NM 6		0998-45		1036-73		1020-44		1035-41		F5H 37		F5H 64	
Lignin Amount (nmol mg^{-1} cell wall)	Amount	Std. Dev.	Amount	Std. Dev.	Amount	Std. Dev.	Amount	Std. Dev.	Amount	Std. Dev.	Amount	Std. Dev.	Amount	Std. Dev.	Amount	Std. Dev.
H	1.7	0	1.2	0.4	0	0	1.5	0.1	2.2	0.6	0.7	1	1.7	0.4	1.6	0.2
G	80.5	0.3	75.1	0.6	65.3	2.6	88.3	0.7	101.3	9.3	100.7	2.4	17.2	0.2	35.4	1
S	164.1	0.9	152.2	12.2	17.2	1.3	107.3	0.5	136.1	6.4	111.8	0.8	253.4	13.8	221.6	0.5
total	246.3	1.2	228.5	11.2	82.5	3.9	197.1	0.1	238.7	16.4	213.2	2.5	272.2	14	258.5	1.6
Percent	%	Std. Dev.	%	Std. Dev.	%	Std. Dev.	%	Std. Dev.	%	Std. Dev.	%	Std. Dev.	%	Std. Dev.	%	Std. Dev.
H	0.7	0%	0.5	0.2%	0	0.0%	0.8	0.0%	0.9	0.2%	0.3	0.5%	0.6	0.1%	0.6	0.1%
G	32.7	0%	32.9	1.9%	79.2	0.6%	44.8	0.3%	42.2	1.0%	47.2	0.6%	6.3	0.4%	13.7	0.3%
S	66.6	0%	66.5	2.1%	20.8	0.6%	54.4	0.3%	56.8	12.0%	52.4	1.0%	93.1	0.3%	85.7	0.4%

Table A.6: Glucan and xylan conversion efficiencies for *Piromyces* sp. UH3-1 when grown on untreated corn stover for 168 hours

	Corn Stover		
Glucan in Raw (g)	0.377	±	0.011
Glucan in Spent (g)	0.156	±	0.010
Xylan in Raw (g)	0.221	±	0.004
Xylan in Spent (g)	0.157	±	0.002
Glucan Released	58.6%	±	3.6%
Xylan Released	29.0%	±	2.0%

APPENDIX B ADDITIONAL DATA FOR *PIROMYCES* SP. UH3-1

Formal Species Description

Taxonomy

Piromyces sp. UH3-1   Ethan Hillman, Adrian Ortiz-Velez, Kevin Solomon, sp. nov.

Index Fungorum number: IF554555   JMRC: SF:012426

Typification: The holotype (Figure 2.1A) derived from the following: USA, INDIANA: Independence, 40.34° N, 18.17° W, ~170m above sea level, 3 day old culture of isolate UH3-1, originally isolated from the feces of a donkey (*Equus africanus asinus*), July 2016, Ethan Hillman. Ex-type strain: UH3-1. GenBank: ITS1-5.8S-ITS2 = KY494854

Etymology: The epithet honors the host organisms from which this fungus was isolated.

*A*n anaerobic fungus with a determinate (finite) life cycle displaying a monocentric thallus. The fungi exhibit endogenous zoosporangial development where the encysted zoospore retain the nucleus. The encysted zoospore geminates to form a rhizoidal network and a single oval or balloon-shaped sporangium (20-75 µm long and 20-30 µm wide), which on maturity liberates many zoospores. The rhizoidal system is devoid of nuclei (as seen under DAPI staining; Figure 2.1D-E) and is highly branched. Free swimming zoospores (Figure 2.1C) are typically spherical (10 µm diameter) and the species is characterized by the presence of a single posteriorly directed flagella (~30 µm long); the flagella propels the zoospore forward toward plant material/nutrient sources (chemotaxis) [71].

The clade is defined by the sequence KY494854 for ITS1-5.8S-ITS2
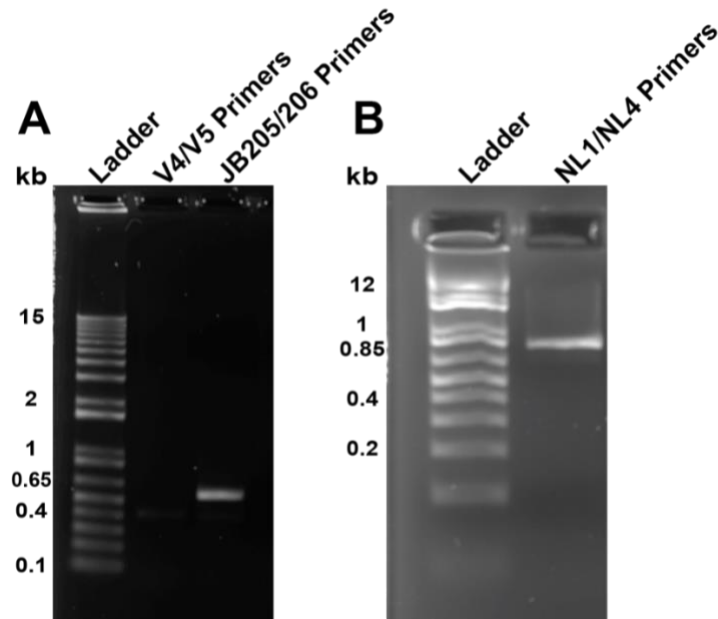
Figure B.1: *Piromyces* sp. UH3-1 DNA controls: Lane 2 V4/V5 primers don't lead to amplification of *Piromyces* sp. UH3-1 DNA, Lane 3 JB206/JB205 primers lead to amplification of the ITS1 region of the *Piromyces* sp. UH3-1 genome [65].
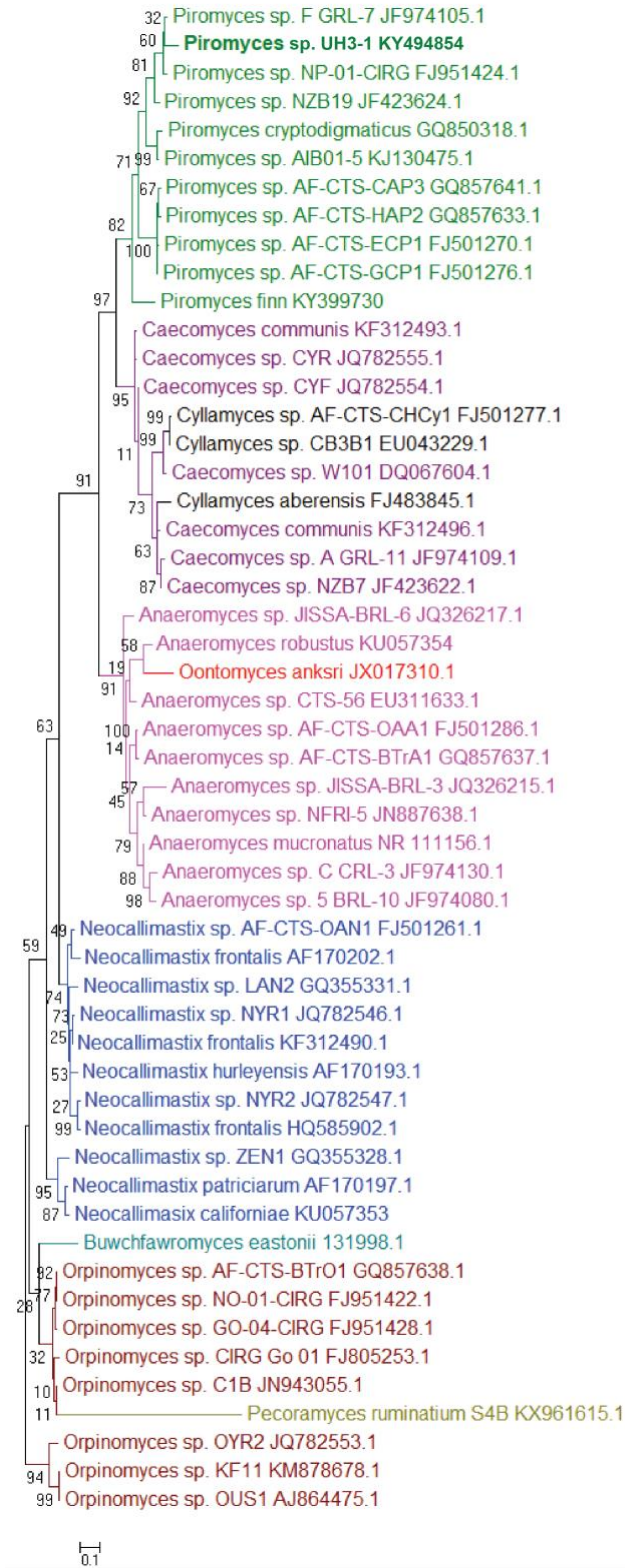
Figure B.2: Expanded *Piromyces* sp. UH3-1 ITS1 phylogenetic tree with accession numbers [65].
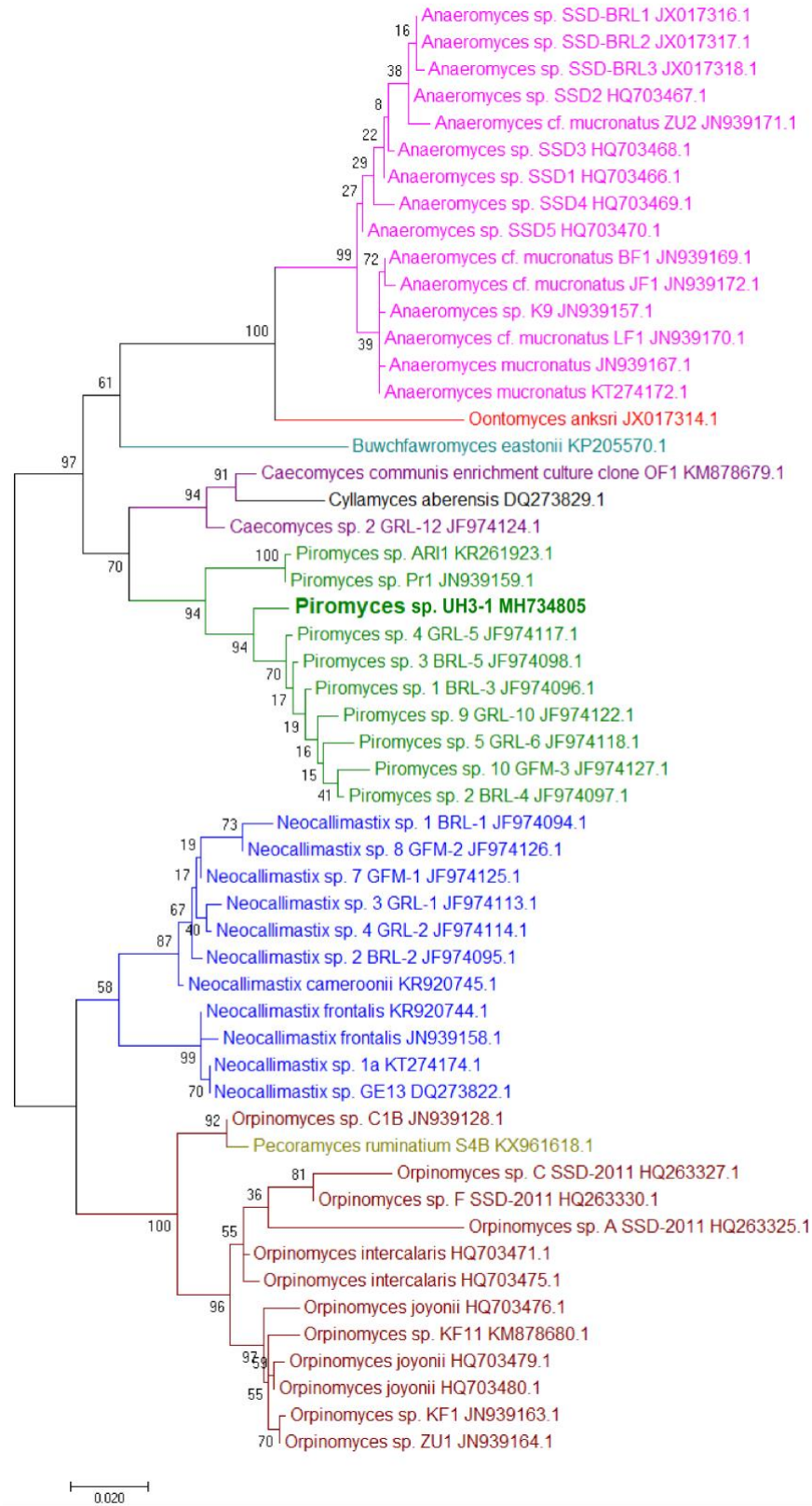
Figure B.3: Expanded *Piromyces* sp. UH3-1 LSU phylogenetic tree with accession numbers [65]
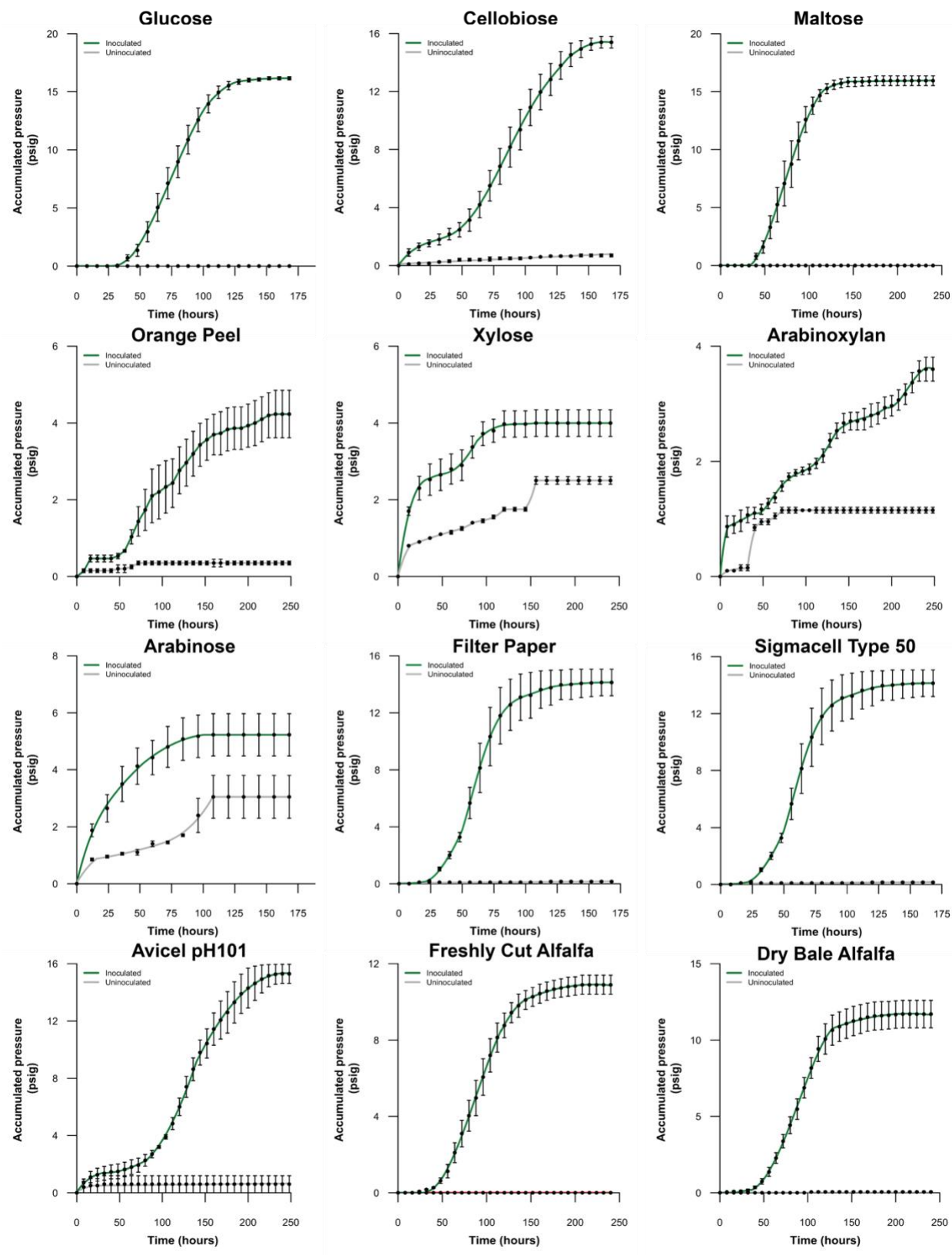
Figure B.4: Selected growth curves for *Piromyces* sp. UH3-1 [65].
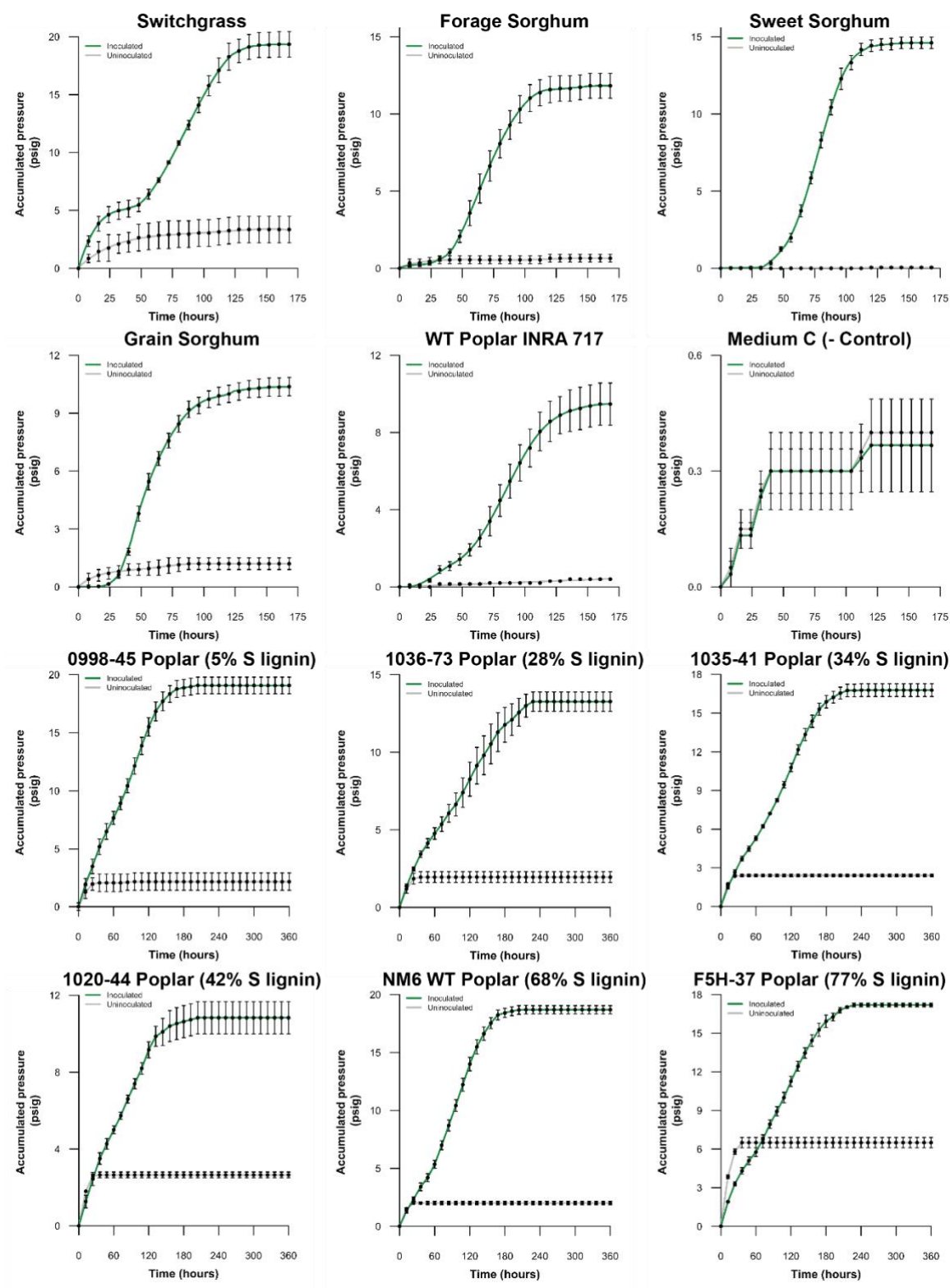
Figure B.5: Growth curves of *Piromyces* sp. UH3-1 on genetically modified lines of poplar [65].
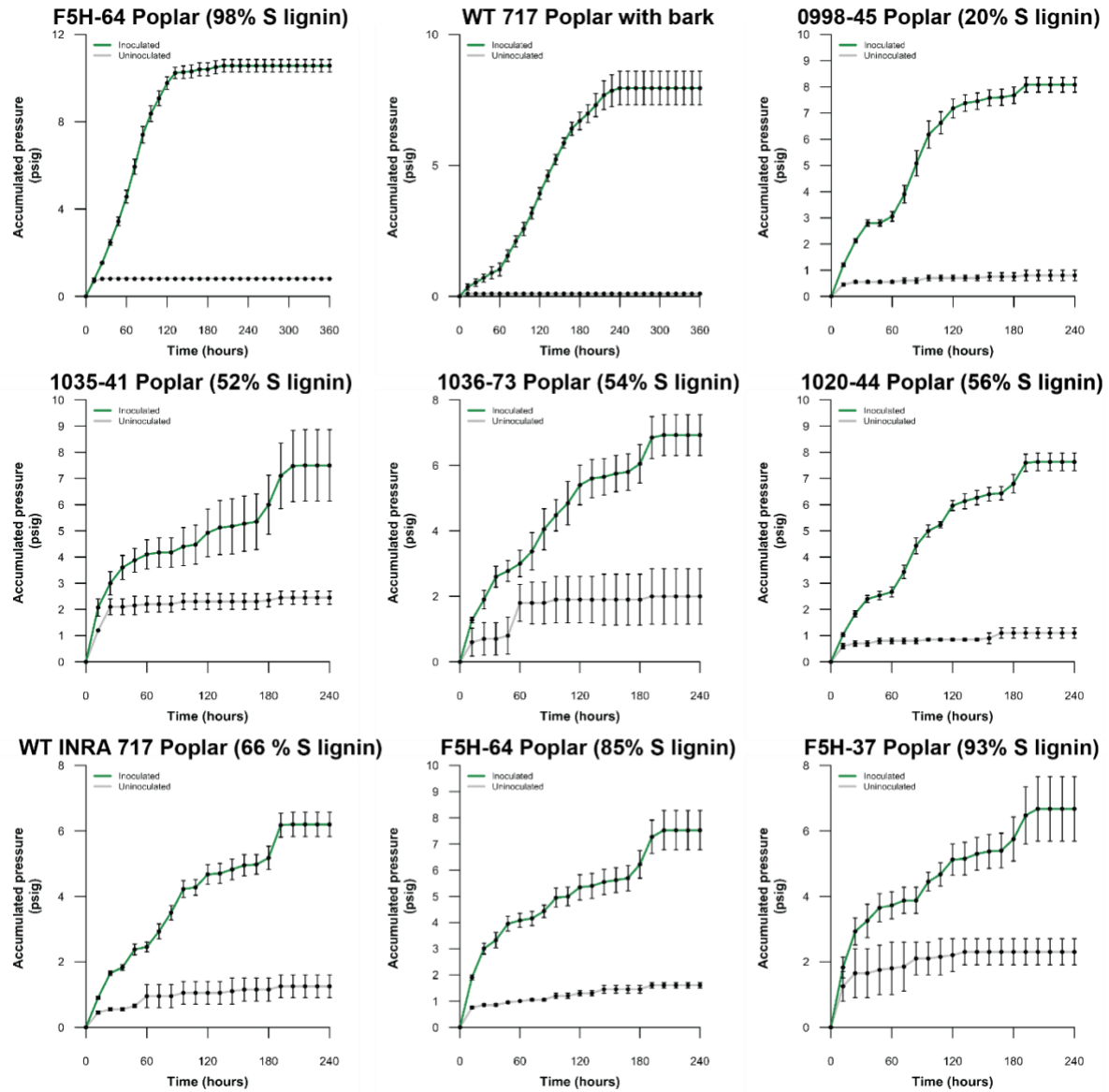
Figure B.6: Growth curves of *Piromyces* sp. UH3-1 of genetically modified lines of poplar from the 2017 harvest [65].
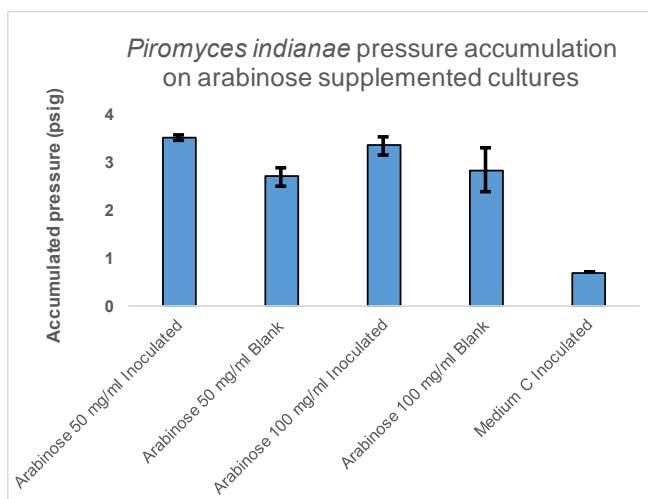
Figure B.7: *Piromyces* sp. UH3-1 fails to reproducibly accumulate pressure on arabinose supplemented cultures regardless of substrate loading.

Figure B.8: *Piromyces* sp. UH3-1 shows visible fungal biomass accumulation on media containing xylose. The top tube was inoculated and shows a high amount of fungal biomass, while the bottom tube was used as a negative control, and was not inoculated [65].
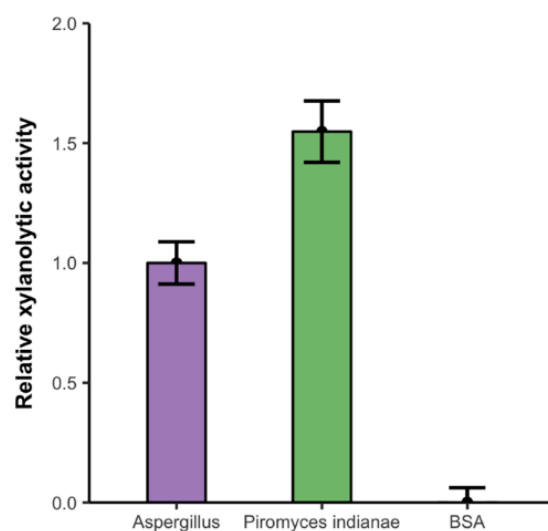
Figure B.9: *Piromyces* sp. UH3-1 shows strong xylanolytic activity on xylan from beechwood at 50° C, pH 7 for six hours of hydrolysis. Values normalized to Viscozyme (*Aspergillus spp.*) [65].
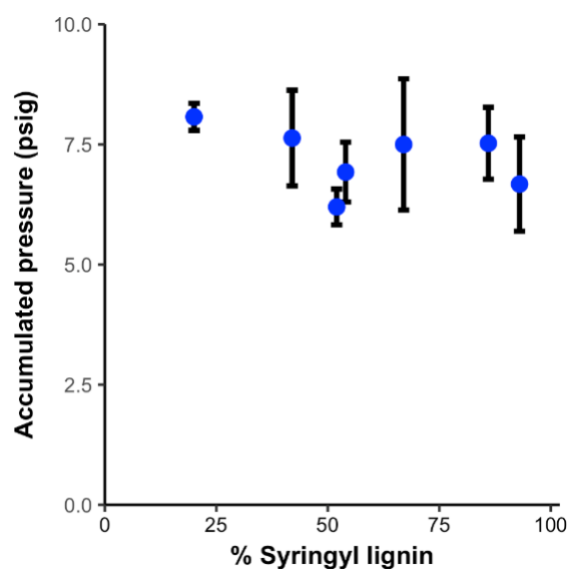
Figure B.10: *Piromyces* sp. UH3-1 degrades untreated poplar to a similar extent on the 2017 harvested poplar having distinct S lignin molar ratios to those of the 2014 harvested lines. $R^2 = 0.008$, $p = 0.647$.
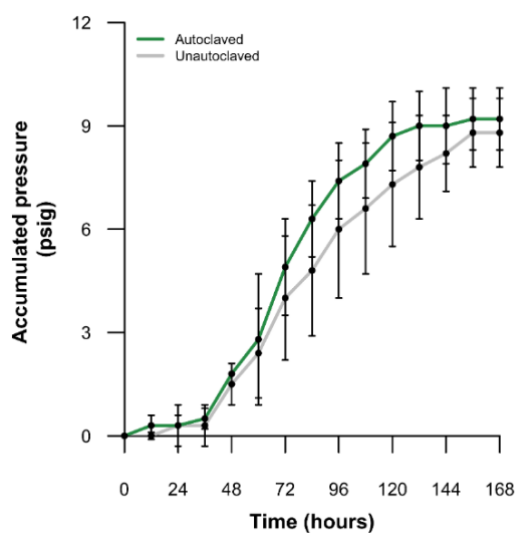
Figure B.11: Autoclaving corn stover at 120 °C for 30 minutes does not significantly enhance fungal growth rate or total accumulated pressure for *Piromyces* sp. UH3-1 This autoclaved corn stover was not washed to remove any potential fermentation inhibitors that would be expected to reduce fungal growth. N=4 [65].

APPENDIX C: EPIGENETICS DATA
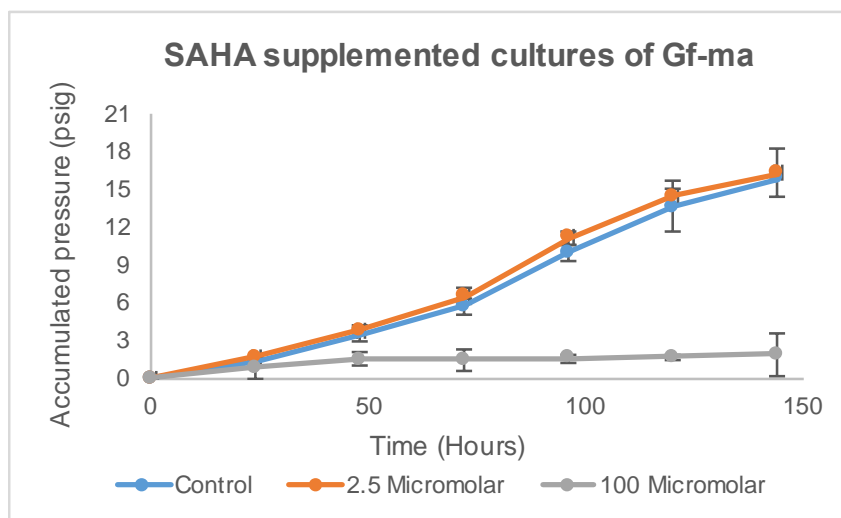
Figure C.1: Growth curves for anaerobic fungal isolate *Neocallimastix* sp. Gf-ma cultured on Medium B with glucose as the substrate with different concentrations of SAHA (See Methods 3.2) N=3.
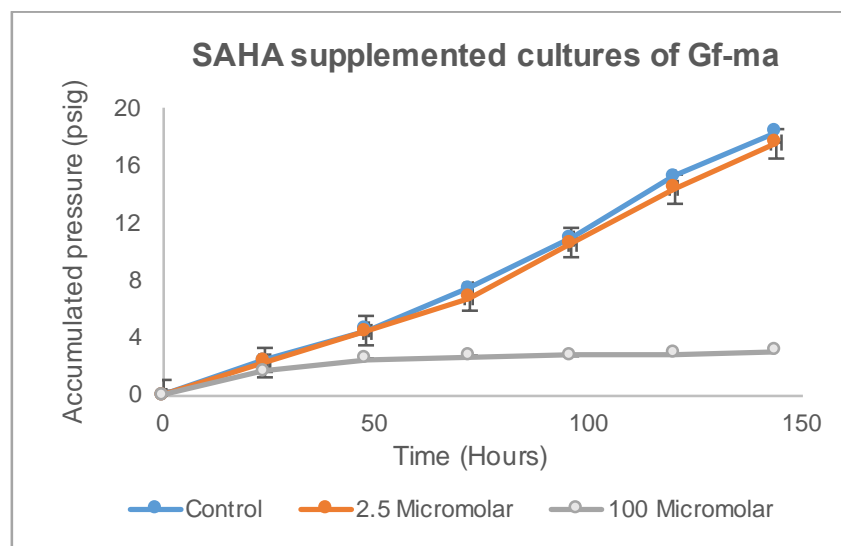
Figure C.2: Growth curve for anaerobic fungal isolate *Neocallimastix* sp. Gf-ma cultured on Medium B with WT INRA 717 poplar as the substrate with different concentrations of SAHA (See Methods 3.2) N=3.
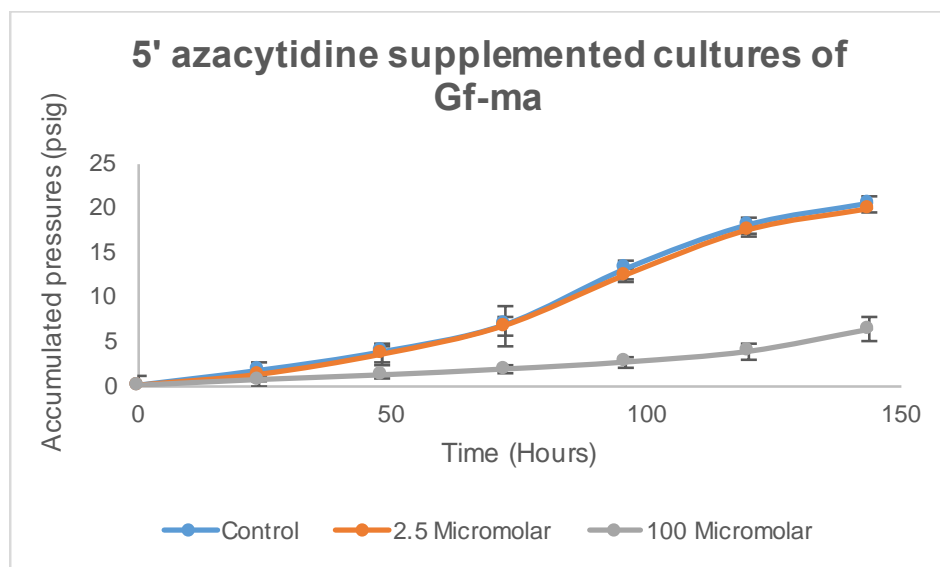
Figure C.3 Growth curve for anaerobic fungal isolate *Neocallimastix* sp. Gf-ma cultured on Medium B with glucose as the substrate with different concentrations of 5' azacytidine (See Methods 3.2) N=3.
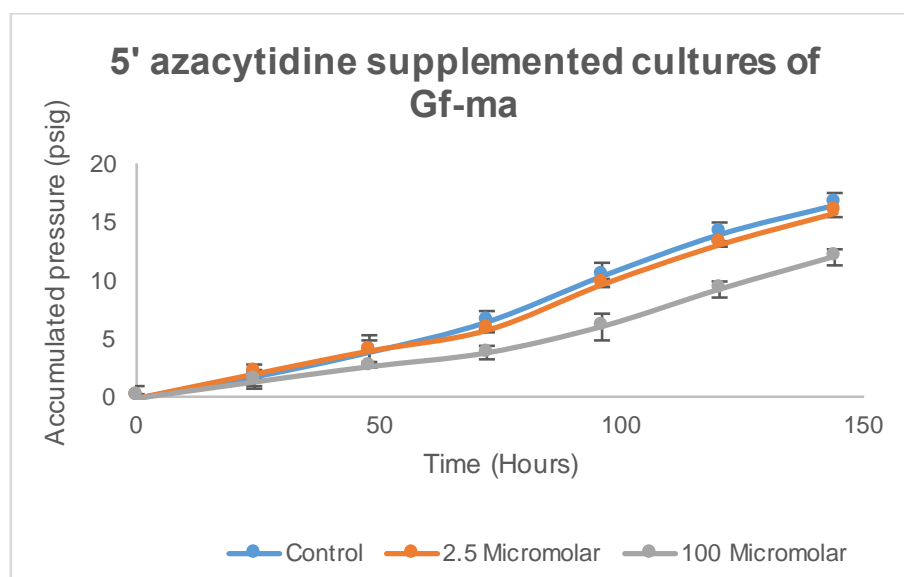
Figure C.4: Growth curve for anaerobic fungal isolate *Neocallimastix* sp. Gf-ma cultured on Medium B WT INRA 717 poplar as the substrate with different concentrations of 5' azacytidine (See Methods 3.2) N=3.

REFERENCES

1. Pauly M, Keegstra K. Cell wall carbohydrates and their modification as a resource for biofuels. *Plant Journal*. 2008;54(4):559-68.

2. Liu Q, Luo L, Zheng L. Lignins: Biosynthesis and biological functions in plants. *International Journal of Molecular Sciences*. 2018;19(2):335.

3. Park S, Baker JO, Himmel ME, Parilla PA, Johnson DK. Cellulose crystallinity index: Measurement techniques and their impact on interpreting cellulase performance. *Biotechnology for Biofuels*. 2010;3(1):10.

4. Zou G, Shi S, Jiang Y, van den Brink J, de Vries RP, Chen L, Zhang J, Ma L, Wang C, Zhou Z. Construction of a cellulase hyper-expression system in Trichoderma reesei by promoter and enzyme engineering. *Microbial Cell Factories*. 2012;11(1):21.

5. Zhang YHP, Lynd LR. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: Noncomplexed cellulase systems. *Biotechnology and Bioengineering*. 2004;88(7):797-824.

6. Mansfield SD, Meder R. Cellulose hydrolysis - The role of monocomponent cellulases in crystalline cellulose degradation. *Cellulose*. 2003;10(2):159-69.

7. Asghari A, Bothast RJ, Doran JB, Ingram LO. Ethanol production from hemicellulose hydrolysates of agricultural residues using genetically engineered Escherichia coli strain KO11. *Journal of Industrial Microbiology and Biotechnology*. 1996;16(1):42-7.

8. Scheller HV, Ulvskov P. Hemicelluloses. *Annual Review of Plant Biology*. 2010;61(1):263-89.

9. Theander O. Cellulose, Hemicellulose and Extractives. In: Overend RP, Milne TA, Mudge LK, editor. Fundamentals of Thermochemical Biomass Conversion. Dordrecht: Springer Netherlands; 1985. p.35-60.

10. Schuetz M, Benske A, Smith RA, Watanabe Y, Tobimatsu Y, Ralph J, Demura T, Ellis B, Samuels AL. Laccases direct lignification in the discrete secondary cell wall domains of protoxylem. *Plant Physiology*. 2014;166(2):798-807.

11. Buranov AU, Mazza G. Lignin in straw of herbaceous crops. *Industrial Crops and Products*. 2008;28(3):237-59.

12. Mäkelä MR, Hildén KS, de Vries RP. Degradation and modification of plant biomass by fungi. In: Nowrousian M, editor. Fungal Genomics. Berlin, Heidelberg: Springer Berlin Heidelberg; 2014. p.175-208.

13. Porter JR, Howell FM, Mason PB, Blanchard TC. Existing biomass infrastructure and theoretical potential biomass production in the U.S. *Journal of Maps*. 2009;5(1):206-18.

14. Zhao X, Zhang L, Liu D. Biomass recalcitrance. Part I: the chemical compositions and physical structures affecting the enzymatic hydrolysis of lignocellulose. *Biofuels, Bioproducts and Biorefining*. 2012;6(4):465-82.

15. Isikgor FH, Becer CR. Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polymer Chemistry*. 2015;6(25):4497-559.

16. Mathur S, Umakanth AV, Tonapi VA, Sharma R, Sharma MK. Sweet sorghum as biofuel feedstock: recent advances and available resources. *Biotechnology for Biofuels*. 2017;10(1):146.

17. Gelfand I, Sahajpal R, Zhang X, Izaurralde RC, Gross KL, Robertson GP. Sustainable bioenergy production from marginal lands in the US Midwest. *Nature*. 2013;493(7433):514.

18. Skyba O, Douglas C, Mansfield S. Syringyl-rich lignin renders poplars more resistant to degradation by wood decay fungi. *Applied and Environmental Microbiology*. 2013;79(8):2560-71.

19. Vertès AA, Inui M, Yukawa H. Implementing biofuels on a global scale. *Nature Biotechnology*. 2006;24(7):761-4.

20. U.S. DOE. Lignocellulosic Biomass for Advanced Biofuels and Bioproducts. U.S. Department of Energy Office of Science. 2015; https://genomicscience.energy.gov/biofuels/lignocellulose/.

21. Buhler DD, Netzer DA, Riemenschneider DE, Hartzler RG. Weed management in short rotation poplar and herbaceous perennial crops grown for biofuel production. *Biomass & Bioenergy*. 1998;14(4):385-94.

22. Allwright MR, Taylor G. Molecular breeding for improved second generation bioenergy crops. *Trends in Plant Science*. 2016;21(1):43-54.

23. Wilson P, Glithero NJ, Ramsden SJ. Prospects for dedicated energy crop production and attitudes towards agricultural straw use: The case of livestock farmers. *Energy Policy*. 2014;74:101-10.

24. Lynd LR, Laser MS, Bransby D, Dale BE, Davison B, Hamilton R, Himmel M, Keller M, McMillan JD, Sheehan J *et al*. How biotech can transform biofuels. *Nature Biotechnology*. 2008;26(2):169-72.

25. Harkki A, Mäntylä A, Penttilä M, Muttilainen S, Bühler R, Suominen P, Knowles J, Nevalainen H. Genetic engineering of *Trichoderma* to produce strains with novel cellulase profiles. *Enzyme and Microbial Technology*. 1991;13(3):227-33.

26. Mohanram S, Amat D, Choudhary J, Arora A, Nain L. Novel perspectives for evolving enzyme cocktails for lignocellulose hydrolysis in biorefineries. *Sustainable Chemical Processes*. 2013;1(1):15.

27. Yeoman CJ, Han Y, Dodd D, Schroeder CM, Mackie RI, Cann IKO. Thermostable enzymes as biocatalysts in the biofuel industry. *Advances in Applied Microbiology*. 2010;70:1-55.

28. Klein-Marcuschamer D, Oleskowicz-Popiel P, Simmons BA, Blanch HW. The challenge of enzyme cost in the production of lignocellulosic biofuels. *Biotechnology and Bioengineering*. 2012;109(4):1083-7.

29. Hirasawa H, Shioya K, Furukawa T, Tani S, Sumitani J-i, Kawaguchi T, Morikawa Y, Shida Y, Ogasawara W. Engineering of the Trichoderma reesei xylanase3 promoter for efficient enzyme expression. *Applied Microbiology and Biotechnology*. 2018;102(6):2737-52.

30. Bischof RH, Ramoni J, Seiboth B. Cellulases and beyond: the first 70 years of the enzyme producer *Trichoderma reesei*. *Microbial Cell Factories*. 2016;15(1):106.

31. de Vries RP, Visser J. Aspergillus enzymes involved in degradation of plant cell wall polysaccharides. *Microbiology and Molecular Biology Reviews*. 2001;65(4):497-522.

32. Druzhinina IS, Kubicek CP. Genetic engineering of Trichoderma reesei cellulases and their production. *Microbial Biotechnology*. 2017;10(6):1485-99.

33. Kawai T, Nakazawa H, Ida N, Okada H, Tani S, Sumitani J-i, Kawaguchi T, Ogasawara W, Morikawa Y, Kobayashi Y. Analysis of the saccharification capability of high-functional cellulase JN11 for various pretreated biomasses through a comparison with commercially available counterparts. *Journal of Industrial Microbiology and Biotechnology*. 2012;39(12):1741-9.

34. Sun Y, Cheng JJ. Dilute acid pretreatment of rye straw and bermudagrass for ethanol production. *Bioresource Technology*. 2005;96(14):1599-606.

35. Palonen H, Tjerneld F, Zacchi G, Tenkanen M. Adsorption of Trichoderma reesei CBH I and EG II and their catalytic domains on steam pretreated softwood and isolated lignin. *Journal of Biotechnology*. 2004;107(1):65-72.

36. Bruce D. A sober view of the difficulties in scaling cellulosic biofuels. *Biofuels, Bioproducts and Biorefining*. 2017;11(1):5-7.

37. Yang M, Zhang W, Rosentrater K. Anhydrous ammonia pretreatment of corn stover and enzymatic hydrolysis of glucan from pretreated corn stover. *Fermentation*. 2017;3(1):9.

38. Mosier N, Wyman C, Dale B, Elander R, Lee YY, Holtzapple M, Ladisch M. Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresource Technology*. 2005;96(6):673-86.

39. Lynd LR. Overview and evaluation of fuel ethanol from cellulosic biomass: Technology, economics, the environment, and policy. *Annual Review of Energy and the Environment*. 1996;21(1):403-65.

40. Holtzapple MT, Jun JH, Ashok G, Patibandla SL, Dale BE. The ammonia freeze explosion (AFEX) process: a practical lignocellulose pretreatment. *Applied Biochemistry and Biotechnology*. 1991;28(1):59-74.

41. Bunterngsook B, Laothanachareon T, Chotirotsukon C, Inoue H, Fujii T, Hoshino T, Roongsawang N, Kuboon S, Kraithong W, Techanan W *et al*. Development of tailor-made synergistic cellulolytic enzyme system for saccharification of steam exploded sugarcane bagasse. *Journal of Bioscience and Bioengineering*. 2018;125(4):390-6.

42. Wang S, Chen H, Tang X, Zhang H, Chen W, Chen YQ. Molecular tools for gene manipulation in filamentous fungi. *Applied Microbiology and Biotechnology*. 2017;101(22):8063-75.

43. Bauchop T. Rumen anaerobic fungi of cattle and sheep. *Applied and Environmental Microbiology*. 1979;38(1):148-58.

44. Youssef NH, Couger MB, Struchtemeyer CG, Liggenstoffer AS, Prade RA, Najar FZ, Atiyeh HK, Wilkins MR, Elshahed MS. The genome of the anaerobic fungus Orpinomyces sp. strain C1A reveals the unique evolutionary history of a remarkable plant biomass degrader. *Applied and Environmental Microbiology*. 2013;79(15):4620-34.

45. Orpin CG. The rumen flagellate Piromonas communis: Its life-history and invasion of plant material in the rumen. *Journal of General Microbiology*. 1977;99(1):107-17.

46. Theodorou MK, Mennim G, Davies DR, Zhu W-Y, Trinci APJ, Brookman JL. Anaerobic fungi in the digestive tract of mammalian herbivores and their potential for exploitation. *Proceedings of The Nutritional Society*. 1996;55(3):913-26.

47. Grigoriev IV, 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 Research*. 2014;42(D1):D699-D704.

48. Akin D, Benner R. Degradation of polysaccharides and lignin by ruminal bacteria and fungi. *Applied and Environment Microbiology*. 1988;54(5):1117-25.

49. Gruninger RJ, Puniya AK, Callaghan TM, Edwards JE, Youssef N, Dagar SS, Fliegerova K, Griffith GW, Forster R, Tsang A *et al*. Anaerobic fungi (phylum Neocallimastigomycota): advances in understanding their taxonomy, life cycle, ecology, role and biotechnological potential. *FEMS Microbiology Ecology*. 2014;90(1):1-17.

50. Orpin CG, Bountiff L. Zoospore chemotaxis in the rumen phycomycete Neocallimastix frontalis. *Microbiology*. 1978;104(1):113-22.

51. Wubah DA, Kim DS. Chemoattraction of anaerobic ruminai fungi zoospores to selected phenolic acids. *Microbiology Research*. 1996;151(3):257-62.

52. Solomon KV, Henske JK, Theodorou MK, Amp, Apos, Malley MA. Robust and effective methodologies for cryopreservation and DNA extraction from anaerobic gut fungi. *Anaerobe*. 2016;38:39-46.

53. Solomon KV, Haitjema CH, Henske JK, Gilmore SP, Borges-Rivera D, Lipzen A, Brewer HM, Purvine SO, Wright AT, Theodorou MK *et al*. Early-branching gut fungi possess a large, comprehensive array of biomass-degrading enzymes. *Science*. 2016;351(6278):1192-5.

54. Haitjema CH, Solomon KV, Henske JK, Theodorou MK, O'Malley MA. Anaerobic gut fungi: Advances in isolation, culture, and cellulolytic enzyme discovery for biofuel production. *Biotechnology and Bioengineering*. 2014;111(8):1471-82.

55. Novotna Z, Prochazka J, Šimůnek J, Fliegerová K. Xylanases of anaerobic fungus Anaeromyces mucronatus. *Folia Microbiologica*. 2010;55(4):363-7.

56. Callaghan TM, Podmirseg SM, Hohlweck D, Edwards JE, Puniya AK, Dagar SS, Griffith GW. Buwchfawromyces eastonii gen. nov., sp. nov.: a new anaerobic fungus (Neocallimastigomycota) isolated from buffalo faeces. *MycoKeys*. 2015;9:11.

57. Orpin C. Studies on the rumen flagellate Sphaeromonas communis. *Microbiology*. 1976;94(2):270-80.

58. Ozkose E, Thomas BJ, Davies DR, Griffith GW, Theodorou MK. Cyllamyces aberensis gen. nov. sp. nov., a new anaerobic gut fungus with branched sporangiophores isolated from cattle. *Canadian Journal of Botany/Revue Canadienne de Botanique*. 2001;79(6):666-73.

59. Orpin CG. Studies on the rumen flagellate Neocallimastix frontalis. *Journal of General Microbiology*. 1975;91(2):249-62.

60. Dagar SS, Kumar S, Griffith GW, Edwards JE, Callaghan TM, Singh R, Nagpal AK, Puniya AK. A new anaerobic fungus (Oontomyces anksri gen. nov., sp. nov.) from the digestive tract of the Indian camel (Camelus dromedarius). *Fungal Biology*. 2015;119(8):731-7.

61. Borneman W, Akin D, Ljungdahl L. Fermentation products and plant cell wall-degrading enzymes produced by monocentric and polycentric anaerobic ruminal fungi. *Applied and Environmental Microbiology*. 1989;55(5):1066-73.

62. Hanafy RA, Elshahed MS, Liggenstoffer AS, Griffith GW, Youssef NH. Pecoramyces ruminantium, gen. nov., sp. nov., an anaerobic gut fungus from the feces of cattle and sheep. *Mycologia*. 2017;109(2):231-43.

63. Orpin CG. Isolation of cellulolytic phycomycete fungi from the caecum of the horse. *Journal of General Microbiology*. 1981;123(2):287-96.

64. Haitjema CH, Gilmore SP, Henske JK, Solomon KV, de Groot R, Kuo A, Mondo SJ, Salamov AA, LaButti K, Zhao Z *et al*. A parts list for fungal cellulosomes revealed by comparative genomics. *Nature Microbiology*. 2017;2:17087.

65. Hooker CA, Hillman ET, Overton JC, Ortiz-Velez A, Schacht M, Hunnicutt A, Mosier NS, Solomon KV. Hydrolysis of untreated lignocellulosic feedstock is independent of S-lignin composition in newly classified anaerobic fungal isolate, Piromyces sp. UH3-1. *Biotechnology for Biofuels*. 2018;11(1):293.

66. Rubin EM. Genomics of cellulosic biofuels. *Nature*. 2008;454(7206):841-5.

67. Dondelinger E, Aubry N, Ben Chaabane F, Cohen C, Tayeb J, Rémond C. Contrasted enzymatic cocktails reveal the importance of cellulases and hemicellulases activity ratios for the hydrolysis of cellulose in presence of xylans. *AMB Express*. 2016;6(1):24.

68. Schülein M. Cellulases of *Trichoderma reesei*. In: Methods in Enzymology. 1 edn: Academic Press; 1988. p.234-42.

69. Nakazawa H, Kawai T, Ida N, Shida Y, Kobayashi Y, Okada H, Tani S, Sumitani J-i, Kawaguchi T, Morikawa Y *et al*. Construction of a recombinant *Trichoderma reesei* strain expressing *Aspergillus aculeatus* β-glucosidase 1 for efficient biomass conversion. *Biotechnology and Bioengineering*. 2012;109(1):92-9.

70. Nicholson MJ, Theodorou MK, Brookman JL. Molecular analysis of the anaerobic rumen fungus *Orpinomyces* – insights into an AT-rich genome. *Microbiology*. 2005;151(1):121-33.

71. Lowe SE, Griffith GG, Milne A, Theodorou MK, Trinci APJ. Life cycle and growth kinetics of an anaerobic rumen fungus. *Microbiology*. 1987;133(7):1815-27.

72. Meilan R, Ma C. Poplar (Populus spp.). In: Wang K, editor. Agrobacterium Protocols Volume 2. Totowa, New Jersey: Humana Press; 2007. p.143-51.

73. Anderson NA, Tobimatsu Y, Ciesielski PN, Ximenes E, Ralph J, Donohoe BS, Ladisch M, Chapple C. Manipulation of guaiacyl and syringyl monomer biosynthesis in an *Arabidopsis* cinnamyl alcohol dehydrogenase mutant results in atypical lignin biosynthesis and modified cell wall structure. *Plant Cell*. 2015;27(8).

74. Franke R, McMichael CM, Meyer K, Shirley AM, Cusumano JC, Chapple C. Modified lignin in tobacco and poplar plants over-expressing the *Arabidopsis* gene encoding ferulate 5-hydroxylase. *Plant Journal*. 2000;22(3):223-34.

75. Overend RP, Chornet E. Fractionation of lignocellulosics by steam-aqueous pretreatments. *Philosophical Transactions of the Royal Society of London, Series A*. 1987;321(1561):523-36.

76. Jönsson LJ, Martín C. Pretreatment of lignocellulose: Formation of inhibitory by-products and strategies for minimizing their effects. *Bioresource Technology*. 2016;199:103-12.

77. Cao G, Ximenes E, Nichols NN, Frazer SE, Kim D, Cotta MA, Ladisch M. Bioabatement with hemicellulase supplementation to reduce enzymatic hydrolysis inhibitors. *Bioresource Technology*. 2015;190:412-5.

78. Tuckwell DS, Nicholson MJ, McSweeney CS, Theodorou MK, Brookman JL. The rapid assignment of ruminal fungi to presumptive genera using ITS1 and ITS2 RNA secondary structures to produce group-specific fingerprints. *Microbiology*. 2005;151(Pt 5):1557.

79. Dagar SS, Kumar S, Mudgil P, Singh R, Puniya AK. D1/D2 domain of large-subunit ribosomal DNA for differentiation of *Orpinomyces spp*. *Applied and Environmental Microbiology*. 2011;77(18):6722-5.

80. Theodorou MK, Williams BA, Dhanoa MS, McAllan AB, France J. A simple gas production method using a pressure transducer to determine the fermentation kinetics of ruminant feeds. *Animal Feed Science and Technology*. 1994;48(3):185-97.

81. Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*. 1976;72(1-2):248-54.

82. Tseng M-J, Yap M-N, Ratanakhanokchai K, Kyu KL, Chen S-T. Purification and characterization of two cellulase free xylanases from an alkaliphilic *Bacillus firmus*. *Enzyme and Microbial Technology*. 2002;30(5):590-5.

83. Xiao Z, Storms R, Tsang A. Microplate-based filter paper assay to measure total cellulase activity. *Biotechnology and Bioengineering*. 2004;88(7):832-7.

84. Hames B, Ruiz R, Scarlata C, Sluiter A, Sluiter J, Templeton D. Preparation of Samples for Compositional Analysis. National Renewable Energy Laboratory. 2008; https://www.nrel.gov/docs/gen/fy08/42620.pdf.

85. Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D, Crocker D. Determination of Structural Carbohydrates and Lignin in Biomass. National Renewable Energy Laboratory. 2008; https://www.nrel.gov/docs/gen/fy13/42618.pdf.

86. Sluiter A, Ruiz R, Scarlata C, Sluiter J, Templeton D. Determination of Extractives in Biomass. National Renewable Energy Laboratory. 2008; https://www.nrel.gov/docs/gen/fy08/42619.pdf.

87. Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding C, Fungal Barcoding Consortium Author L, Bolchacova E *et al*. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences, USA*. 2012;109(16):6241-6.

88. Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G, Parada A, Gilbert JA, Jansson JK, Caporaso JG, Fuhrman JA *et al*. Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems*. 2016;1(1):e00009-15.

89. Liggenstoffer AS, Youssef NH, Couger MB, Elshahed MS. Phylogenetic diversity and community structure of anaerobic gut fungi (phylum Neocallimastigomycota) in ruminant and non-ruminant herbivores. *ISME J*. 2010;4(10):1225-35.

90. Dollhofer V, Podmirseg SM, Callaghan TM, Griffith GW, Fliegerová K. Anaerobic Fungi and Their Potential for Biogas Production. In: Güebitz GM, Bauer A, Bochmann G, Gronauer A, S W, editor. Biogas Science and Technology. Cham Switzerland: Springer International Publishing; 2015. p.41-61.

91. Juhász T, Egyházi A, Réczey K. β-Glucosidase production by *Trichoderma reesei*. *Applied Biochemistry and Biotechnology*. 2005;121(1):243-54.

92. Mikkelsen D, Flanagan BM, Wilson SM, Bacic A, Gidley MJ. Interactions of arabinoxylan and (1,3)(1,4)-β-glucan with cellulose networks. *Biomacromolecules*. 2015;16(4):1232.

93. Edwards MC, Doran-Peterson J. Pectin-rich biomass as feedstock for fuel ethanol production. *Applied Microbiology and Biotechnology*. 2012;95(3):565-75.

94. Mohnen D. Pectins and their manipulation. In: Knox GSaJ, editor. Biosynthesis of Pectins. Oxford: Blackwell Publishing, Ltd; 2002.

95. Mohnen D. Pectin structure and biosynthesis. *Current Opinion in Plant Biology*. 2008;11(3):266-77.

96. Teather RM, Wood PJ. Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *Applied and Environmental Microbiology*. 1982;43(4):777-80.

97. Kumar D, Murthy GS. Impact of pretreatment and downstream processing technologies on economics and energy in cellulosic ethanol production. *Biotechnology for Biofuels*. 2011;4(1):27.

98. Hall M, Bansal P, Lee JH, Realff MJ, Bommarius AS. Cellulose crystallinity a key predictor of the enzymatic hydrolysis rate. *FEBS Journal*. 2010;277(6):1571-82.

99. Peciulyte A, Karlström K, Larsson PT, Olsson L. Impact of the supramolecular structure of cellulose on the efficiency of enzymatic hydrolysis. *Biotechnology for Biofuels*. 2015;8(1):56.

100. Barbehenn R, Bernays E. Relative nutritional quality of C 3 and C 4 grasses for a graminivorous lepidopteran, *Paratrytone melane* (Hesperiidae). *Oecologia*. 1992;92(1):97-103.

101. Rao PS, Kumar CG, Prakasham RS, Rao AU, Reddy BVS. Sweet Sorghum: Breeding and Bioproducts. In: Cruz VMV, Dierig DA, editor. Industrial Crops: Breeding for BioEnergy and Bioproducts. New York, NY: Springer New York; 2015. p.1-28.

102. Parrish DJ, Fike JH. The biology and agronomy of switchgrass for biofuels. *Critical Reviews in Plant Sciences*. 2005;24(5-6):423-59.

103. Littlewood J, Guo M, Boerjan W, Murphy RJ. Bioethanol from poplar: a commercially viable alternative to fossil fuel in the European Union. *Biotechnology for Biofuels*. 2014;7(1):113.

104. Balatinecz JJ KD. Properties and utilization of poplar wood. Ontario, Canada: Canadian Science Publishing NRC Research Press; 2001.

105. Wiselogel A, Tyson S, Johnson D. Biomass feedstock resources and composition. In: Wyman C, editor. Handbook on bioethanol: Production and utilization. London: Taylor and Francis; 1996.

106. Kim Y, Mosier NS, Ladisch MR. Enzymatic digestion of liquid hot water pretreated hybrid poplar. *Biotechnology Progress*. 2009;25(2):340-8.

107. Christersson L. Wood production potential in poplar plantations in Sweden. *Biomass & Bioenergy*. 2010;34(9):1289-99.

108. Jørgensen H, Kristensen JB, Felby C. Enzymatic conversion of lignocellulose into fermentable sugars: challenges and opportunities. *Biofuels, Bioproducts and Biorefining*. 2007;1(2):119-34.

109. Joblin KN, Naylor GE. Fermentation of woods by rumen anaerobic fungi. *FEMS Microbiology Letters*. 1989;65(1):119-22.

110. Kim TH, Kim JS, Sunwoo C, Lee YY. Pretreatment of corn stover by aqueous ammonia. *Bioresource Technology*. 2003;90(1):39-47.

111. Yang B, Wyman CE. Pretreatment: the key to unlocking low-cost cellulosic ethanol. *Biofuels, Bioproducts and Biorefining*. 2008;2(1):26-40.

112. Kim SM, Dien BS, Tumbleson ME, Rausch KD, Singh V. Improvement of sugar yields from corn stover using sequential hot water pretreatment and disk milling. *Bioresource Technology*. 2016;216:706-13.

113. Henske JK, Wilken SE, Solomon KV, Smallwood CR, Shutthanandan V, Evans JE, Theodorou MK, O'Malley MA. Metabolic characterization of anaerobic fungi provides a path forward for bioprocessing of crude lignocellulose. *Biotechnology and Bioengineering*. 2018;115(4):874-84.

114. Rodrigues AC, Haven MØ, Lindedam J, Felby C, Gama M. Celluclast and Cellic® CTec2: Saccharification/fermentation of wheat straw, solid–liquid partition and potential of enzyme recycling by alkaline washing. *Enzyme and Microbial Technology*. 2015;79-80:70-7.

115. Hu J, Arantes V, Saddler JN. The enhancement of enzymatic hydrolysis of lignocellulosic substrates by the addition of accessory enzymes such as xylanase: is it an additive or synergistic effect? *Biotechnology for Biofuels*. 2011;4(1):36-.

116. Chahal DS, McGuire S, Pikor H, Noble G. Production of cellulase complex by Trichoderma reesei rut-C30 on lignocellulose and its hydrolytic potential. *Biomass*. 1982;2(2):127-37.

117. Montenecourt BS, Eveleigh DE. Selective screening methods for the isolation of high yielding cellulase mutants of *Trichoderma reesei*. In: Hydrolysis of Cellulose: Mechanisms of Enzymatic and Acid Catalysis. Am Chem Soc; 1979. p.289-301.

118. Nevalainen KM. Induction, isolation, and characterization of *Aspergillus niger* mutant strains producing elevated levels of beta-galactosidase. *Applied and Environmental Microbiology*. 1981;41(3):593-6.

119. Primrose SB, Twyman R. Principles of gene manipulation and genomics: John Wiley & Sons; 2013.

120. Holoch D, Moazed D. RNA-mediated epigenetic regulation of gene expression. *Nature Reviews: Genetics*. 2015;16(2):71-84.

121. Seiboth B, Karimi RA, Phatale PA, Linke R, Hartl L, Sauer DG, Smith KM, Baker SE, Freitag M, Kubicek CP. The putative protein methyltransferase LAE1 controls cellulase gene expression in Trichoderma reesei. *Molecular Microbiology*. 2012;84(6):1150-64.

122. Xin Q, Gong Y, Lv X, Chen G, Liu W. Trichoderma reesei histone acetyltransferase Gcn5 regulates fungal growth, conidiation, and cellulase gene expression. *Current Microbiology*. 2013;67(5):580-9.

123. Soyer JL, El Ghalid M, Glaser N, Ollivier B, Linglin J, Grandaubert J, Balesdent M-H, Connolly LR, Freitag M, Rouxel T *et al*. Epigenetic control of effector gene expression in the plant pathogenic fungus Leptosphaeria maculans. *PLoS Genetics*. 2014;10(3):e1004227.

124. Breen J, Mur L, Sivakumaran A, Akinyemi A, Wilkinson M, Rodriguez Lopez CM. Botrytis cinerea loss and restoration of virulence during in vitro culture follows flux in global DNA methylation. *bioRxiv*. 2016.

125. Mondo SJ, Dannebaum RO, Kuo RC, Louie KB, Bewick AJ, LaButti K, Haridas S, Kuo A, Salamov A, Ahrendt SR *et al*. Widespread adenine N6-methylation of active genes in fungi. *Nature Genetics*. 2017;49(6):964-8.

126. Freitag M. Histone methylation by SET domain proteins in fungi. *Annual Review of Microbiology*. 2017;71(1):413-39.

127. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Research*. 2011;21(3):381-95.

128. Jenuwein T, Allis CD. Translating the histone code. *Science*. 2001;293(5532):1074-80.

129. Gary JD, Lin W-J, Yang MC, Herschman HR, Clarke S. The predominant protein-arginine methyltransferase from Saccharomyces cerevisiae. *Journal of Biological Chemistry*. 1996;271(21):12585-94.

130. da Silva Delabona P, Rodrigues GN, Zubieta MP, Ramoni J, Codima CA, Lima DJ, Farinas CS, da Cruz Pradella JG, Seiboth B. The relation between xyr1 overexpression in Trichoderma harzianum and sugarcane bagasse saccharification performance. *Journal of Biotechnology*. 2017;246:24-32.

131. Weber CM, Henikoff S. Histone variants: dynamic punctuation in transcription. *Genes & Development*. 2014;28(7):672-82.

132. Lowe SE. Growth of anaerobic rumen fungi on defined and semi-defined media lacking rumen fluid. *Journal of General Microbiology*. 1985;131(9):2225-30.

133. Christman JK. 5-Azacytidine and 5-aza-2′-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene*. 2002;21(35):5483.

134. Richon VM, Emiliani S, Verdin E, Webb Y, Breslow R, Rifkind RA, Marks PA. A class of hybrid polar inducers of transformed cell differentiation inhibits histone deacetylases. *Proceedings of the National Academy of Sciences, USA*. 1998;95(6):3003-7.

135. Berni Canani R, Di Costanzo M, Leone L. The epigenetic effects of butyrate: potential therapeutic implications for clinical practice. *Clinical Epigenetics*. 2012;4(1):4.

136. Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, Klenk DC. Measurement of protein using bicinchoninic acid. *Analytical Biochemistry*. 1985;150(1):76-85.

137. Hockly E, Richon VM, Woodman B, Smith DL, Zhou X, Rosa E, Sathasivam K, Ghazi-Noori S, Mahal A, Lowden PAS *et al*. Suberoylanilide hydroxamic acid, a histone deacetylase inhibitor, ameliorates motor deficits in a mouse model of Huntington's disease. *Proceedings of the National Academy of Sciences, USA*. 2003;100(4):2041-6.

138. Verma MS, Fink MJ, Salmon GL, Fornelos N, Ohara TE, Ryu SH, Vlamakis H, Xavier RJ, Stappenbeck TS, Whitesides GM. A common mechanism links activities of butyrate in the colon. *ACS Chemical Biology*. 2018;13(5):1291-8.

139. Li X-L, Chen H, Ljungdahl LG. Monocentric and polycentric anaerobic fungi produce structurally related cellulases and xylanases. *Applied and Environmental Microbiology*. 1997;63(2):628-35.

140. Frandsen KEH, Simmons TJ, Dupree P, Poulsen J-CN, Hemsworth GR, Ciano L, Johnston EM, Tovborg M, Johansen KS, von Freiesleben P *et al*. The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases. *Nature Chemical Biology*. 2016;12(4):298.