

# UC Santa Barbara

## UC Santa Barbara Previously Published Works

### Title

Co-cultivation of the anaerobic fungus *Caecomyces churrovis* with *Methanobacterium bryantii* enhances transcription of carbohydrate binding modules, dockerins, and pyruvate formate lyases on specific substrates

### Permalink

<https://escholarship.org/uc/item/Or62m212>

### Journal

Biotechnology for Biofuels and Bioproducts, 14(1)

### ISSN

2731-3654

### Authors

Brown, Jennifer L  
Swift, Candice L  
Mondo, Stephen J  
et al.

### Publication Date

2021

### DOI

10.1186/s13068-021-02083-w


Peer reviewed

RESEARCH

Open Access



# Co-cultivation of the anaerobic fungus *Caecomyces churrovis* with *Methanobacterium bryantii* enhances transcription of carbohydrate binding modules, dockerins, and pyruvate formate lyases on specific substrates

Jennifer L. Brown<sup>1</sup>, Candice L. Swift<sup>1</sup>, Stephen J. Mondo<sup>2,3</sup>, Susanna Seppala<sup>1</sup>, Asaf Salamov<sup>2</sup>, Vasanth Singan<sup>2</sup>, Bernard Henrissat<sup>4,5</sup>, Elodie Drula<sup>6,7</sup>, John K. Henske<sup>1</sup>, Samantha Lee<sup>1</sup>, Kurt LaButti<sup>2</sup>, Guifen He<sup>2</sup>, Mi Yan<sup>2</sup>, Kerrie Barry<sup>2</sup>, Igor V. Grigoriev<sup>2,8</sup> and Michelle A. O'Malley<sup>1,9\*</sup> 

## Abstract

Anaerobic fungi and methanogenic archaea are two classes of microorganisms found in the rumen microbiome that metabolically interact during lignocellulose breakdown. Here, stable synthetic co-cultures of the anaerobic fungus *Caecomyces churrovis* and the methanogen *Methanobacterium bryantii* (not native to the rumen) were formed, demonstrating that microbes from different environments can be paired based on metabolic ties. Transcriptional and metabolic changes induced by methanogen co-culture were evaluated in *C. churrovis* across a variety of substrates to identify mechanisms that impact biomass breakdown and sugar uptake. A high-quality genome of *C. churrovis* was obtained and annotated, which is the first sequenced genome of a non-rhizoid-forming anaerobic fungus. *C. churrovis* possess an abundance of CAZymes and carbohydrate binding modules and, in agreement with previous studies of early-diverging fungal lineages, N6-methyldeoxyadenine (6mA) was associated with transcriptionally active genes. Co-culture with the methanogen increased overall transcription of CAZymes, carbohydrate binding modules, and dockerin domains in co-cultures grown on both lignocellulose and cellulose and caused upregulation of genes coding associated enzymatic machinery including carbohydrate binding modules in family 18 and dockerin domains across multiple growth substrates relative to *C. churrovis* monoculture. Two other fungal strains grown on a reed canary grass substrate in co-culture with the same methanogen also exhibited high log<sub>2</sub>-fold change values for upregulation of genes encoding carbohydrate binding modules in families 1 and 18. Transcriptional upregulation indicated that co-culture of the *C. churrovis* strain with a methanogen may enhance pyruvate formate lyase (PFL) function for growth on xylan and fructose and production of bottleneck enzymes in sugar utilization pathways, further supporting the hypothesis that co-culture with a methanogen may enhance certain fungal metabolic functions. Upregulation of CBM18 may play a role in fungal–methanogen physical associations and fungal cell wall development and remodeling.

\*Correspondence: momalley@ucsb.edu

<sup>1</sup> Department of Chemical Engineering, University of California Santa Barbara, Santa Barbara, CA 93106, USA

Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Keywords:** Anaerobic fungi, Methanogen, Metabolism, Genome, RNA-Seq, Consortia, CAZymes

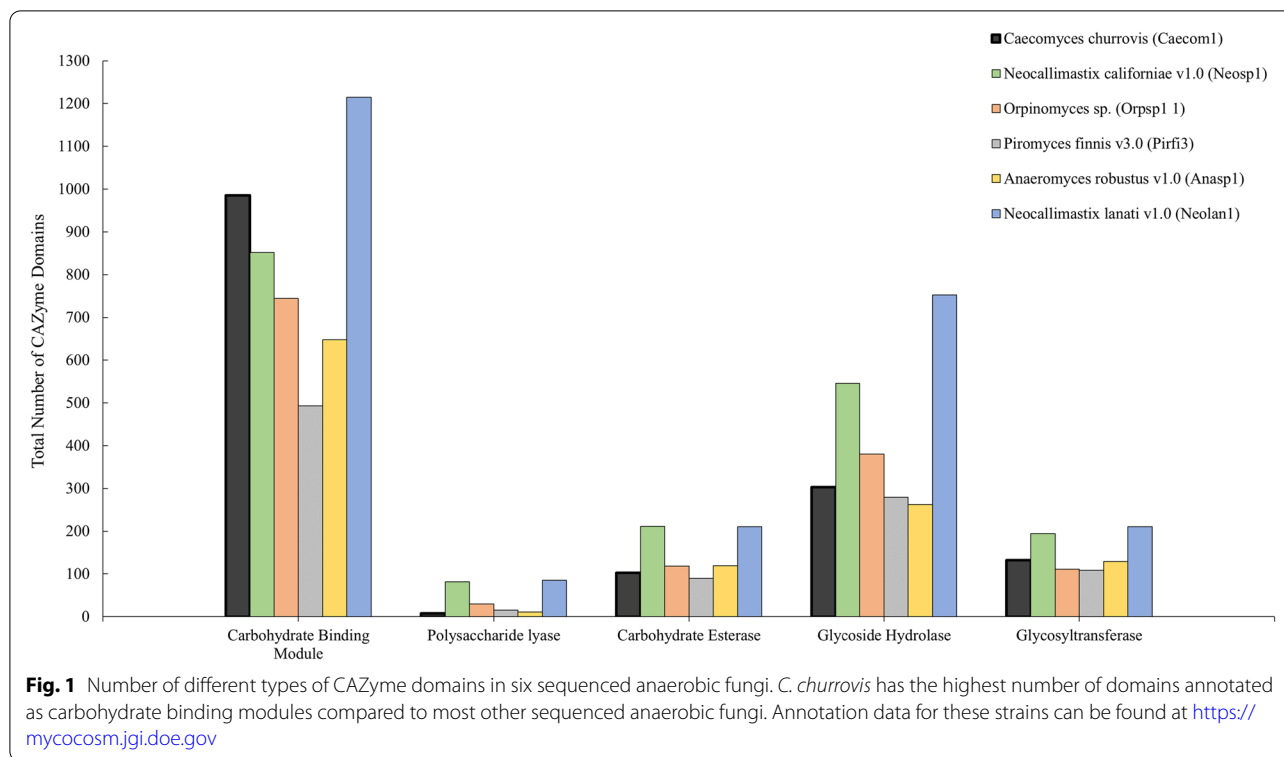
## Introduction

Anaerobic fungi are efficient degraders of recalcitrant lignocellulosic biomass that are found in the guts of herbivores. The high number of CAZymes (carbohydrate active enzymes) that anaerobic fungi produce has driven efforts to collect genomic and transcriptomic data for a variety of emerging anaerobic fungal species, with a focus on the differential transcriptional response of anaerobic fungi to complex carbohydrates versus monomeric sugars [1–6]. Gut fungi function within a community of biomass-degrading bacteria, protozoa, and methanogenic archaea linked by complex metabolic interactions and functional redundancy [7]. Isolating individual members of these natural consortia is one approach to develop a more detailed understanding of microbial interactions, which can then be used to design optimized consortia for biotechnological applications to break down lignocellulose-rich waste. These microbes can be selected through “top-down” enrichment techniques such as serial cultivation or antibiotic treatment to isolate syntrophic pairs of fungi and methanogens from naturally occurring consortia. Alternatively, communities can be formed using “bottom up” methods mixing separate axenic cultures of these microbes to create synthetic pairings linked by metabolic dependency [7–9].

Fungal-methanogen co-cultures have been extensively studied due to the mutually beneficial relationship between the two organisms resulting from their complementary metabolism—fungi produce hydrogen ( $H_2$ ) as an unwelcome byproduct of their own metabolism, which methanogens use in the biosynthesis and release of methane [9–15]. Many previous studies report that co-cultivation of anaerobic fungi with methanogens can enhance biomass breakdown, but the metabolic mechanisms responsible for this outcome are unclear and not uniformly reproducible [14, 16–19]. For example, a recent study concluded that the removal of fungal metabolites by methanogens does not increase the rate of gas production or the rate of substrate deconstruction by a synthetic community of fungi and methanogens relative to fungal monocultures [9]. It has also been hypothesized that co-cultivation of fungi and methanogens results in increased sugar utilization and flux through the fungal hydrogenosome through increased transport and carbon conversion [15, 20]. Additionally, we recently reported that *M. bryantii* enhances the transcription of genes encoding ABC transporters, MFS transporters and G-protein coupled receptors (GPCRs) in the fungus *Anaeromyces robustus*, indicating that co-cultivation may increase the rate

of sugar utilization through the increased expression of sugar transporters [10]. Although many studies have been conducted to determine how co-cultivation with methanogens affects fungal metabolism and biomass breakdown, none have characterized transcriptional and metabolic outcomes across a variety of relevant substrates, which is critical to detangling competing effects of substrate response [10, 11].

Here, we present the first genome of an anaerobic non-rhizoid-forming fungus of the *Caecomyces* genus, and further examine its transcriptional response to the presence of methanogens in multiple synthetic co-cultures supported on lignocellulose, hemicellulose, cellulose, and sugars. *Caecomyces churrovius* lacks the extensive rhizoid network formed by other previously sequenced anaerobic gut fungi to aid in biomass breakdown. Improvements in long-read sequencing technologies enabled assembly and annotation of CAZymes and associated cellular machinery despite the complex fungal physiology, unknown ploidy, AT-content, and repeat-richness. By combining RNA-seq with growth and chemical data, we determine how the fungus responds to co-cultivation with a non-native methanogen in synthetic co-culture. The ability to pair two microbes based on complementary metabolism alone presents the opportunity to combine non-native microbes in a desired technological application without the constraint of naturally developed syntrophy. While other studies have examined global transcriptomic response and CAZyme regulation in anaerobic fungi cultivated with methanogens on a single substrate, none to date have explored regulation across a range of substrates or differences occurring in transcriptional regulation between multiple fungal strains on the same substrate [10, 11]. Through a combination of genomic, transcriptomic, and metabolomic data we found that the *C. churrovius* genome possesses an abundance of both CAZymes and carbohydrate binding modules as shown in Fig. 1. Co-culture of *C. churrovius* with a non-native methanogen enhances transcription of gene sets associated with fungal substrate binding and fungal–methanogen interactions such as carbohydrate binding modules in families 1 and 18, pyruvate formate lyase (PFL) function in the cytosol or possibly the hydrogenosome, and enzymes that are potential bottlenecks for sugar utilization in fungi across multiple substrates. Overall, understanding how methanogen co-culture influences the fibrolytic and metabolic behavior of anaerobic fungi aids in the design of new strategies for conversion of lignocellulose to fermentable sugars and value-added products, and reveals



the genetic mechanisms that underpin fungal–methanogen interactions.

## Results and discussion

### The *Caecomyces churrovis* genome encodes an abundance of CAZymes and carbohydrate binding modules

Anaerobic fungi are emerging platforms for hydrolysis of crude lignocellulose, as they produce powerful CAZymes and mechanically associate with and often penetrate plant cell walls [5, 21, 22]. The first high-quality genome of a non-rhizoid-forming anaerobic fungus from the *Caecomyces* genera was sequenced with PacBio SMRT sequencing using high molecular weight DNA fragments, a method that is critical to high-quality genome assemblies for anaerobic fungi [2, 23, 24]. Previously, we assembled a de novo transcriptome of *C. churrovis* by pooling RNA from batch cultures grown on glucose, fructose, cellobiose, cellulose, and reed canary grass, obtaining an inclusive set of expressed genes for these substrates [5]. The acquisition of the *C. churrovis* genome now enables more detailed investigation of genetic regulatory mechanisms, splicing, ploidy, and comparative genomics that cannot be accomplished with a sole transcriptome. Based on genome sequencing, 15,009 genes were annotated/identified, compared to the predicted 33,437 genes based on the sequenced transcriptome (predicted by taking into account the number of transcripts less isoforms);

this difference in gene number prediction between transcriptomes and genomes is consistent across anaerobic fungi and likely reflective of ploidy [2, 5]. This discrepancy is largely explained by our observation that this strain of *Caecomyces* is likely a diploid (or dikaryon), as we detected ~10k gene models on smaller scaffolds in regions that were >95% identical to regions on larger scaffolds. These scaffolds were designated as secondary scaffolds and these secondary models/alleles were not included in further analyses but are available from MycoCosm [25]. Table 1 depicts genomic features for high-resolution sequenced anaerobic fungi, as reported by the JGI MycoCosm pipeline [25].

As noted in Table 1, the *C. churrovis* genome is GC depleted on the same order of magnitude as the other sequenced anaerobic fungal strains. Such extreme codon biases have made it challenging to heterologously express and evaluate the function of anaerobic fungal genes (like CAZymes) in model systems [26–28]. Homopolymeric runs of amino acids found in the *C. churrovis* genome are common in the CAZyme machinery of anaerobic fungi and could serve as glycosylation sites that prevent proteolytic cleavage [28]. Collectively, the function of such features needs to be better characterized if gut fungal CAZymes from strains such as *C. churrovis* are to be heterologously produced in a model organism [28].

**Table 1** Overview of sequenced anaerobic fungal genome features and statistics<sup>2–4</sup>

	<i>Caecomyces churrovius</i>	<i>Anaeromyces robustus</i>	<i>Neocallimastix californiae</i>	<i>Neocallimastix lanati</i>	<i>Piromyces finnis</i>	<i>Pecoramyces ruminantium*</i>
Genome size (Mbp)	165.50	71.69	193.03	200.97	56.46	100.95
No. scaffolds	7737	1035	1801	970	232	32,574
% GC content	19	16	22	18	21	17
Scaffold L50 (Mbp)	0.03	0.14	0.44	1.03	0.75	0.00
No. of gene models	15,009	12,832	20,219	25,350	10,992	18,936
Gene % CAZymes	7.22	6.73	7.23	7.05	6.45	5.67
No. of DDPs <sup>a</sup>	389	276	422	586	227	318
No. of scaffoldins	36	26	55	93	14	83
No. of diploid gene pairs	10,972	147	1154	497	146	3113

\*Formerly named *Orpinomyces* sp. strain C1A

<sup>a</sup> dockerin domain proteins

Anaerobic gut fungi possess an abundance of CAZymes with diverse functions, and are particularly rich in hemicellulases (especially glycosyl hydrolase 10 family) and polysaccharide deacetylases [22]. Some CAZymes are anchored by non-catalytic fungal dockerin domains (NCDDs) to cohesin domains on large scaffoldin proteins to form enzymatic complexes called fungal cellulosomes [2]. The high-resolution genome presented here enabled a Hidden Markov Model (HMM) analysis of the *C. churrovius* genome, which annotated 36 genes as fungal scaffoldins, compared to the 38 transcripts predicted based on tblastn alignment of the previously sequenced transcriptome [5, 29]. The quantity of predicted proteins identified as cellulases, hemicellulases, and other accessory enzymes along with the total number of CAZymes for each of the 6 sequenced fungal strains are listed in Additional file 1: Table S1. Fewer total CAZymes in the above categories were identified using predicted proteins found in the sequenced genome (338) than were identified by counting the number of transcripts in the sequenced transcriptome (512), which did not take ploidy into account. The highest abundance accessory enzymes identified in the genome were pectin lyases (15.7% of all CAZymes), in contrast to the transcriptome, in which carbohydrate esterases containing SGNH (defined by four invariant residues—serine, glycine, asparagine, and histidine) hydrolase domains were identified as the most abundant (Additional file 1: Table S1) [30, 31]. However, the *C. churrovius* genome also contains the smallest number of polysaccharide lyase domains (PLs) of any of the 6 fungal genomes characterized (Fig. 1).

Proteins containing non-catalytic fungal dockerin domains (NCDDs) were also identified and found to be relatively consistent across strains, in agreement with what was observed for transcript counts (Table 1). However, in contrast to the observation that *C. churrovius* NCDD containing transcripts represented only 15% of

all CAZyme transcripts in comparison to 27.9–31.4% for the three other fungal strains examined, the number of NCDD containing proteins represented 35.9% of all CAZyme proteins for *C. churrovius*, similar to the other three fungal strains (Table 1). This suggests that while *C. churrovius* may place greater emphasis on secreted uncomplexed, free enzymes to attack plant biomass and release fermentable sugars compared to rhizoid-forming anaerobic fungi based on previously collected transcriptional data, its genome still contains a proportion of NCDD proteins similar to that observed in the genomes of rhizoid-forming anaerobic fungal genera. *C. churrovius* also has the second highest number of carbohydrate binding module domains (CBMs) compared to five other high-quality anaerobic fungal genomes (Fig. 1). Further analysis revealed that of these genes, *C. churrovius* also possessed the highest number of CBM family 18 domains among anaerobic fungi sequenced to date (Additional file 1: Fig. S1).

It was previously reported that N6-methyldeoxyadenine (6mA) is associated with transcriptionally active genes in early-diverging fungal lineages in a study using single-molecule long-read sequencing to determine which adenines were methylated [32]. Of the 6692 genes that were methylated when the *C. churrovius* genome was sequenced, 4063 had KOG annotations, 1002 had KEGG annotations, 3450 had GO annotations, and 401 were annotated as CAZymes. Almost 1% of all adenines are methylated, and 93% of modifications are at AT dinucleotides, as shown in Additional file 1: Figure S2A. Very few symmetric runs were present, consistent with avoidance of TAT/ATA reported previously [32]. Modifications are primarily at the start of genes, specifically ramping up in presence at the start of transcription (Additional file 1: Figure S2B). 6mA was rare in repetitive regions of the genome (Additional file 1: Figure S2C) and a large

proportion of total 6mA was restricted to genic space (Additional file 1: Figure S2D).

These results agree with the trends observed for other anaerobic fungal species, further serving to identify 6mA as a widespread epigenetic mark in early-diverging fungi that is associated with transcriptionally active genes [32]. Note that only ~6% of methylated genes in the genome are annotated as CAZymes, indicating that these genes are not always highly transcribed, but rather the majority of CAZymes are transcribed as needed in response to external stimuli, such as co-culture, growth substrate, etc. Nevertheless, association of gene expression with adenine methylation is necessary to understand and develop transformation techniques, which has proven difficult in anaerobic fungi and other non-model eukaryotic systems to date [22, 33]. Accounting for methylated adenine cluster (MAC) positioning and other epigenetic features could help achieve the methylation required to sufficiently overexpress target genes, such as the CAZymes involved in applications requiring biomass breakdown in both fungal monoculture and in anaerobic biomass-degrading consortia [32].

#### Synthetic co-cultures of *C. churrovii* with methanogen *M. bryantii* produce methane

Establishing synthetic co-cultures of anaerobic fungi with methanogens is a valuable tool to probe the impact of co-culture on plant biomass breakdown, substrate uptake, and growth of the individual microbes [9]. Once plant biomass has been broken down into its constituent sugars by fungal CAZymes, they are catabolized by the fungi and other organisms in the native rumen environment [21]. Sugars consumed by the fungi undergo glycolysis in the fungal cytoplasm, and the resulting malate and pyruvate are taken up by the fungal hydrogenosome, where they are converted to H<sub>2</sub> and formate via hydrogenase and pyruvate formate lyase, respectively [2, 34, 35]. The hydrogen and formate produced are then exported and available to neighboring methanogens, which assimilate these products and ultimately generate methane [23]. As such, the metabolic exchange between anaerobic fungi and methanogens benefits both microbes, since it is hypothesized that fungal metabolic end products such as H<sub>2</sub> and formate may inhibit fungal growth and function if allowed to accumulate, while the methanogens are provided with their required growth substrates [36].

Figure 2A summarizes the design of this experiment. Cumulative pressure was measured daily (as a proxy for microbial growth) in order to determine when mid-log growth phase had been reached, at which time the cultures were harvested for RNA extraction as shown in Fig. 2B and C [9]. Gas chromatography was used to determine the concentration of methane and hydrogen

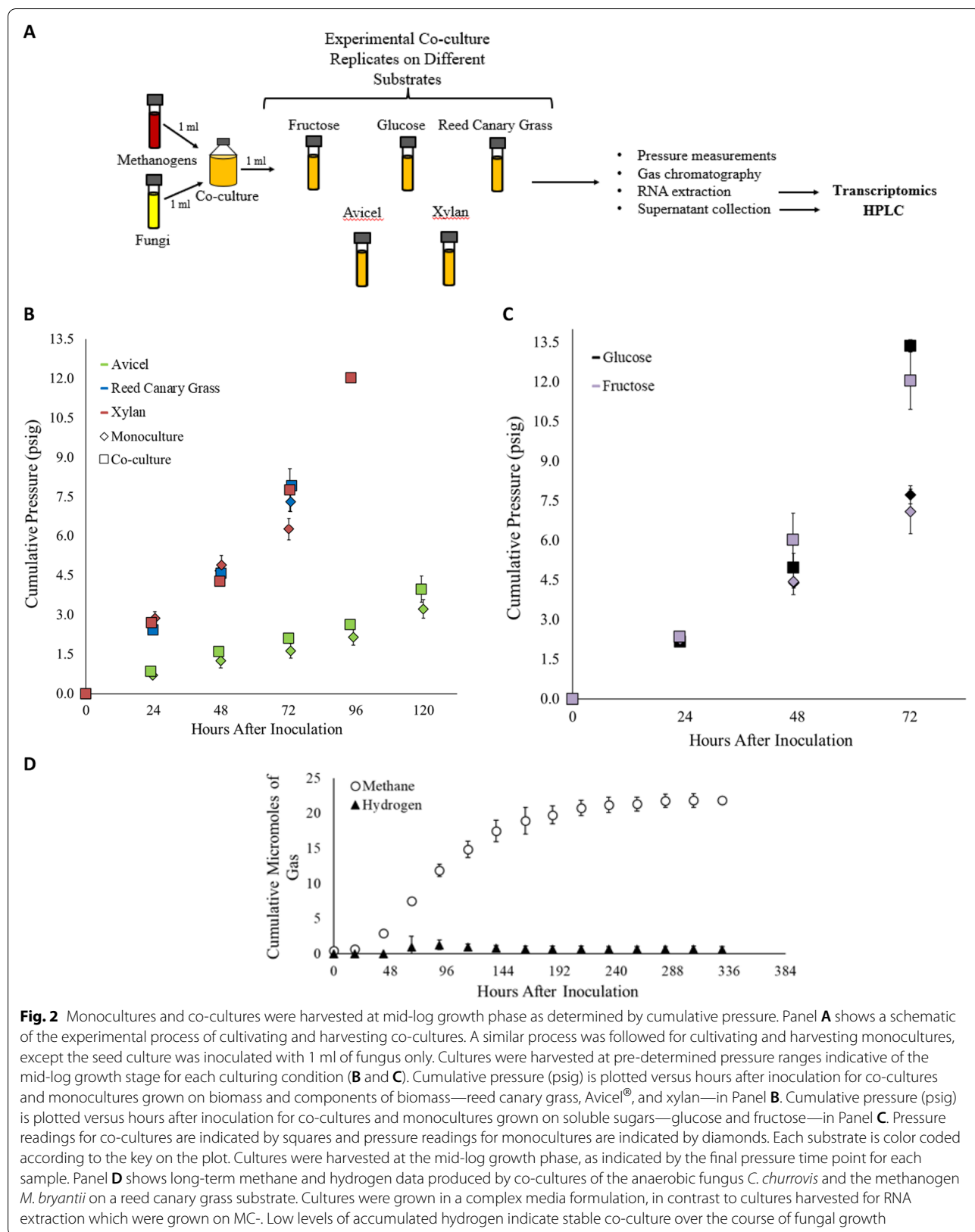
in the headspace gas of synthetic co-cultures and fungal monocultures on each substrate prior to harvest for RNA extraction at mid-log growth phase. No significant amount of hydrogen was detected in the co-cultures, and no methane was detected in the fungal monocultures, in agreement with *M. bryantii*'s H<sub>2</sub>/CO<sub>2</sub> requirement for methane production [37], as shown in Additional file 1: Figure S3. The absence of hydrogen in the co-cultures indicates that stable pairings of the fungus and methanogen were formed on all substrates, including cultivation of the pairing extended to stationary growth phase on the reed canary grass substrate (Fig. 2D), which is consistent with previous observations for the *N. californiae* and *A. robustus* anaerobic fungal strains paired with the same methanogen and grown on cellulose and lignocellulosic reed canary grass [9, 10]. Subsequently, transcriptional regulation coupled with HPLC analysis was used to determine the impact of co-cultivation on fungal sugar utilization, hydrogenosome function, secondary metabolite production, and membrane protein regulation in stable, non-native fungal–methanogen co-cultures.

#### Co-culture with a methanogen enhances production of fungal carbohydrate binding modules and fungal dockerins across multiple substrates

Changes in the transcriptional regulation of anaerobic fungi when challenged by different substrates indicate how the fungal CAZyme repertoire and fungal metabolism are adjusted in response to an altered environment. Often, waste streams containing biomass in industrial settings can vary in composition, potentially affecting bioreactor function through shifts in community composition and metabolic function [38, 39]. Examining these changes using RNA-seq reveals how variations in the composition of growth substrates impact biomass breakdown and product generation. Differential regulation of CAZymes and associated enzymatic machinery was examined for *C. churrovii* co-cultivated with *M. bryantii* and was compared to *C. churrovii* fungal monocultures, both grown on Avicel<sup>®</sup>, reed canary grass, glucose, fructose, and xylan. A proportionally greater number of genes annotated as CAZymes and enzymatic machinery was upregulated in fungal–methanogen co-cultures relative to fungal monocultures than were downregulated on lignocellulose- and hemicellulose-rich substrates, reed canary grass and Avicel<sup>®</sup>. The opposite was true for co-cultures grown on substrates rich in soluble sugars, glucose, fructose, and xylan as shown in the Additional file 1: Figure S4A. The genes upregulated or downregulated for individual CBM, GH, CE, PL, and GT families is shown in Additional file 1: Figure S4B–D.

However, the majority of the ten most highly upregulated genes in these categories in fungal–methanogen





**Table 2** Table of the top ten upregulated fungal genes annotated as CAZymes or associated enzymatic machinery in co-cultures of the anaerobic fungus *C. churrovii* and the methanogen *M. bryantii* relative to fungal monocultures of *C. churrovii* grown on multiple substrates

Avicel	Reed canary grass			Glucose			Fructose			Xylan		
	Protein Id	Log2-fold change	CAZyme annotation	Protein Id	Log2-fold change	CAZyme annotation	Protein Id	Log2-fold change	CAZyme annotation	Protein Id	Log2-fold change	CAZyme annotation
607438	8.00	5.94	CBM 18	407913	9.08	CBM 18	407913	7.72	CBM 18	198053	8.95	CBM 18
547795	5.96	5.08	CBM 18	596610	7.02	GH 73	403091	6.48	GT 71	126623	7.72	CE 6/DOC
629343	5.46	4.87	DOC	555679	5.95	CBM 18	198053	6.13	CBM 18	524258	6.25	CBM 18
620648	5.14	3.91	GT 17	622576	5.87	CBM 18	529683	5.74	DOC	407913	5.73	CBM 18
513365	4.96	3.55	CBM 87	590554	5.55	DOC	436379	4.91	CBM 18	563945	5.29	CBM 18
548447	4.85	3.37	CBM 1	126623	5.39	CE 6/DOC	498097	4.52	DOC	531011	5.23	CBM 18
100311	4.71	3.29	CBM 18	198053	4.81	CBM 18	401262	4.42	DOC	627323	5.06	CBM 18
528501	4.58	2.94	GH 78	498097	4.52	DOC	136193	4.19	DOC	513365	4.53	CBM 87
430401	4.42	2.93	DOC/GH 30	621851	4.45	GH 25	620648	4.10	GT 17	622031	4.11	CBM 18
207551	4.15	2.83	CBM 18	593248	4.40	DOC	607438	4.08	CBM 18	527510	3.92	CBM 18

Co-cultures of the anaerobic fungus and the methanogen and fungal monocultures were grown on Avicel® (A), reed canary grass, glucose, fructose, and xylan. Differential expression of fungal genes in co-cultures relative to fungal monocultures was determined using DESEQ2. The ten genes with the highest log2-fold change in expression in co-culture relative to fungal monoculture are shown in the table above for each substrate and organized into the following classifications: carbohydrate binding module family (CBM), dockerins (DOC), carbohydrate esterase family (CE), glycoside hydrolase family (GH), and glycosyltransferase family (GT). Dockerin-fused CAZymes are indicated by a forward slash between annotations. CBMs were highly upregulated, indicating that there may be an increase in enzymatic machinery that aids in anchoring CAZymes to substrates in co-culture, even when grown on soluble sugars

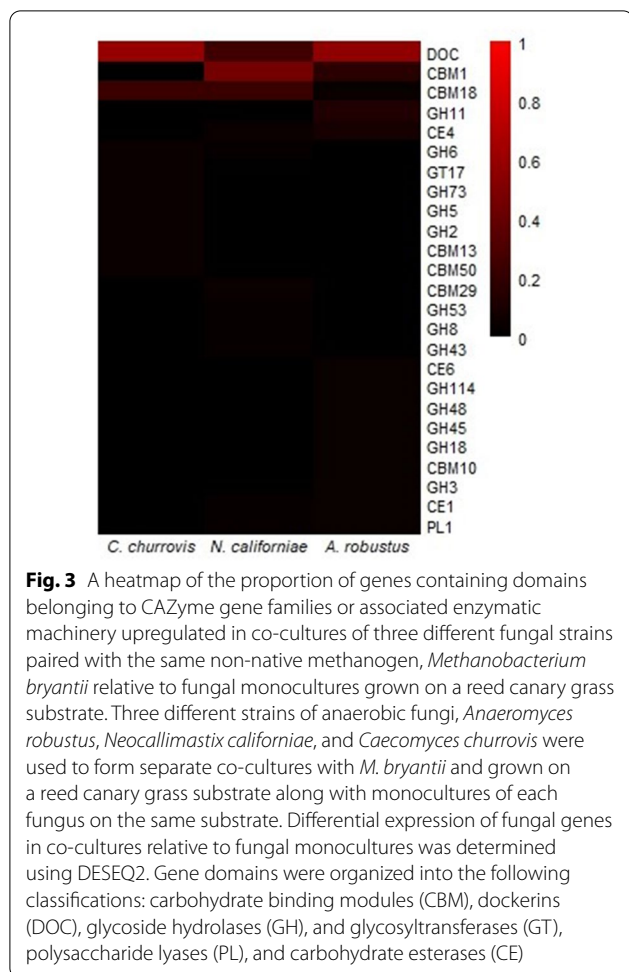


co-culture relative to fungal monoculture on all substrates were annotated as either CBM 18 family proteins or fungal dockerin domains, the majority of which were associated with genes of unknown function. Table 2 shows the top ten most highly upregulated fungal genes according to log<sub>2</sub>-fold change values annotated as CAZymes or associated enzymatic machinery in co-cultures of the anaerobic fungus *C. churrovis* and the methanogen *M. bryantii* relative to monocultures of *C. churrovis* grown on multiple substrates. The CBM family with the most abundant number of genes in the sequenced genome, CBM 18, was consistently the gene classification with the greatest log<sub>2</sub>-fold change of any CAZyme or enzymatic machinery on all substrates in fungal–methanogen co-cultures relative to fungal monocultures. Furthermore, the same CBM 18 gene (*Caecomyces churrovis* protein Id 407913) had the greatest log<sub>2</sub>-fold change in fungal–methanogen co-cultures relative to fungal monocultures on reed canary grass, glucose, and fructose substrates. CBM family 18 modules contain approximately 40 amino acid residues and include members with functions linked to modules with chitinase activity or which are lectins [40, 41]. The modules may therefore either be attached to chitinase catalytic domains or in non-catalytic proteins in isolation or as multiple repeats. These carbohydrate binding proteins possess diversity in ligand specificity and the ability to maintain enzymes in proximity of the substrate, increasing enzyme concentration and potentially leading to more rapid degradation of polysaccharides. These features make these proteins excellent candidates for use in biotechnological applications designed for biomass breakdown [42–45].

The observation that CAZymes, fungal dockerins, and other biomass-degrading machinery are upregulated in all co-cultures, even those grown on glucose, is in agreement with previous studies conducted for fungal–methanogen co-cultures on reed canary grass and glucose at mid-log growth stage [10, 11]. Since the majority of the top ten genes upregulated on all substrates were annotated as either CBM 18 family proteins or fungal dockerin domains, this strongly suggests that co-culture with the methanogen *M. bryantii* results in the transcriptional upregulation of enzymatic machinery associated with biomass degradation. Although no transcriptional upregulation of scaffoldin-encoding genes was initially detected in this study, likely due to the more stringent log<sub>2</sub>-fold change cutoff used to determine significant upregulation, Pre-ranked Gene Set Enrichment Analysis (GSEA) of the entire set of regulated genes revealed that upregulated scaffoldins are significantly enriched in co-cultures grown on Avicel® and reed canary grass [46, 47]. These results agree with the finding by Swift et al. that transcription of fungal cellulosome components

increases in co-culture [10]. Another possibility is that the production of CBM18 transcripts is not related to plant biomass breakdown but instead to interactions between the fungus and methanogen since differential expression is observed across all conditions, including growth on glucose. Many of the dockerin domains not attached to CAZymes contain a CotH kinase protein domain. Previous work showed that approximately 20% of DDPs identified in five previously sequenced anaerobic fungi belonged to spore coat protein CotH and were also present in bacterial cellulosomes [2]. These dockerin domain proteins belonging to spore coat protein CotH have been speculated to be involved in plant cell wall binding, although this remains to be experimentally validated [48]. A list of the top ten upregulated genes containing dockerin domains along with any associated spore coat protein CotH annotations is included in Additional file 2.

The top ten most highly upregulated genes according to log<sub>2</sub>-fold change annotated as CAZymes, CBMs, or fungal dockerins in co-cultures of *C. churrovis* with *M. bryantii* grown on reed canary grass were compared to those upregulated in co-cultures of the same methanogen, *M. bryantii*, with fungal strains *A. robustus* (previously published) and *N. californiae*, grown on the same substrate [10]. A plot of the proportion of genes containing domains belonging to CAZyme gene families or associated enzymatic machinery upregulated in co-cultures of the three different fungal strains paired with the same non-native methanogen, *Methanobacterium bryantii* relative to fungal monocultures grown on a reed canary grass substrate is included in Fig. 3. The genes regulated in CBM, GT, PL, CE, and GH families or containing dockerin domains in the three fungal strains in co-culture versus fungal monoculture on reed canary grass substrate is shown in Additional file 1: Figure S5. The most highly upregulated gene for each strain was a CBM family 18 protein for both the *N. californiae* strain and the *C. churrovis* strain and a Carbohydrate Esterase (family 1) protein for the *A. robustus* strain. For each strain, at least three of the top ten genes were fungal dockerin domains, fused to CAZymes or genes of other function. A high proportion of upregulated genes for all three strains contained dockerin domains and a relatively high proportion of genes containing CBM family 1 or CBM family 18 domains were upregulated for multiple strains as well, as shown in Fig. 3. This comparison suggests that co-cultivation with a methanogen likely encourages substrate channeling between synergistic enzymes for both rhizoid-forming fungal strains (*A. robustus* and *N. californiae*) and non-rhizoid-forming fungi (*C. churrovis*) [2, 10]. Previously, it was suggested that a smaller proportion of CAZyme transcripts containing dockerin domains



in the transcriptome of *C. churrovius* indicated a greater dependence on free enzymes compared to rhizoid-forming gut fungal genera [5]. Nevertheless, with comparative transcriptomic data, upregulation of these non-catalytic modules and CBMs is clearly observed when *C. churrovius* is cultured with *M. bryantii*. This could indicate that anaerobic fungi, regardless of their usual mode of biomass deconstruction, will respond to the presence of other microbes by increasing binding to fibrous substrates. This would allow them more direct access to sugars released during biomass breakdown, which might otherwise be consumed by other microbes.

#### Fungal co-culture with a methanogen may enhance PFL function and production of bottleneck enzymes in sugar pathways

Transcriptional regulation coupled with HPLC analysis was used to determine the impact of methanogen co-cultivation on fungal sugar utilization, genes potentially associated with hydrogenosome function, secondary metabolite production, and membrane protein

regulation in stable, non-native fungal–methanogen co-cultures. Previous studies of fungal–methanogen co-cultures described increased sugar utilization in co-culture [15, 49]. As such, we hypothesized that genes encoding enzymes involved in sugar catabolism would be upregulated in *C. churrovius* and *M. bryantii* co-cultures relative to fungal monocultures. While some enzymes within these pathways showed changes for each substrate, no co-culture condition resulted in uniform upregulation or downregulation of all enzymes within a given sugar pathway, as shown in Additional file 1: Figure S6. The enzymes that were upregulated in fungal–methanogen co-culture relative to fungal monoculture on the same substrate may represent bottlenecks in these catabolic pathways. We suspected that sugar utilization in co-cultures could also be increased through upregulation of sugar transporters in the co-culture condition. We instead observe that in the presence of Avicel® and xylan, *M. bryantii* induces transcriptional upregulation of genes that appear to encode proteins homologous to prokaryotic substrate binding proteins (SBPs), as well as Class C G-protein coupled receptors (GPCRs) as seen in Additional file 1: Table S2 [50–52]. While the function of these protein domains and receptors remains unknown, we speculate that they may be involved in the increased binding of sugar polymers in the presence of the methanogen; or in establishing physical interactions between the methanogens and fungi [53].

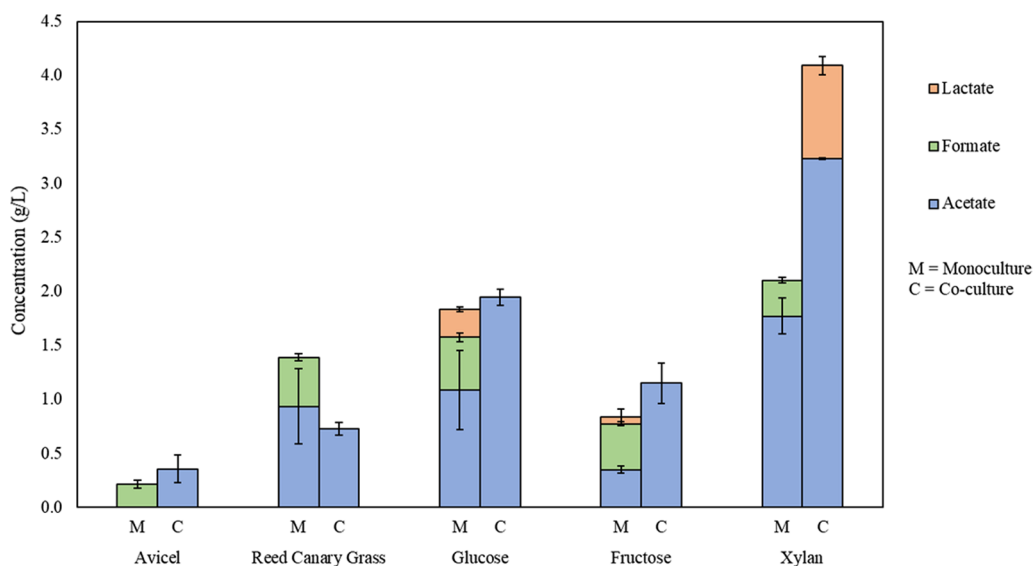
A previous study showed that anaerobic fungal genomes encode a wide array of biosynthetic enzymes of natural products including secondary metabolites—small, bioactive molecules known to mediate a variety of interactions between microorganisms [54–57]. The majority of these genes were not significantly differentially expressed between co-culture and monoculture conditions on the various substrates in this study. However, two of these fungal genes were highly upregulated in co-culture ( $p$ -adjusted < 0.01). The first is a non-ribosomal peptide synthetase (NRPS)-like gene (protein Id 604712), which was upregulated eightfold during growth on fructose and on Avicel®. The second, a polyketide synthase (PKS; protein Id 402343) was fourfold upregulated in co-culture compared to monoculture during growth on xylan and reed canary grass, suggesting that some fungal secondary metabolites may mediate the interaction between *C. churrovius* and *M. bryantii*, depending on the specific substrate. Co-culture interaction may be most notable on Avicel® and xylan substrates, as both transporters and secondary metabolite biosynthesis genes were upregulated in co-culture for both of these substrates.

Based on previous studies noting an increase in metabolites produced by the ATP-generating fungal

hydrogenosome during co-culture with methanogens, we hypothesized that genes associated with hydrogenosomal function would be upregulated in methanogen co-culture [11, 15]. A list of genes associated with the fungal hydrogenosome of the *C. churrovii* strain was constructed based on homology with known hydrogenosome components, shown in Additional file 1: Table S3. FASTA sequences from known hydrogenosomal components identified in the fungal strain *Neocallimastix lanati* [4] were aligned to filtered model proteins of *C. churrovii* using the blastp alignment program in MycoCosm [25]. One or more genes within the *C. churrovii* genome aligned to all listed hydrogenosomal enzymes found in *N. lanati*. Regulation of these genes in co-culture compared to monoculture was examined for each substrate. As shown in Additional file 1: Table S3, 21 genes were homologous to both pyruvate formate lyases (PFLs) that were identified in the *N. lanati* genome [4]. This enzyme reversibly converts pyruvate and CoA into acetyl-CoA and formate, which plays a central role in anaerobic glucose fermentation [58]. It has been shown that this enzyme is functional in hydrogenosomes of the anaerobic fungal species *Piromyces* sp. E2 and *Neocallimastix* sp. L2 [59]. The most notable upregulation of PFLs was observed in cultures grown on xylan and fructose, where 15 of the 21 PFL genes identified by homology were upregulated in co-cultures compared to monocultures grown on xylan and two genes identified by homology

were upregulated in co-cultures compared to monocultures grown on fructose as shown in Additional file 1: Table S3. Five additional genes annotated as PFLs (or formate C acetyltransferases) according to Enzyme Commission (EC) number rather than homology to the *N. lanati* genome were upregulated on xylan and one additional gene was upregulated on fructose. One of these genes (Protein Id 428490) was upregulated in co-culture on all substrates examined except reed canary grass. A previous study examining transcriptional regulation of co-cultures of the native fungus–methanogen pairing *Pecoramyces* sp. F1 with the methanogen *Methanobrevibacter thaueri* versus monoculture of the fungus grown on glucose did not detect a difference in expression levels of PFL genes (although upregulation was detected at the protein level) [11].

Although we hypothesized that genes associated with the hydrogenosome would be transcriptionally upregulated in the co-culture relative to the fungal monocultures based on the metabolic data collected in previous work, transcriptional upregulation of genes associated with hydrogenosomal function is limited, with the exception of pyruvate formate lyases in co-cultures grown on xylan and fructose. It is important to note that further studies are needed to confirm that this transcriptional upregulation of PFLs is associated specifically with the hydrogenosome, as PFLs function in both the cytosol and the hydrogenosome. However, as a complement to the



**Fig. 4** Accumulated metabolites for co-cultures of *C. churrovii* paired with *M. bryantii* versus monocultures of *C. churrovii* upon harvest. HPLC data are shown for co-culture and monoculture grown on each substrate. No formate was observed in co-culture on any substrate, suggesting that *M. bryantii* is capable of metabolizing formate. Trace amounts of ethanol were present in the cultures but fell below the 0.1 g/L limit of detection. This, in conjunction with increased levels of acetate in co-culture, indicates that some of the PFLs upregulated in co-cultures grown on xylan and fructose may be functioning within the hydrogenosome

transcriptional information regarding metabolic function in this study, end-point metabolites present in the supernatant were measured using HPLC upon harvest of the co-cultures and monocultures (Fig. 4). Increases in the amount of acetate produced in co-culture and the absence of significant amounts of ethanol and lactate indicate that some of these genes may potentially be associated with hydrogenosome function for cultures grown on fructose, since pyruvate can either be converted to lactate or ethanol by PFLs functioning in the cytosol or converted to acetate by PFLs functioning within the hydrogenosome. Ethanol was also absent in cultures grown on xylan, although higher levels of lactate were observed in co-culture in addition to higher levels of acetate, indicating that both cytosolic and hydrogenosomal PFLs may be upregulated in co-culture. GSEA Preranked analysis also indicated that upregulated genes were enriched in pathways associated with pyruvate metabolism and glycolysis for co-cultures grown on xylan, in agreement with the observed upregulation of PFLs [46, 47].

While analysis of the end-point metabolites of *A. robustus* paired with *M. bryantii* in previous work did not indicate a statistically significant difference in the level of formate in co-culture versus monoculture, formate was absent in the *C. churrovius* and *M. bryantii* co-culture samples but present in fungal monocultures [10]. Earlier studies concluded that this type strain of *M. bryantii* (DSM 863 M.o.H.) was unable to produce methane from formate in pure culture [60, 61]. The discovery of a formate transporter and several copies of formate dehydrogenase genes upon sequencing the methanogen's genome has suggested the possibility of growth on formate [37]. The observed upregulation of PFL genes and the absence of formate in co-cultures in the current study provide evidence that this strain of *M. bryantii* can utilize formate under certain conditions. A similar phenomenon has been observed for co-cultivation of a formate-producing *Piromyces* fungal species and the natively associated methanogen *Methanobrevibacter thaueri*, a methanogen that has been shown incapable of growth on formate [20, 62]. It is possible that cultivating these methanogens under the conditions required for co-culture with rumen anaerobic fungi stimulates formate utilization by inducing function of the formate transporter and formate dehydrogenase discovered upon sequencing the genome [37].

## Conclusions

Here, we have sequenced the first high-quality genome of a non-rhizoidal fungus, *Caecomyces churrovius*, revealing an abundance of diverse CAZymes and the highest number of CBM family 18 domains among anaerobic

fungi sequenced to date. We found that co-cultivation of the *C. churrovius* fungal strain with the non-native methanogen *M. bryantii* enhanced production of transcripts containing these chitin-binding CBM 18 domains across a variety of substrates. Upregulation of CBMs and dockerin domains in fungal–methanogen co-culture with the same non-native methanogen relative to fungal monoculture on a lignocellulose-rich substrate was also observed for two other previously sequenced fungal strains, *A. robustus* and *N. californiae*. We hypothesize that the function of CBMs belonging to family 18 may not be directly related to plant biomass breakdown but instead to interactions between the fungus and methanogen since upregulation of transcripts containing these domains is observed across multiple cultivation conditions, including both cellulose and lignocellulose-rich substrates as well as soluble sugars. Upregulation of genes associated with sugar pathways and the functioning of the hydrogenosome for *C. churrovius* and *M. bryantii* co-cultures relative to fungal monocultures of *C. churrovius* also suggests that co-culture with a methanogen may enhance pyruvate formate lyase (PFL) function under certain cultivation conditions and production of key enzymes in sugar utilization pathways. Overall, these observations enhance our understanding of the mechanistic interactions between anaerobic fungi and associated methanogens, which aids in our ability to design synthetic biomass-degrading microbial consortia.

## Methods

### Growing and harvesting cultures for RNA extractions

Anaerobic serum bottles containing 80 mL of modified medium C (“MC-”) with 0.8 mL 100× vitamin solution and 0.8 g reed canary grass were inoculated with cultures of *C. churrovius* and *M. bryantii*: 1.0 mL of *C. churrovius* or a combination of 1.0 mL of *C. churrovius* and 1.0 mL of *M. bryantii* (DSM No.-863, DSMZ) (routine cultures were cultivated as described previously by Swift, et al.) [10]. The fungal and methanogen co-cultures and fungal monocultures were grown anaerobically at 39 °C in Hungate tubes filled with 9.0 mL of autoclaved modified medium C [63] (“MC-”), containing 1.25 g/L yeast extract, 5 g/L Bacto™ Casitone, and 7.5 vol% clarified rumen fluid, with either 0.1 g of milled reed canary grass, 0.1 g Avicel®, 0.1 g xylan, 0.5 ml of a 0.1 g/ml sterile filtered glucose stock solution, or 0.1 g/ml of a sterile filtered fructose stock solution as the growth substrate, and supplemented with vitamin solution post-autoclaving [64]. Pressure production was used as a proxy for fungal growth, as described previously [65]. Daily pressure measurements were taken using a probe pressure transducer to determine when the cultures reached the mid-log growth phase, based upon previous pressure growth curves measured to stationary



phase growth. Upon reaching mid-log growth phase, cultures were harvested and stored for later RNA extraction. After sampling the headspace gas of the culture to determine end-point methane and hydrogen concentrations for monocultures and co-cultures, a volume of 1.2 mL of the culture supernatant was pipetted off of the top of the culture and stored at  $-20\text{ }^{\circ}\text{C}$  for later HPLC analysis. The remainder of the culture was transferred to a 15-mL falcon tube and spun down at 10,000 g and  $4\text{ }^{\circ}\text{C}$  for 6 min. The remaining supernatant was then decanted or pipetted off depending upon the integrity of the remaining cell pellet and replaced with 1 mL of RNA-later and mixed by pipetting. Samples were then stored at  $-80\text{ }^{\circ}\text{C}$  until extraction.

#### Measuring hydrogen and methane production

End-point methane and hydrogen measurements for both monocultures and co-cultures were taken from the headspace of the culture tubes before harvesting the cultures. Daily measurements and sampling were performed to monitor the growth of the co-cultures and monocultures. First the pressure in each sample was measured using a pressure transducer [66], and the headspace composition was measured on a gas chromatograph (GC)-pulsed, discharge helium ionization detector (Thermo Fisher Scientific TRACE 1300) [67]. Finally, the headspace pressure of the sample was vented to return the headspace to atmospheric pressure. The total moles of headspace gas were calculated using the ideal gas law. Gas concentrations for  $\text{H}_2$  and methane were calculated using an external standard calibration method. The gas concentration could then be multiplied by the number of moles present both before and after the pressure sampling in order to determine the moles of  $\text{H}_2$  or methane produced. It was assumed that the amount of gas dissolved in the liquid media was negligible for these calculations.

#### HPLC analysis

Levels of volatile fatty acids present in the supernatant of both co-cultures and monocultures were measured using an Agilent 1260 Infinity HPLC (Agilent). Samples were prepared by acidifying to 5 mM using sulfuric acid and subsequently incubating at room temperature for 5 min. Samples were then centrifuged for 5 min at 21,000 g. The supernatant was syringe filtered into an HPLC vial (Eppendorf FA-45-24-11) using a  $0.22\text{ }\mu\text{m}$  PVDF filter. Samples were analyzed on an Agilent 1260 Infinity high-performance liquid chromatography system (HPLC, Agilent, Santa Clara, CA) equipped with an auto-sampler unit (1260 ALS). Separation of formate, acetate, and lactate was achieved with a Bio-Rad Aminex<sup>®</sup> 87H Ion Exclusion Column for organic acids (Part No. 1250140,

Bio-Rad, Hercules, CA) with a mobile phase of 5 mM sulfuric acid. In-house standards were prepared with MC- blank culture medium as a base and sodium formate (ACS Grade, Fisher Chemical S648500), sodium acetate (ACS Grade, Fisher Chemical S210500), and L-lactic acid sodium (99%, extra pure, Acros Organics 439,220,100) at VFA concentrations of 0.1 and 1 g/L.

#### Genome sequencing and annotation of anaerobic fungus *Caecomyces churrovii*

The *Caecomyces churrovii* fungal strain was isolated as described by Henske, et al. [5] Genomic DNA was isolated from cultures grown for 5–7 days on glucose to reduce the interference of plant material during cell lysis. DNA was extracted using the MoBio PowerPlant Pro kit. DNA was isolated from 5–10 cultures grown in 40 mL volumes and pooled together by collecting the DNA in the same silica column. This process was repeated until the total amount of DNA isolated was greater than 12  $\mu\text{g}$ . The *C. churrovii* genome was sequenced using the PacBio sequencing platform.  $>10\text{-kb}$  fragments were size selected using Blue Pippin Size Selection, then 10  $\mu\text{g}$  of genomic DNA was sheared to  $>10\text{-kb}$  fragments using Covaris g-Tubes. The sheared DNA was treated with exonuclease to remove single-stranded ends and DNA damage repair mix followed by end repair and ligation of blunt adapters using SMRTbell Template Prep Kit 1.0 (Pacific Biosciences). The library was purified with AMPure PB beads and size selected with BluePippin (Sage Science) at  $>10\text{-kb}$  cutoff size. PacBio Sequencing primer was then annealed to the SMRTbell template library and sequencing polymerase was bound to them using Sequel Binding kit 2.0. The prepared SMRTbell template libraries were then sequenced on a Pacific Biosystems' Sequel sequencer using v3 sequencing primer, 1 M v2 SMRT cells, and Version 2.0 sequencing chemistry with 6 h & 10 h movie run times. 6 mA modifications were detected using the PacBio SMRT analysis platform (pb\_basemods package; smrtanalysis version: smrtlink/8.0.0.80529). 6 mA modifications were then filtered and methylated genes were identified following the methods described in Mondo et al. [32]. The assembly was completed with Falcon which generates better assemblies than competing methods likely due to an improvement in isolation of high molecular weight DNA and sequencing larger DNA fragments [1, 68, 69]. While annotating fungal genomes present a challenge due to the lack of anaerobic fungal gene content in existing databases, the genome was annotated using the JGI Annotation Pipeline, which employs a variety of gene modelers to discover genes [25]. In addition to homology-based modelers, ab initio gene discovery tools and RNAseq based methods were used for annotation. Models were determined to be allelic if they were

located in regions on smaller scaffolds that were >95% identical at the nucleic acid level and >50% of the smaller scaffold was covered by these regions. The CAZymes of the *C. churrovii* genome were detected and assigned to families by the curators of the CAZy database using the methods used for the daily updates of the CAZy database [41, 70]. Other fungal genomes included in comparisons were sequenced previously [2–4].

#### Extracting RNA from experimental samples

Samples were removed from storage at  $-80^{\circ}\text{C}$  and thawed on ice. After thawing, samples were spun down for 6 min at  $4^{\circ}\text{C}$  and 10,000 g and RNA later was removed. Cells were lysed for the reed canary grass and Avicel<sup>®</sup> cultures using bead beating for 1 min in 30 s intervals and cells were lysed for glucose, fructose, and xylan cultures using liquid nitrogen grinding. Total RNA was extracted using the RNeasy Mini kit (QIAGEN) following the protocol for “Purification of Total RNA from Plant Cells and Tissues and Filamentous Fungi” including an on-column DNase digest. An Agilent TapeStation was used to determine the quality of the sequenced RNA and Qubit High Sensitivity RNA Assay was used to determine concentrations.

#### RNA sequencing and data analysis

Stranded RNASeq library(s) were created and quantified by qPCR for both monoculture and co-culture samples. Stranded cDNA libraries were generated using the Illumina Truseq Stranded mRNA Library Prep kit. mRNA was purified from 1  $\mu\text{g}$  of total RNA using magnetic beads containing poly-T oligos. mRNA was fragmented and reversed transcribed using random hexamers and SSII (Invitrogen) followed by second strand synthesis. The fragmented cDNA was treated with end-pair, A-tailing, adapter ligation, and 8 cycles of PCR. The prepared library was quantified using KAPA Biosystems' next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument. For genome annotation, the quantified library was then prepared for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit, v4. Sequencing of the flow cell was performed on the Illumina HiSeq 2500 sequencer using HiSeq TruSeq SBS sequencing kits, v4, following a  $2 \times 150$  indexed run recipe. Reads filtered for artifacts and trimmed for quality were assembled into consensus sequences using Trinity v. 2.3.2 [71]. For differential gene expression analysis, sequencing of the libraries was performed on the Illumina NovaSeq sequencer using NovaSeq XP V1 reagent kits, S4 flowcell, and following a  $2 \times 150$  indexed run recipe. The filtered reads from each library were aligned to the *Caecomycetes churrovii* genome using HISAT2 version 2.1.0 [72].

Strand-specific coverage was generated using deepTools v3.1 [73]. Raw gene counts were generated using featureCounts, with only primary hits assigned to the reverse strand were included in the raw gene counts [74]. Raw gene counts were used to evaluate the level of correlation between biological replicates using Pearson's correlation and determine which replicates would be used in the DGE analysis. DESeq2 (version 1.18.1) [75] was subsequently used to determine which genes were differentially expressed between pairs of conditions. The parameters used to call a gene DE between conditions were  $p$ -value < 0.05 and a log<sub>2</sub>-fold change greater than 2. This log<sub>2</sub>-fold change cutoff is more stringent than the typical cutoff used in previous studies to account for variation in undefined rumen fluid components across different batches of media. Raw gene counts, not normalized counts, were used for DGE analysis since DESeq2 uses its own internal normalization. Subsequent analysis was done using the filtered model gene catalog for *C. churrovii* provided for download on the MycoCosm website [25]. Pre-ranked Gene Set Enrichment Analysis (GSEA) of regulated genes in co-cultures relative to fungal monocultures for each substrate condition was conducted using 1,000 permutations and weighted enrichment statistics [46, 47]. The TOPCONS web server was used to determine consensus prediction of membrane protein topology for upregulated and downregulated gene sets and sequences were annotated using Pfam and the HMMER web server [50, 76, 77].

#### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13068-021-02083-w>.

**Additional file 1.** Additional tables and figures.

**Additional file 2.** Additional data.

#### Acknowledgements

The work conducted by the US DOE Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the US DOE under contract no. DE-AC02-05CH11231. The authors acknowledge the use of the Biological Nanostructures Laboratory within the California NanoSystems Institute, supported by the University of California, Santa Barbara and the University of California, Office of the President. Data were prepared for submission to the TOPCONS webserver using the computational resources of the Center for Scientific Computing from the CNSI, MRL: an NSF MRSEC (DMR-1121053) and NSF CNS-0960316. The authors thank Patrick Leggieri, Stephen Lillington, and Amy Eisenberg for helpful discussions and revision of the manuscript.

#### Authors' contributions

JLB performed all experiments and analysis for the *C. churrovii* co-cultures and monocultures except for RNA sequencing and subsequent analysis to the point of obtaining raw gene counts. VS, GH, and MY performed RNA sequencing and subsequent analysis to the point of obtaining raw gene counts and independently determined differential gene expression using an alternate version of DESeq2. Experiments, sequencing, and differential gene expression analysis for the *N. californiae* co-culture and monoculture samples were performed by CLS. JLB performed subsequent analysis comparing differential

gene expression in the *C. churrovii* and *N. californiae* co-cultures and monocultures. The genome was collected by JKH and SL and sequenced and annotated by KL, AS, BH, and ED with project management from IG and KB. SJM performed the genome methylation analysis. SS annotated transmembrane sequences using Pfam and the HMMER web server. JLB and MAO conceived the project and wrote the manuscript, with contributions from SJM, SS, and CLS regarding methylation analysis, transmembrane protein regulation, and secondary metabolite regulation, respectively. All authors read and approved the final manuscript.

#### Funding

We thank the following for funding support: the National Science Foundation (NSF, Grant No. MCB-1553721), the Office of Science (BER) of the US Department of Energy (DOE) (Grant No. DE-SC0010352), the Institute for Collaborative Biotechnologies (Grant Nos. W911NF-09-D-0001, W911NF-19-2-0026, and W911NF-19-1-0010) from the US Army Research Office, and the Camille Dreyfus Teacher-Scholar Awards Program.

#### Availability of data and materials

Annotation data for the fungal genomes discussed can be found at <https://mycocosm.jgi.doe.gov>. The *Caecomyces churrovii* genome is listed under Project ID 1099319. RNA-Seq data for *C. churrovii* monocultures, *C. churrovii* and *M. bryantii* co-cultures, *N. californiae* monocultures, and *N. californiae* and *M. bryantii* co-cultures have been deposited under the BioProject accession number PRJNA757726.

#### Declarations

##### Ethics approval and consent to participate

Not applicable.

##### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Department of Chemical Engineering, University of California Santa Barbara, Santa Barbara, CA 93106, USA. <sup>2</sup>US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. <sup>3</sup>Department of Agricultural Biology, Colorado State University, Fort Collins, CO 80523, USA. <sup>4</sup>DTU Bioengineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark. <sup>5</sup>Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia. <sup>6</sup>Architecture Et Fonction Des Macromolécules Biologiques, CNRS/Aix-Marseille University, Marseille, France. <sup>7</sup>INRAE USC1408, AFMB, 13009 Marseille, France. <sup>8</sup>Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, CA, USA. <sup>9</sup>Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA.

Received: 17 August 2021 Accepted: 19 November 2021

Published online: 10 December 2021

#### References

- Edwards JE, Forster RJ, Callaghan TM, Dollhofer V, Dagar SS, Cheng Y, et al. PCR and omics based techniques to study the diversity, ecology and biology of anaerobic fungi: Insights, challenges and opportunities. *Front Microbiol.* 2017;8:1.
- Haitjema CH, Gilmore SP, Henske JK, Solomon KV, De Groot R, Kuo A, et al. A parts list for fungal cellulosomes revealed by comparative genomics. *Nat Microbiol.* 2017;2:1–8.
- Youssef NH, Couger MB, Struchtemeyer CG, Ligginstoff AS, Prade RA, Najjar FZ, et al. The genome of the anaerobic fungus *Orpinomyces* sp. strain C1A reveals the unique evolutionary history of a remarkable plant biomass degrader. *Appl Environ Microbiol.* 2013;79:4620–34.
- Wilken SE, Monk JM, Leggieri PA, Lawson CE, Lankiewicz TS, Seppälä S, et al. Experimentally validated reconstruction and analysis of a genome-scale metabolic model of an anaerobic neocallimastigomycota fungus. *mSystems.* 2021;6:1–22.
- Henske JK, Gilmore SP, Knop D, Cunningham FJ, Sexton JA, Smallwood CR, et al. Transcriptomic characterization of *Caecomyces churrovii*: A novel, non-rhizoid-forming lignocellulolytic anaerobic fungus. *Biotechnol Biofuels.* 2017;10:1–12.
- Couger MB, Youssef NH, Struchtemeyer CG, Ligginstoff AS, Elshahed MS. Transcriptomic analysis of lignocellulosic biomass degradation by the anaerobic fungal isolate *Orpinomyces* sp. strain C1A. *Biotechnol Biofuels.* 2015;8:1–17.
- Peng X, Wilken SE, Lankiewicz TS, Gilmore SP, Brown JL, Henske JK, et al. Genomic and functional analyses of fungal and bacterial consortia that enable lignocellulose breakdown in goat gut microbiomes. *Nat Microbiol.* 2021;6(4):499–511.
- Peng X, Gilmore SP, O'Malley MA. Microbial communities for bioprocessing: lessons learned from nature. *Curr Opin Chem Eng.* 2016;14:103–9.
- Gilmore SP, Lankiewicz TS, Wilken SE, Brown JL, Sexton JA, Henske JK, et al. Top-down enrichment guides in formation of synthetic microbial consortia for biomass degradation. *ACS Synth Biol.* 2019;8:2174–85.
- Swift CL, Brown JL, Seppälä S, O'Malley MA. Co-cultivation of the anaerobic fungus *Anaeromyces robustus* with *Methanobacterium bryantii* enhances transcription of carbohydrate active enzymes. *J Ind Microbiol Biotechnol.* 2019;1:16–9.
- Li Y, Li Y, Jin W, Sharpton TJ, Mackie RI, Cann I, et al. Combined genomic, transcriptomic, proteomic, and physiological characterization of the growth of *Pecoramyces* sp. F1 in monoculture and co-culture with a syntrophic methanogen. *Front Microbiol.* 2019;10:1–12.
- Mountfort DO, Asher RA, Bauchop T. Fermentation of cellulose to methane and carbon dioxide by a rumen anaerobic fungus in a triculture with *Methanobrevibacter* sp. strain RA1 and *Methanosarcina barkeri*. *Appl Environ Microbiol.* 1982;44:128–34.
- Nakashimada Y, Srinivasan K, Murakami M, Nishio N. Direct conversion of cellulose to methane by anaerobic fungus *Neocallimastix frontalis* and defined methanogens. *Biotechnol Lett.* 2000;22:223–7.
- Jin W, Cheng YF, Mao SY, Zhu WY. Isolation of natural cultures of anaerobic fungi and indigenously associated methanogens from herbivores and their bioconversion of lignocellulosic materials to methane. *Bioresour Technol.* 2011;102:7925–31.
- Li Y, Jin W, Cheng Y, Zhu W. Effect of the Associated Methanogen *Methanobrevibacter thaueri* on the Dynamic Profile of End and Intermediate Metabolites of Anaerobic Fungus *Piromyces* sp. F1. *Curr Microbiol.* 2016;73:434–41.
- Teunissen MJ, Kets EPW, den Camp HJM, Veld JHJ, Vogels GD. Effect of coculture of anaerobic fungi isolated from ruminants and non-ruminants with methanogenic bacteria on cellulolytic and xylanolytic enzyme activities. *Arch Microbiol.* 1992;157:176–82.
- Marvin-Sikkema FD, Richardson AJ, Stewart CS, Gottschal JC, Prins RA. Influence of hydrogen-consuming bacteria on cellulose degradation by anaerobic fungi. *Appl Environ Microbiol.* 1990;56:3793–7.
- Bauchop T, Mountfort DO. Cellulose fermentation by a rumen anaerobic fungus in both the absence and the presence of rumen methanogens. *Appl Environ Microbiol.* 1981;42:1103–10.
- Hungate RE. The rumen and its microbes. New York: Academic Press Inc.; 1966.
- Li Y, Jin W, Mu C, Cheng Y, Zhu W. Indigenously associated methanogens intensified the metabolism in hydrogenosomes of anaerobic fungi with xylose as substrate. *J Basic Microbiol.* 2017;57:933–40.
- Henske JK, Wilken SE, Solomon KV, Smallwood CR, Shutthanandan V, Evans JE, et al. Metabolic characterization of anaerobic fungi provides a path forward for bioprocessing of crude lignocellulose. *Biotechnol Bioeng.* 2018;115:874–84.
- Solomon KV, Haitjema CH, Henske JK, Gilmore SP, Borges-Rivera D, Lipzen A, et al. Early-branching gut fungi possess a large, comprehensive array of biomass-degrading enzymes. *Science.* 2016;351:1192–5.
- 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. *Biotechnol Bioeng.* 2014;111:1471–82.
- Solomon KV, Henske JK, Theodorou MK, O'Malley MA. Robust and effective methodologies for cryopreservation and DNA extraction from anaerobic gut fungi. *Anaerobe.* 2016;38:39–46.
- Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, et al. Myco-Cosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Res.* 2014;42:699–704.



26. Seppälä S, Yoo JI, Yur D, O'Malley MA. Heterologous transporters from anaerobic fungi bolster fluoride tolerance in *Saccharomyces cerevisiae*. *Metab Eng Commun*. 2019;9:1.
27. O'Malley MA, Theodorou MK, Kaiser CA. Evaluating expression and catalytic activity of anaerobic fungal fibrolytic enzymes native to *Piromyces* sp. E2 in *Saccharomyces cerevisiae*. *Environ Prog Sustain Energy*. 2011;31:37–46.
28. Wilken SE, Seppälä S, Lankiewicz TS, Saxena M, Henske JK, Salamov AA, et al. Genomic and proteomic biases inform metabolic engineering strategies for anaerobic fungi. *Metab Eng Commun*. 2020;10:1.
29. Eddy SR. What is a hidden Markov model? *Nat Biotechnol*. 2004;22:1315–6.
30. Mølgaard A, Kauppinen S, Larsen S. Rhamnogalacturonan acetyltransferase elucidates the structure and function of a new family of hydrolases. *Structure*. 2000;8:373–83.
31. Akoh CC, Lee GC, Liaw YC, Huang TH, Shaw JF. GDSL family of serine esterases/lipases. *Prog Lipid Res*. 2004;43:534–52.
32. Mondo SJ, Dannebaum RO, Kuo RC, Louie KB, Bewick AJ, LaButti K, et al. Widespread adenine N6-methylation of active genes in fungi. *Nat Genet*. 2017;49:964–8.
33. Obraztsova IN, Prados N, Holzmann K, Avalos J, Cerdá-Olmedo E. Genetic damage following introduction of DNA in *Phycomyces*. *Fungal Genet Biol*. 2004;41:168–80.
34. Mountfort DO, Orpin CG. *Anaerobic Fungi Biology, Ecology, and Function*. 1994.
35. Fontes CMGA, Gilbert HJ. Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Annu Rev Biochem*. 2010;79:655–81.
36. Saye LMG, Navaratna TA, Chong JPI, O'Malley MA, Theodorou MK, Reilly M. The anaerobic fungi: Challenges and opportunities for industrial lignocellulosic biofuel production. *Microorganisms*. 2021;9:1–28.
37. Gilmore SP, Henske JK, Sexton JA, Solomon KV, Seppälä S, Yoo JI, et al. Genomic analysis of methanogenic archaea reveals a shift towards energy conservation. *BMC Genomics*. 2017;18:1–14.
38. De Vrieze J, Verstraete W, Boon N. Repeated pulse feeding induces functional stability in anaerobic digestion. *Microb Biotechnol*. 2013;6:414–24.
39. Sundberg C, Al-Soud WA, Larsson M, Alm E, Yekta SS, Svensson BH, et al. 454 pyrosequencing analyses of bacterial and archaeal richness in 21 full-scale biogas digesters. *FEMS Microbiol Ecol*. 2013;85:612–26.
40. Boraston AB, Bolam DN, Gilbert HJ, Davies GJ. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem J*. 2004;382:769–81.
41. Cantarel BI, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The carbohydrate-active EnZymes database (CAZy): an expert resource for glycoenomics. *Nucleic Acids Res*. 2009;37:233–8.
42. Tomme P, Boraston A, McLean B, Kormos J, Creagh AL, Sturch K, et al. Characterization and affinity applications of cellulose-binding domains. *J Chromatogr B Biomed Appl*. 1998;715:283–96.
43. Bolam DN, Ciruela A, McQueen-mason S, Simpson P, Williamson MP, Rixon JE, et al. Enzyme substrate proximity. *Cultures*. 1998;781:775–81.
44. Shoseyov O, Shani Z, Levy I. Carbohydrate binding modules: biochemical properties and novel applications. *Microbiol Mol Biol Rev*. 2006;70:283–95.
45. Stern J, Morais S, Lamed R, Bayer EA. Adaptor scaffoldins: An original strategy for extended designer cellulosomes, inspired from nature. *MBio*. 2016;7:1.
46. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;102:15545–50.
47. Mootha VK, Lindgren CM, Eriksson K, Subramanian A, Sihag S, Lehar J, et al. PGC-1 $\alpha$ -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genetics*. 2003;34:267–73.
48. Nguyen KB, Sreelatha A, Durrant ES, Lopez-Garrido J, Muszewska A, Dudkiewicz M, et al. Phosphorylation of spore coat proteins by a family of atypical protein kinases. *Proc Natl Acad Sci USA*. 2016;113:E3482–91.
49. Joblin KN, Williams AG. Effect of cocultivation of ruminant chytrid fungi with *Methanobrevibacter smithii* on lucerne stem degradation and extra-cellular fungal enzyme activities. *Lett Appl Microbiol*. 1991;12:121–4.
50. Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: The protein families database in 2021. *Nucleic Acids Res*. 2021;49:D412–9.
51. Ter Beek J, Guskov A, Slotboom DJ. Structural diversity of ABC transporters. *J Gen Physiol*. 2014;143:419–35.
52. Møller TC, Moreno-Delgado D, Pin J-P, Kniazeff J. Class C G protein-coupled receptors: reviving old couples with new partners. *Biophys Rep*. 2017;3:57–63.
53. Seppälä S, Solomon KV, Gilmore SP, Henske JK, O'Malley MA. Mapping the membrane proteome of anaerobic gut fungi identifies a wealth of carbohydrate binding proteins and transporters. *Microb Cell Fact*. 2016;15:1–14.
54. Swift CL, Louie KB, Bowen BP, Olson HM, Purvine SO, Salamov A, et al. Anaerobic gut fungi are an untapped reservoir of natural products. *Proc Natl Acad Sci*. 2021;1:1–10.
55. Sharrar A, Crits-Christoph A, Meheust R, Diamond S, Starr E, Banfield J. crossm Bacterial Secondary Metabolite Biosynthetic Potential in Soil. 2020;11:1–17.
56. Bärenstrauch M, Mann S, Jacquemin C, Bibi S, Sylla OK, Baudouin E, et al. Molecular crosstalk between the endophyte *Paraconiothyrium variabile* and the phytopathogen *Fusarium oxysporum* – Modulation of lipoxygenase activity and beauvericin production during the interaction. *Fungal Genet Biol*. 2020;139:103383.
57. Bayliak UYS and MM. Legume-Rhizobium Symbiosis: Secondary Metabolites, Free Radical Processes and Effects of Heavy Metals. *Co-Evolution Second Metab*. 2020. p. 291–322.
58. Becker A, Fritz-Wolf K, Kabsch W, Knappe J, Schultz S, Volker Wagner AF. Structure and mechanism of the glycol radical enzyme pyruvate formate-lyase. *Nat Struct Biol*. 1999;6:969–75.
59. Akhmanova A, Voncken FGJ, Hosea KM, Harhangi H, Keltjens JT, den Camp HJM, et al. A hydrogenosome with pyruvate formate-lyase: Anaerobic chytrid fungi use an alternative route for pyruvate catabolism. *Mol Microbiol*. 1999;32:1103–14.
60. Guyot J, Brauman A. Methane production from formate by syntrophic association of *Methanobacterium bryantii* and *Desulfovibrio vulgaris* JJ. *Appl Environ Microbiol*. 1986;52:1436–7.
61. Benstead J, Archer D, Lloyd D. Formate utilization by members of the genus *Methanobacterium*. *Arch Microbiol*. 1991;156:34–7.
62. Miller CL. Description of *Methanobrevibacter gottschalkii* sp. nov., *Methanobrevibacter thaueri* sp. nov., *Methanobrevibacter woesei* sp. nov. and *Methanobrevibacter wolnii* sp. nov. *Int J Syst Evol Microbiol*. 2017;52:819–22.
63. Theodorou MK, Brookman J, Trinci APJ. *Anaerobic fungi*. Berlin: Methods in gut microbial ecology for ruminants. Springer; 2005.
64. Teunissen MJ, den Camp HJM, Orpin CG, Huis In 't Veld JHJ, Vogels GD. Comparison of growth characteristics of anaerobic fungi isolated from ruminant and non-ruminant herbivores during cultivation in a defined medium. *J Gen Microbiol*. 1991;137:1401–8.
65. Theodorou MK, Davies DR, Nielsen BB, Lawrence MIG, Trinci APJ. Determination of growth of anaerobic fungi on soluble and cellulosic substrates using a pressure transducer. *Microbiology*. 1995;141:671–8.
66. Theodorou MK. A simple gas production method using a pressure transducer to determine the fermentation kinetics of ruminant feeds. *Anim Feed Sci Technol*. 1994;48:185–97.
67. Cai H, Stearns SD. Pulsed discharge helium ionization detector with multiple combined bias/collecting electrodes for gas chromatography. *J Chromatogr A*. 2013;1284:163–73.
68. Pacific Biosciences Falcon Github <https://github.com/PacificBiosciences/FALCON>
69. Gordon D, Huddleston J, Chaisson MJP, Hill CM, Kronenberg ZN, Munson KM, et al. Long-read sequence assembly of the gorilla genome. *Science*. 2016;352.
70. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 2014;42:490–5.
71. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29:644–52.
72. Kim D, Langmead B, Salzberg SL. HISAT: A fast spliced aligner with low memory requirements. *Nat Methods*. 2015;12:357–60.

73. Ramírez F, Dündar F, Diehl S, Grüning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* 2014;42:W187–91.
74. Liao Y, Smyth GK, Shi W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30:923–30.
75. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550.
76. Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Res.* 2015;43:W401–7.
77. Finn RD, Clements J, Eddy SR. HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Res.* 2011;39:29–37.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

