Metabolic characterization of anaerobic fungi provides a path forward for consolidated bioprocessing of crude lignocellulose

by Henske, J.K., Wilken, S.E., Solomon, K.V., Smallwood, C.R., Shutthanandan, V., Evans, J.E., Theodorou, M.K. and O'Malley, M.A.

Harper Adams University

1 **Metabolic characterization of anaerobic fungi provides a path forward for bioprocessing of crude**

2 **lignocellulose**

3 John K. Henske[1], St. Elmo Wilken[1], Kevin V. Solomon[1,2], Chuck R. Smallwood[3], Vaithiyalingam

4 Shutthanandan[3], James E. Evans[3], Michael K. Theodorou[4], Michelle A. O'Malley[1‡]

5 [1]Department of Chemical Engineering, University of California, Santa Barbara, CA 93106, USA

6 [2]Agriculture and Biological Engineering, Purdue University, W. Lafayette, IN 47907, USA

7 [3]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA

8 99354, USA

9 [4]Animal Production, Welfare and Veterinary Sciences, Harper Adams University, Newport, Shropshire,

10 TF10 8NB, UK

11 [‡]Corresponding Author. E-mail: momalley@engineering.ucsb.edu

12

13

14

15

16

17

18

19

20

21

22

23

24 **Highlights**

25   • Gut fungi efficiently degrade complex biomass with a combination of enzymatic hydrolysis and
26   mechanical disruption

27   • Regulation patterns of carbohydrate active enzymes by substrate availability provide insight into the
28   optimal conditions required for enzyme production

29   • Excess amounts of glucose, fructose, xylose, and arabinose are released from biomass during fungal
30   growth

31   • Identification of complete and incomplete sugar catabolic processes in gut fungi identify sugars
32   suitable for hand-off to additional organisms

33   • Hydrolyzed sugars can be fed to model microbes for production of value added products in a two-
34   stage consolidated bioprocessing approach

35

36   **Abstract**

37   The conversion of lignocellulose-rich biomass to bio-based chemicals remains a grand challenge, as single-
38   microbe approaches cannot drive both deconstruction and fermentation steps. In contrast, consortia
39   based bioprocessing leverages the strengths of different microbes to distribute metabolic loads and
40   achieve process synergy, product diversity, and bolster yields. Here, we describe a biphasic fermentation
41   scheme that combines the lignocellulolytic action of anaerobic fungi isolated from large herbivores with
42   domesticated microbes for bioproduction. When grown in batch culture, anaerobic fungi release excess
43   sugars from both cellulose and crude biomass due to a wealth of highly expressed carbohydrate active
44   enzymes (CAZymes), converting as much as 49% of cellulose to free glucose. This sugar-rich hydrolysate
45   readily supports growth of *S. cerevisiae*, which can be engineered to produce a range of value-added
46   chemicals. Further, reconstruction of metabolic pathways from transcriptomic data reveals that anaerobic
47   fungi do not catabolize all sugars that their enzymes hydrolyze from biomass, leaving other carbohydrates
48   such as galactose, arabinose, and mannose available as nutritional links to other microbes in their

49   consortium. Although basal expression of CAZymes in anaerobic fungi is high, it is drastically amplified by

50   cellobiose breakout products encountered during biomass hydrolysis. Overall, these results suggest that

51   anaerobic fungi provide a nutritional benefit to the rumen microbiome, which can be harnessed to design

52   synthetic microbial communities that compartmentalize biomass degradation and bioproduct formation.

53

54   **Keywords**

55   Consolidated bioprocessing; Sugar metabolism; Co-culture; biomass hydrolysis

56

57   **1. Introduction**

58   New approaches to harness lignocellulosic feedstocks for energy and chemical production are

59   needed to grow a sustainable bio-based economy (1). Most fermentation processes utilize microbes that

60   require simple sugars as feedstocks. Lengthy, expensive, and often harsh pretreatments are used to

61   separate carbohydrate fractions from crude biomass (2) that must then be hydrolyzed into fermentable

62   sugars using large cocktails of enzymes (3). Combining lignocellulose hydrolysis, biocatalysis, and

63   conversion in a single bioprocess would improve the efficiency of bio-based chemical production.

64   Typically, consolidated bioprocessing (CBP) approaches rely on endowing model organisms with

65   cellulolytic activity or engineering natively cellulolytic organisms for bioproduction (4). However, the

66   ability to compartmentalize breakdown and production steps within different microbes offers a third path

67   forward, and capitalizes on the strengths of specialist microbes to combine deconstruction, fermentation,

68   and conversion steps (5-9). While such strategies are promising, existing consortia-enabled technologies

69   still require extensive pretreatment to remove lignin from biomass prior to breakdown and conversion.

70   The use of environmental microbes that effectively degrade crude lignocellulose eliminates the

71   need for these pretreatment steps. In this regard, anaerobic gut fungi are members of a natural microbial

72   community found in the guts of many large that evolved to break down crude plant material (10-12).

3

73   These primitive fungi degrade lignin-rich biomass (13) through the secretion of cellulases, hemicellulases,

74   and other yet unknown hydrolytic mechanisms aided by secretion of extracellular fungal cellulosomes (14,

75   15). Gut fungi are critical members in the gut microbiome of large herbivores, where they form syntrophic

76   relationships with rumen methanogens that convert the carbon dioxide and hydrogen they produce into

77   methane (13, 16). While largely unexplored, it is likely that these fungi liberate additional micro or

78   macronutrients during lignocellulose hydrolysis that benefit other members within their community.

79   Despite their powerful natural lignocellulolytic activity, gut fungi have not been incorporated into

80   industrial biomass processing strategies, largely due to a lack of genetic information, genetic tools, and a

81   detailed understanding of their metabolism.

82   Here, we evaluated the potential of two recently isolated (17) strains of anaerobic gut fungi,

83   *Neocallimastix californiae* (IF551675) and *Anaeromyces robustus* (IF551676)*, for their use in a CBP co-

84   culture strategy with the model production microbe *Saccharomyces cerevisiae*. Through transcriptomic

85   analysis we established the catabolic pathways of biomass derived sugars to predict the carbohydrates

86   utilized by gut fungi, and those likely to be left behind for potential microbial partners. Differential

87   expression analysis also identified culture conditions required to enhance biomass degrading enzyme

88   production. In *A. robustus,* growth on cellobiose triggered expression of cellulases, hemicellulase, and

89   accessory enzymes, yet in *N. californiae* only cellulases were upregulated by cellobiose while

90   hemicellulases were activated by biomass substrates. Batch fermentation experiments revealed that high

91   production of fungal enzymes led to the accumulation of excess sugars in the culture medium, enabling

92   biphasic fermentation opportunities that harness the excess sugars to support growth of non-cellulolytic

93   organisms, like *S. cerevisiae*. Overall, this work shows that anaerobic gut fungi can consolidate

94   pretreatment and hydrolysis steps, providing sugar rich hydrolysate to support growth of model microbes

95   for bioproduction directly from lignocellulose.

96

2. Materials/Methods

2.1. Culture maintenance of gut fungi and methodology for growth experiments.

Anaerobic media preparation and gut fungal culture procedures were used throughout this work. Anaerobic gut fungi were routinely grown at 39°C in 10 mL cultures of Medium C (18) containing ground reed canary grass (4 mm particle size) in 15 mL Hungate tubes with 100% CO_2 headspace. Cultures were transferred to new media every 3-5 days. For differential expression experiments, source cultures grown in 80 mL of medium C in 120 ml serum bottles containing reed canary grass were used to inoculate all 10 mL experimental cultures. Fungi were grown on a variety of carbon sources including glucose (anhydrous, Thermo Fisher Scientific, Canoga Park, CA), maltose (Sigma-Aldrich, St Louis, MO), cellobiose (Sigma-Aldrich), Avicel (PH-101, 50 µm particle size, Sigma-Aldrich), corn stover, reed canary grass, switchgrass, and alfalfa stems; biomass substrates were provided by the USDA-ARS Research Center (Madison, WI). Soluble substrates were added to a final concentration of 5 g/L and particulate substrates to a final concentration of 10 g/L.

To monitor fungal proliferation, the pressure of fermentation gases was measured during growth (19). Cultures that accumulated pressure significantly more than the blank control (inoculated 10 mL Medium C culture lacking a carbon source) were considered positive for growth. Effective net specific growth rates were determined from pressure accumulation data during the phase of exponential gas accumulation.

For sugar release experiments, fungal cultures were grown on Avicel and reed canary grass (4mm particles) in 10 mL cultures containing anaerobic Medium C. Cellulose cultures contained 100 or 200 mg of cellulose, and biomass cultures contained either 100 mg or 500 mg of reed canary grass. Pressure measurements were taken three times per day to track growth of the fungi. Aliquots of 0.1 mL supernatant were removed from cultures for sugar determinations using either a YSI 2900 substrate

120   analyzer with YSI 2365 glucose detection membrane kits (YSI Inc., Yellow Springs, OH) or HPLC as described

121   below.

122

123   *2.2. Analysis of Sugars (HPLC)*

124   Sulfuric acid (0.85 M) was added (1 in 10 volumes) to fungal hydrolysate samples, that were then

125   vortexed and allowed to stand for 5 min at room temperature. Nine volumes of water were added and

126   the sample again vortexed briefly, centrifuged for 5 minutes at 21000xg, and the supernatants were

127   extracted with a syringe and filtered into HPLC vials using a 0.22µm filter. Samples were run on an Agilent

128   1260 Infinity HPLC (Agilent) using a Bio-Rad Aminex HPX-87P column (Part No. 1250098, Bio-Rad,

129   Hercules, CA) with inline filter (Part No. 5067-1551, Agilent), Bio-rad Micro-Guard De-Ashing column (Part

130   No. 1250118, Bio-Rad), and Bio-Rad Micro-Guard CarboP column (Part No. 1250119, Bio-Rad) in the

131   following orientation: Inline filter>De-Ashing>CarboP>HPX-87P. Samples were run with a water mobile

132   phase at a flow rate of 0.5 mL/min and column temperature of 80°C. Signals were detected using a

133   refractive index detector. HPLC standards were created for cellobiose, maltose, sucrose, glucose, fructose,

134   galactose, xylose, mannose, and arabinose at 1%, 0.1%, and 0.01% w/v concentrations in Medium C and

135   the above protocol was followed to run each standard.

136

137   *2.3. Helium Ion Microscopy*

138   Fungi grown on various substrates were chemically fixed with 2% glutaraldehyde (Sigma Aldrich)

139   and dehydrated through a series of 10 mL step-gradients from 0% to 70% ethanol then centrifuged at 4°C

140   (3000Xg for 2 mins). The biomass was washed twice more with 10mL of 100% ethanol for 15 mins, then

141   centrifuged and finally resuspended in 5mL of 100% ethanol to remove any residual water. Fungal and/or

142   plant biomass suspensions in 100% ethanol were gently extracted by wide-mouth pipet and placed onto

143   stainless steel carriers for automatic critical point drying (CPD) using an Autosamdri-815 (Tousimis,

144 Rockville, MD), with CO_2 as a transitional fluid. The CPD-processed biomass was mounted onto aluminum

145 stubs and sputter coated with approximately 10 to 20nm of conductive carbon to preserve the sample

146 surface information and minimize charge effects. Secondary electron images of the samples were

147 obtained using Orion helium ion microscope (HIM) (Carl Zeiss Microscopy, Peabody, MA) at 25 or 30 keV

148 beam energy, with a probe current range of 0.1 to 1 pA. Prepared samples were transferred into the HIM

149 via load-lock system and were maintained at $\sim 3 \times 10^{-7}$ Torr during imaging. Use of a low energy electron

150 flood gun (~ 500 eV) was applied briefly interlaced with the helium ion beam that enabled charge control

151 to be maintained from sample to sample. The image signal was acquired in line-averaging mode, with 16

152 lines integrated into each line in the final image with a dwell time of 1µs at a working distance range of 7

153 to 8 mm. Charge neutralization was applied to the sample after each individual line pass of the helium ion

154 beam, which displaced charges on the surface minimizing charging effects in the final image. No post-

155 processing procedures were applied to the digital images besides standard noise reduction, brightness

156 and contrast adjustment using Photoshop plugins.

157

158 *2.4. Metabolic Map Reconstruction from Annotated Transcriptomes*

159 Transcriptomes were annotated as described by Solomon et al. (20). Enzymes present in the

160 metabolic maps were determined based on the presence of enzyme commission (EC) numbers (21).

161 Metabolic maps present in the KEGG database (22) were completed based on EC numbers identified from

162 the transcriptome annotations. Gaps in metabolic maps were then identified and filled by searching the

163 entire annotation, including BLAST (23) and InterPro (24).

164

165 *2.5. RNA Isolation, Library Preparation, and Sequencing*

166 RNA was isolated as described in Solomon et al. (20), and quantity and quality were measured on

167 a Qubit fluorimeter (Qubit, New York, NY) and Tapestation 2200 (Agilent, Santa Clara, CA), respectively.

168 Sequencing libraries were prepared using an Illumina TruSeq Stranded mRNA library prep kit (Illumina

169 Inc., San Diego, CA) following the kit protocol. A separate library was created for each fungus with each

170 growth condition in triplicate. For each sample from *Neocallimastix californiae*, 600 ng of total RNA was

171 used while for each sample from *Anaeromyces robustus*, 400 ng of total RNA was used as input for the

172 library preparation. Once the library preparation was completed, samples from each fungus were pooled

173 together into two separate cDNA libraries with a final concentration of 10 nM. Libraries were sequenced

174 on a NextSeq 500 (Illumina) using High Output 150 Cycle reagent kits in a paired-end 75 base

175 configuration. Samples for *N californiae* and *A. robustus* were sequenced on separate flow cells.

176

177 *2.6. Expression data analysis*

178   Counts of transcripts were quantified by using the RSEM analysis utility within the TRINITY

179 programming package (25). Transcriptomes previously obtained (20) were used as reference templates

180 to obtain count data. Expected counts from this analysis were then fed into the DESeq2 package (26) in

181 the R programming language to determine statistically significant changes in expression as a function of

182 different substrate growth conditions, with a minimum of one $\log_2$ fold change in expression and p-value

183 ≤ 0.01 compared to basal expression on glucose. Bar plots showing changes in expression were made

184 using the transcripts per million (TPM) (27) output from the RSEM analysis. All sequencing data for

185 expression analysis are deposited on GEO (project # GSE95479).

186

187 *2.7. Yeast and Bacteria Culture*

188   Following release of sugar-rich hydrolysates by gut fungi, liquid media was removed from the

189 Hungate tube using a syringe needle and placed in a sterile growth tube that was then inoculated with

190 *Saccharomyces cerevisiae* (BJ5464) or *Escherichia coli* (XL1-Blue). Growth of yeast and bacterial cultures

191 was tracked using optical density measurements at 600 nm (OD_{600}). Cultures were inoculated at a target

192   OD_{600} of 0.5 for yeast cultures and 0.1 for bacteria cultures and grown aerobically in shaker incubators set

193   to 30°C and 225 rpm for yeast, and 37°C and 225 rpm for *E. coli*.

194

195   **3. Results and Discussion**

196   *3.1. Gut fungi are powerful chemical and mechanical degraders of lignocellulose*

197   Anaerobic gut fungi are a valuable untapped resource for lignocellulosic bioprocessing due to their

198   innate ability to degrade crude biomass through abundant secretion of diverse carbohydrate active

199   enzymes (20). Here, we characterized the biomass-degrading activity of two unique anaerobic gut fungal

200   isolates that are attractive to CBP strategies as they effectively degrade plant material without

201   pretreatment.

202   *N. californiae* is a monocentric fungus that forms only a single sporangium on each unit of

203   vegetative growth (thallus) while *A. robustus* is polycentric, capable of forming multiple sporangia from a

204   single center of growth (28). While this results in a significant morphological difference between the two

205   fungi, it is unclear what, if any, metabolic differences are correlated with this attribute. Figure 1 illustrates

206   the vegetative growth of each fungus and their extensive rhizoidal network growing into particles of crude

207   reed canary grass. This growth morphology was consistent with cultures grown on soluble substrates

208   (Figure 1) and additional fibrous substrates (Figure S1). Here, fungal rhizoids aid in plant breakdown via

209   mechanical disruption and work in conjunction with secreted enzymes to deconstruct biomass (29) and

210   increase the biomass surface area to enhance degradation by other cellulolytic bacteria (30).
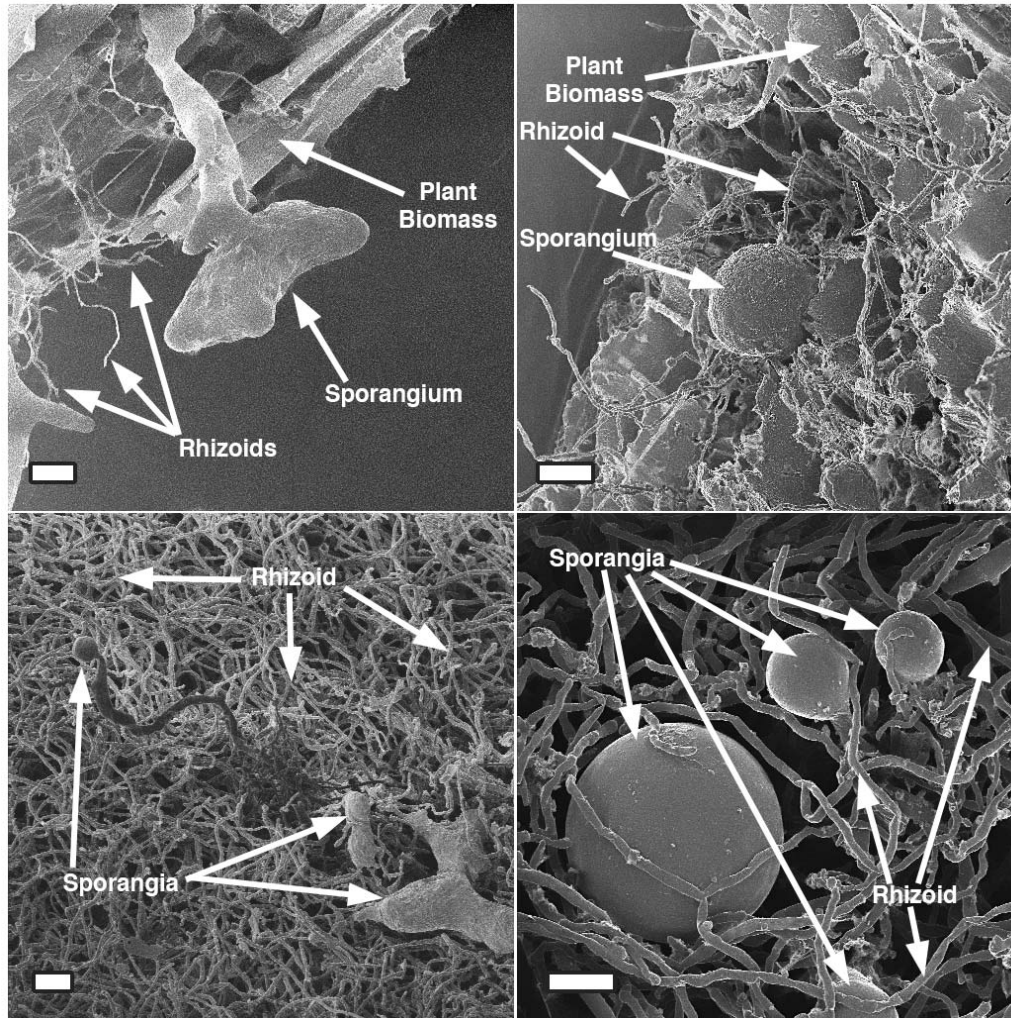
Figure 1. Gut fungi possess extensive rhizoidal network that penetrates into crude biomass. Helium ion micrographs of the sporangial structures of two recently classified gut fungal strains growing on lignocellulosic biomass. *Anaeromyces robustus* (top left) and *Neocallimastix californiae* (top right) grown on reed canary grass form root structures that penetrate the plant material. The same fungi grown on soluble a sugar, glucose, (*A. robustus* bottom left, *N. californiae* bottom right) still grow extensive root networks in the absence of plant biomass. All scale bars represent 10 micrometers.

Both strains of gut fungi thrive on substrates ranging from simple sugars to cellobiose, cellulose, and lignocellulose (20) displaying similar growth rates on complex biomass and simple monosaccharides. Effective net specific growth rates greater than 4.0×10^{-2} hr^{-1} on glucose, fructose, cellobiose, maltose, crystalline cellulose, and lignocellulose (Table S1) suggest that the extra energy required to express and secrete the enzymes required to break down complex biomass did not hinder growth. While some gut

226   fungi have been documented to grow on xylose (31), *N. californiae* displayed no growth while *A. robustus*

227   displayed inconsistent growth on xylose in batch culture, perhaps due to subtle environmental cues (e.g.

228   pH) that may govern xylose assimilation. Neither fungal isolate grew on xylan or carboxymethyl cellulose

229   (Table S1). These results identify strengths and limitations in the carbohydrate utilization profile of each

230   strain that could be exploited for CBP. For example, galactose and arabinose are expected to be liberated

231   during lignocellulose digestion, but did not support growth of the gut fungi; these sugars may serve as

232   metabolic links to a second organism that can catabolize these substrates.

233

234   *3.2. Anaerobic fungi release excess sugars from crude biomass*

235   In nature, gut fungi survive in a competitive microbial community, but in isolation, they have no

236   competition for sugars and other resources and their extracellular cellulolytic enzymes are not subject to

237   extensive proteolytic degradation. Therefore, we hypothesized that fungal enzymes hydrolyze more

238   sugars from biomass than are necessary to support fungal growth. To evaluate this hypothesis, the

239   concentration of glucose was quantified in isolated cultures of *N. californiae* and *A. robustus* grown on

240   crystalline cellulose. From 100 milligrams of crystalline cellulose in a 10-mL culture, *A. robustus* yielded

241   49.1 ± 2 milligrams of soluble excess glucose with a maximum rate of 0.303 mg/hr and *N. californiae*

242   yielded 49.3 ± 4 mg with a maximum rate of 0.287 mg/hr. The bulk of glucose was released after fungal

243   growth had ceased, perhaps due to the continued biocatalysis of secreted enzymes (Figure S2A). The

244   maximum rate of glucose consumption (Figure S2B), 1.470 mg/hr and 0.590 mg/hr for *A. robustus and N.*

245   *californiae*, respectively, was greater than the rates of glucose release. This suggests that the fungal

246   enzymes remained active and continued hydrolysis well beyond fungal death. This excess hydrolytic

247   capacity was highlighted when cellulose loading was increased to 200 mg in 10 mL of media and resulted

248   in nearly doubling the excess glucose released by *A. robustus*, although it had no significant effect on sugar

249   release by *N. californiae* (Figure S3).
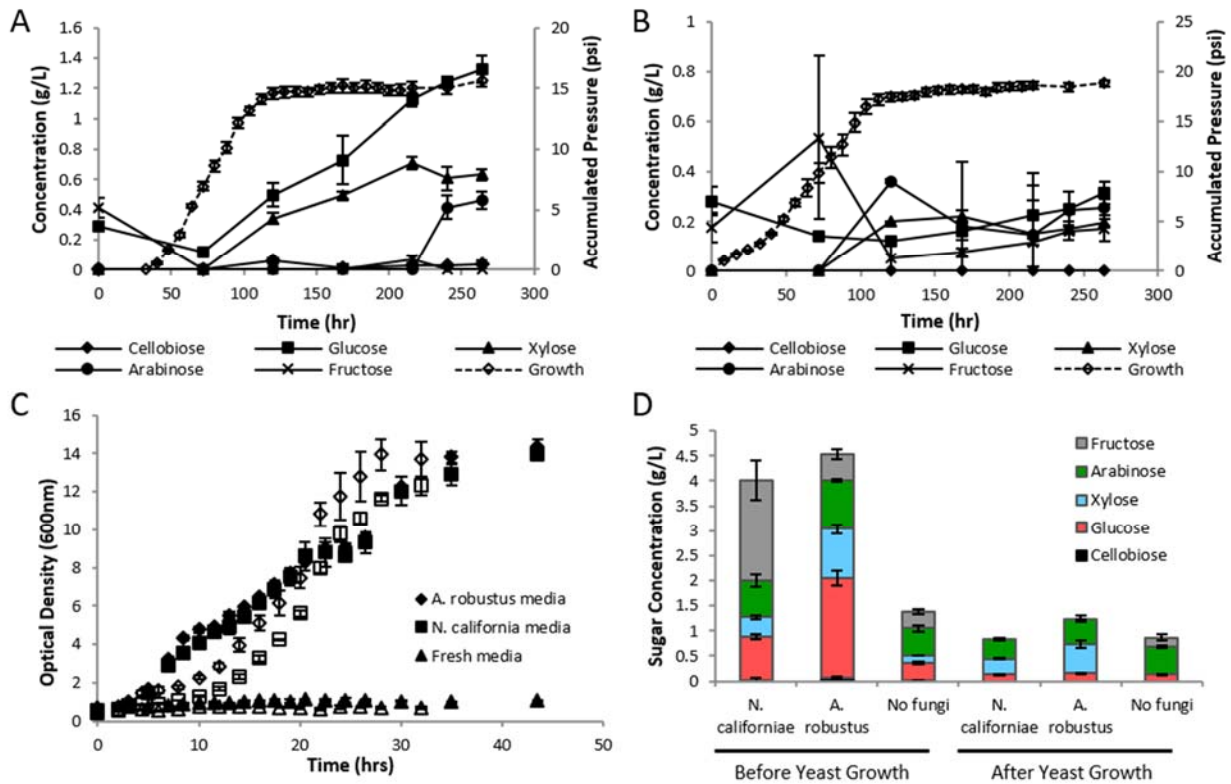
Figure 2. Excess sugars are released from cellulosic and lignocellulosic substrates by anaerobic fungi. A) Growth of *A. robustus* on 0.5 g of reed canary grass in 10 mL culture, and sugar cocentrations released from biomass. Growth (pressure) data is shown in empty symbols and sugar data in solid symbols. B) Growth of *N. californiae* on 0.5 g of reed canary grass in 10 mL culture, and sugar concentrations released from biomass. Growth (pressure) data is shown in empty symbols and sugar data in solid symbols. C) Growth of *S. cerevisiae* on fungal spent media. Spent media containing crystalline cellulose broken down by the fungi into glucose (filled symbols) or reed canary grass broken down into glucose and other sugars (empty symbols). D) End point sugar concentrations produced after fungal growth on reed canary grass, and sugar concentration after yeast proliferation in spent fungal hydrolysate media.

Subsequently, fungi were grown on reed canary grass (lignocellulose) to determine if excess sugars were available following hydrolysis of more industrially-relevant unpretreated biomass substrates (Figure 2A-B). When grown on 500 mg of reed canary grass, *A. robustus* yielded 16.4 ± 1.2 mg of excess glucose and *N californiae* yielded 7.1 ± 0.5 mg glucose in a 10mL batch culture. Considering the reed canary grass cell wall composition with approximately 21% glucose from cellulose (32), this indicates that *A. robustus* released at least 16% of the total cellulose in the reed canary grass as excess glucose. While this yield was significantly lower than the 49% released from pure cellulose and is likely due to the increased complexity of plant material, additional sugars derived from hemicellulose were also present in the

268   hydrolysate in high abundance. Xylose, arabinose, and fructose were also measured in the hydrolysate of

269   each fungus (Figure 2D, Table S2). *A. robustus* and *N. californiae* yielded a total accumulated sugar

270   concentration of 4.5 ± 0.4 and 4.0 ± 0.6 g/L, respectively. We expect that cellobiose is primarily hydrolyzed

271   to glucose or directly taken up due to a wealth of putative cellobiose transporters (33), though trace

272   amounts were detected in the hydrolysate (Table S2). We note that a small amount of sugar was released

273   from the reed canary grass upon autoclaving the media - these are likely soluble sugar components or

274   easily hydrolyzed components of hemicellulose. However, these sugars were rapidly consumed by the

275   fungi (Fig. 2A-B) and the measured quantities were released at later times due to high fungal enzyme

276   activity.

277   Consistent with previous observations, the bulk of the excess sugar release was observed after

278   fungal growth was diminished on the fibrous substrates (Figure 2A-B). While excess xylose and arabinose

279   were expected to be present based on the results of growth experiments (Table S1), glucose was likely

280   present in large quantities because it is the most abundant sugar in biomass, and it is present in greater

281   abundance than needed to support fungal growth. Additional fungal cultures grown on 500 mg of reed

282   canary grass were killed with hygromycin B during exponential growth at 72 hours post-inoculation to

283   evaluate the capability of fungal enzymes alone to hydrolyze crude biomass. These cultures yielded

284   greater amounts of overall sugars, with the largest increases in the amount of glucose released (Figure

285   S4). These results present gut fungi as a source for an improved enzyme cocktails for the hydrolysis of

286   crude lignocellulose, highlighting the hydrolytic capability of the enzymes in the absence of active fungal

287   growth. Because sugars accumulate primarily until after fungal growth has ceased (Figure 2A-B), this

288   suggests that the most feasible application of a co-culture system would be a two-stage approach. In this

289   strategy, biomass is first incubated with gut fungi to produce excess sugar, which can then be fed to a

290   second model organism for production of a value-added bioproducts.

291

292   *3.3. Biomass degrading enzymes are regulated in response to substrate availability*

293   Anaerobic gut fungi possess a large and diverse suite of biomass degrading enzymes (14, 20, 34,

294   35) that allow them to easily break down crude, lignin-rich biomass. Only a few studies have explored

295   how these genes are regulated in response to changing environmental conditions, such as addition of a

296   catabolite repressor (20) or general substrate availability (36). Based on their varied growth and metabolic

297   capabilities, we hypothesized that different fungal genera rely on specific mechanisms to regulate their

298   biomass degrading enzymes in response to substrate availability. In addition, we sought to identify the

299   environmental conditions that optimized biomass degrading enzyme secretion for use in CBP applications.

300   Overall, the transcriptome of *N. californiae* contained more than twice as many carbohydrate

301   active enzyme (CAZy) domain containing transcripts compared to *A. robustus* (657 compared to 306

302   CAZymes), an observation that aligns with the sizes of the genomes for each of these fungi [Haitjema, in

303   review]. However, the relative functional distribution of these CAZymes is conserved across both species

304   with cellulases, hemicellulases, and accessory enzymes each comprising roughly one third of all CAZymes

305   (Table S3). This conserved balance of functional activities suggests that each function is required in equal

306   proportion to efficiently degrade biomass. We isolated RNA, sequenced with greater than 50X coverage

307   (Table S4 & S5), and analyzed transcript abundance using RSEM (27) to obtain expression counts for all

308   transcripts during growth on glucose, maltose, cellobiose, cellulose, corn stover, reed canary grass, and
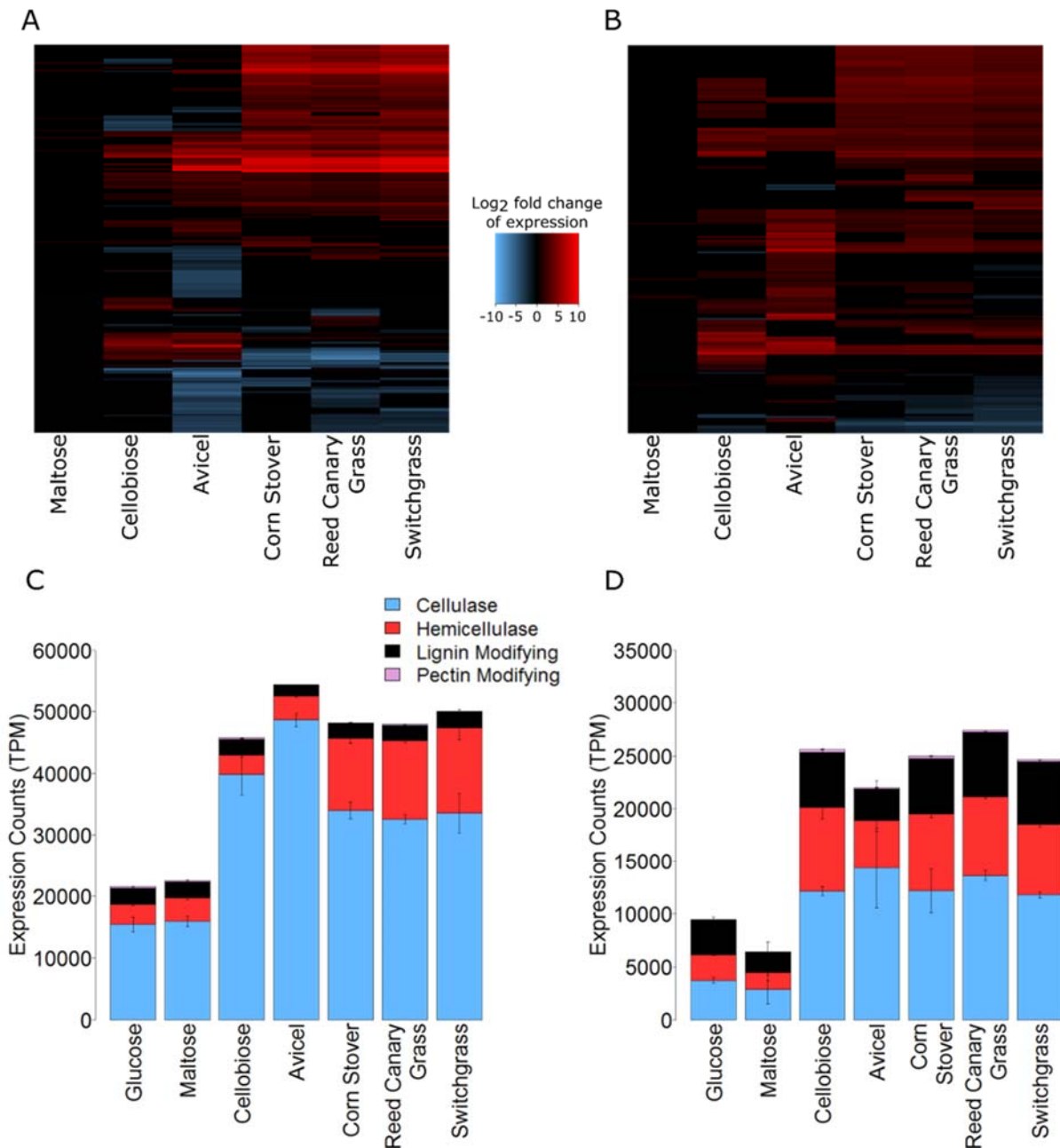
309   switchgrass.

310

311   **Table 1. Summary of up- and down-regulated CAZyme transcripts under different growth conditions compared to basal**
312   **expression on glucose.**

Growth Condition	*A. robustus*		*N. californiae*	
	Down Regulated	Up Regulated	Down Regulated	Up Regulated
Maltose	0	3	0	10
Cellobiose	9	84	36	87
Avicel	4	86	122	124
Corn Stover	11	97	36	168

Reed Canary Grass	19	122	65	177
Switchgrass	34	108	46	168

313

314   Differential expression analysis identified a total of 350 unique CAZymes in *N. californiae* (53% of

315   all CAZymes) and 202 (66%) in *A. robustus* that were significantly regulated (greater than 2-fold change,

316        $p \leq 0.01$) in response to growth on differing substrates compared to glucose. These transcripts were

317   primarily upregulated as substrate complexity increased, though some downregulation was observed

318   (Figure 3) that we expect to be the result of transitioning to more effective CAZymes required to break

319   down complex substrates. Growth on cellobiose, cellulose, and plant biomass triggered large changes in

320   expression of CAZymes, with primarily upregulation of transcripts (Table 1). Only growth of *N. californiae*

321   on Avicel resulted in the downregulation of many CAZyme transcripts, nearly equal to the number

322   upregulated under that condition. There were also many regulated transcripts that contain fungal

323   dockerin (CBM10) domains lacking assigned CAZy functionality; 230 in *N. californiae* and 137 in *A.*

324   *robustus*. While these transcripts cannot be designated as CAZymes, they may play an unknown role in

325   biomass degradation via fungal cellulosomes, representing unclassified carbohydrate active enzymes, or

326   alternate functions involved in improving lignocellulolytic activity of fungal cellulosome complexes.
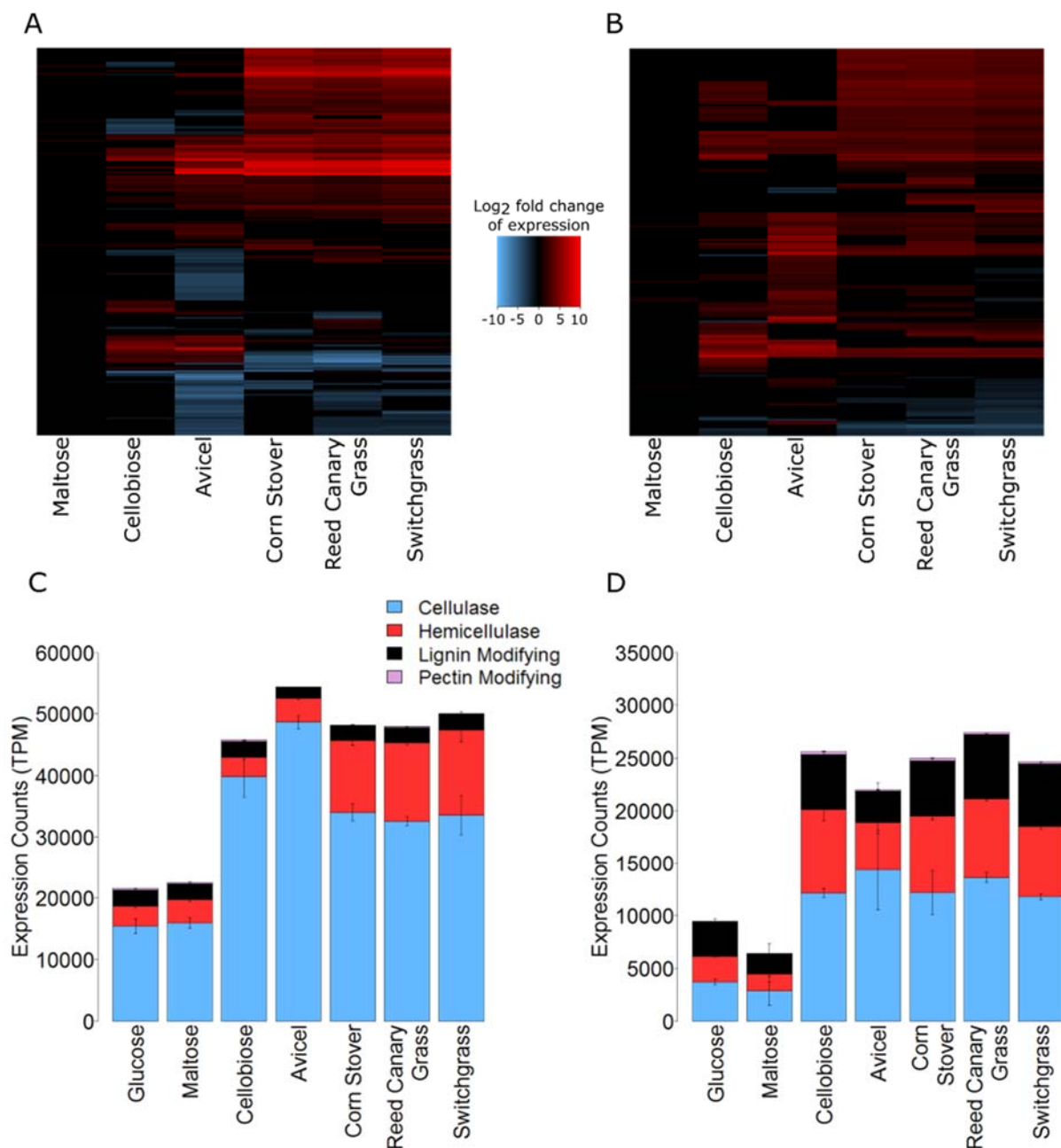
Figure 3. Biomass degrading enzymes of anaerobic fungi are tuned to substrate availability. A and B: Heat maps of the $\log_2$ fold change in expression of biomass degrading enzymes on a variety of indicated substrates compared to basal expression on glucose for *N. californiae* (A) and *A. robustus* (B), respectively. C and D: Normalized expression counts in transcripts per million (TPM) of indicated classes of biomass degrading enzymes under all evaluated subtrate growth conditions for *N. californiae* (C) and *A. robustus* (D), respectively.

We further hypothesized that the overall expression of cellulases, hemicellulases, and accessory enzymes would increase only when their activity was necessary to degrade a given substrate. For example,

336   that hemicellulases were only expressed when hemicellulose was present. This was the case for *N.*

337   *californiae,* with a drastic increase in expression of cellulases on cellobiose and Avicel, yet no change in

338   expression of hemicellulases until hemicellulose was present in biomass substrates. Overall hemicellulase

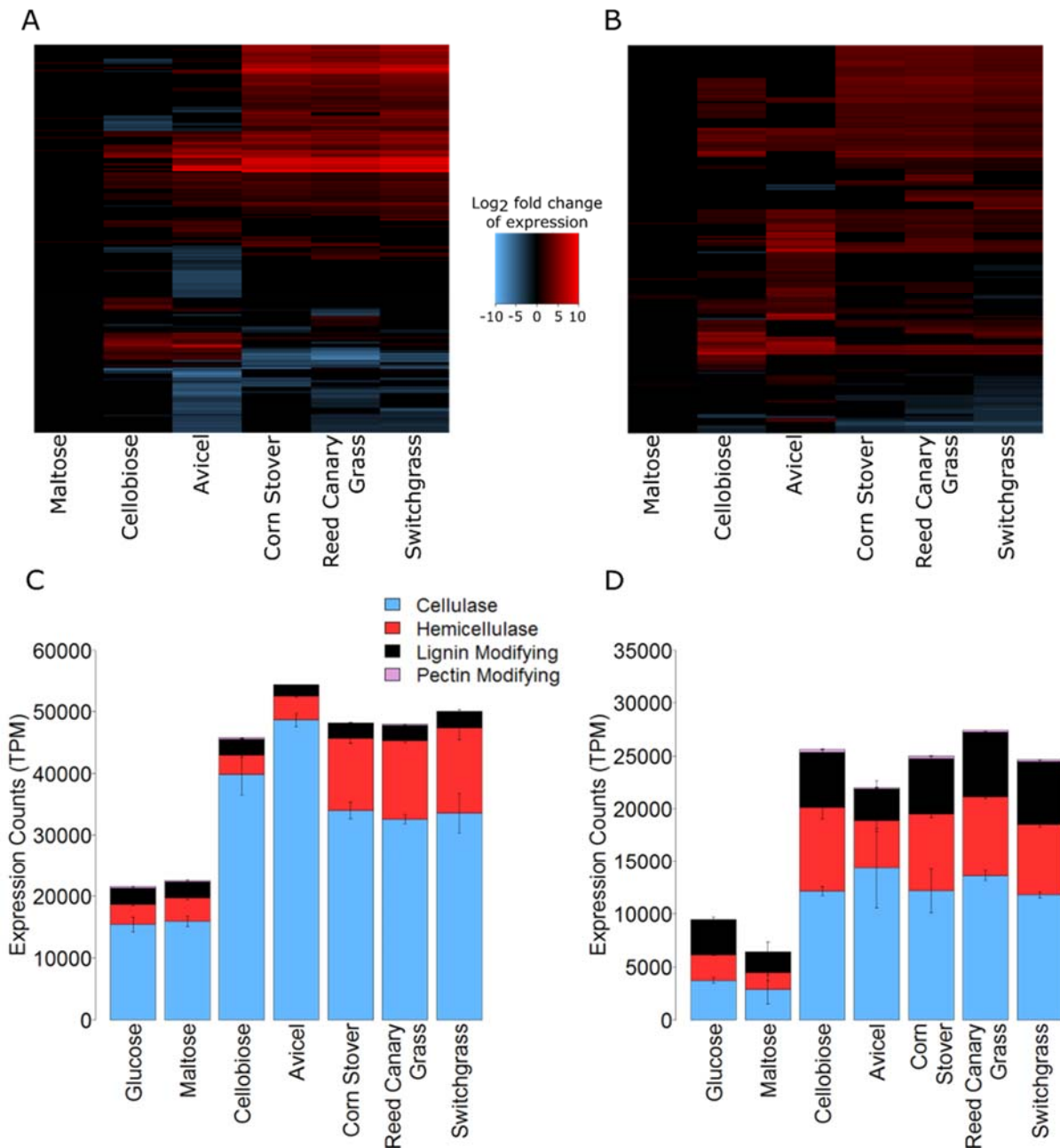339   expression was increased almost 3-fold on reed canary grass as compared to Avicel (

340

341   **Figure 3. Biomass degrading enzymes of anaerobic fungi are tuned to substrate availability**. A and B: Heat maps of
342   the log$_2$ fold change in expression of biomass degrading enzymes on a variety of indicated substrates compared to
343   basal expression on glucose for *N. californiae* (A) and *A. robustus* (B), respectively. C and D: Normalized expression

344 counts in transcripts per million (TPM) of indicated classes of biomass degrading enzymes under all evaluated
345 subtrate growth conditions for *N. californiae* (C) and *A. robustus* (D), respectively.

346

347 3C). This suggests separate mechanisms that rely on different trigger molecules or breakout

348 products to control the expression of cellulases and hemicellulase in *N. californiae*. Alternatively, growth

349 on cellobiose and cellulose, as well as biomass, triggered increased expression of cellulases,

350 hemicellulases, and accessory enzymes in *A. robustus* (

Figure 3. Biomass degrading enzymes of anaerobic fungi are tuned to substrate availability. A and B: Heat maps of the $\log_2$ fold change in expression of biomass degrading enzymes on a variety of indicated substrates compared to basal expression on glucose for *N. californiae* (A) and *A. robustus* (B), respectively. C and D: Normalized expression counts in transcripts per million (TPM) of indicated classes of biomass degrading enzymes under all evaluated substrate growth conditions for *N. californiae* (C) and *A. robustus* (D), respectively.

3D) suggesting a single trigger to regulate all biomass degrading enzymes in this fungal species, a

pattern very different than observed in *N. californiae*.

360   It is important to note that both organisms demonstrated a significant basal expression level of

361   biomass degrading enzymes on glucose, approximately 21,500 and 10,500 TPM (2.15% and 1.05% of total

362   transcriptome expression) in *N. californiae* and *A. robustus*, respectively. This basal activity likely releases

363   break out carbohydrates from lignocellulose, such as cellobiose, that can later trigger increased

364   expression of enzymes required to hydrolyze plant material. In fact, overall expression of CAZymes in both

365   *N. californiae* and *A. robustus* increased most drastically (by greater than 200%) when grown on

366   cellobiose, a low molecular weight cellodextrin, compared to glucose (Figure 3 C&D). This effect revealed

367   that growth of *A. robustus* on cellobiose will induce production of the entire suite of enzymes required to

368   break down crude biomass. Considering that many of these enzymes contain carbohydrate binding

369   domains that keep them tightly bound to lignocellulose, this would allow for simpler purification of

370   enzymes that does not require separation of enzymes from the substrates they act on. Conversely, *N.*

371   *californiae* requires growth on complex biomass to produce all necessary enzymes, making enzyme

372   purification more difficult.

373   Further insight into the regulatory mechanisms of gut fungi can be used to optimize enzyme

374   production and achieve maximum lignocellulolytic activity and sugar handoff to model microorganisms.

375   Possible regulators of biomass degrading enzymes in these gut fungi were previously identified by

376   Solomon et al. (20) by searching for transcripts orthologous to conserved transcription factors in higher

377   fungi, Cre1/CreABC, ACE1-2, ClbR, Clr1-2, and Xyr-1/XlnR that regulate hemicellulase and cellulase

378   production in *Trichoderma reesei*, *Neurospora crassa*, and *Aspergillus niger* (37). A comparison to the

379   current OrthoMCL database in this study identified orthologs to the *creA*, *creB*, *creC*, and Cre-1 regulators

380   from *T. reesei* and *N. crassa* (Table S6). Thus, it is likely that gut fungi possess a similar glucose-based

381   regulation, possibly indicating early evolutionary origin of the CreABC regulatory network. However, only

382   the results for *A. robustus* are consistent with the lack of Xyr-1/XlnR regulators. The lack of orthologs in

383   *N. californiae* may indicate a parallel evolution of this hemicellulase regulation in gut fungi. Solomon et al.

384   identified that glucose concentrations as small as 0.5 g/L (0.05% w/v) can trigger carbon catabolite

385   repression in gut fungi (20). The CreABC regulators are likely candidates for the source of this regulation

386   and knocking them out may alleviate catabolite repression of CAZymes as sugars accumulate during active

387   growth of gut fungi.

388

389   *3.4. Metabolic maps reveal modes of catabolism and opportunities for consolidated bioprocessing*

390   Anaerobic gut fungi are capable of releasing sugars from both cellulose and hemicellulose (Figure

391   4B), but growth experiments (Table S1) revealed that they did not metabolize some of these sugars in

392   monoculture. Using transcriptome annotations, including Enzyme Commission (EC) numbers metabolic

393   maps were built based on KEGG pathways (22, 38) to highlight gaps in sugar catabolism that provide

394   opportunities for co-culture via sugar exchange (Figure 4A).
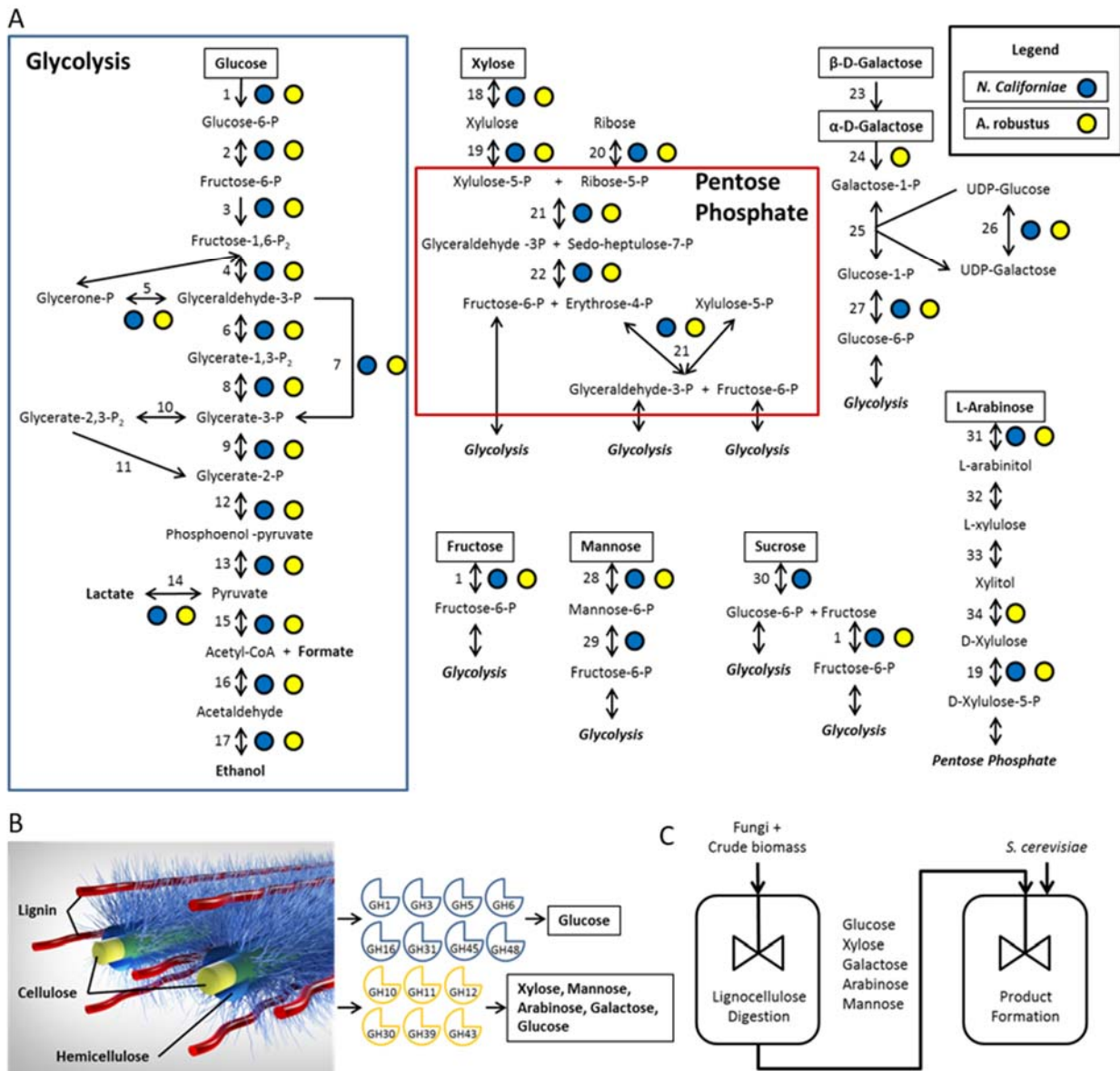
Figure 4. Metabolic reconstruction identifies complete sugar catabolic pathways for biomass derived sugars in anaerobic fungi. A) Enzymatic steps in the indicated pathway are identified as present in each of the fungi. Dots indicate enzymes identified in the transcriptomes of *N. californiae* (blue) and *A. robustus* (yellow). All enzymes were identified via EC number, except for enzyme 9 identified by BLAST annotation. Additional enzyme information is available in the supplementary information (Table S7). Both fungi are capable of glycolysis (with one enzyme identified by BLAST annotation), xylose, and fructose metabolism. Neither fungus contained the necessary enzymes to metabolize arabinose and galactose, and only *Neocallimastix californiae* is capable of metabolizing mannose and sucrose. **B)** Cellulases and hemicellulases release sugar-rich hydrolysates from lignocellulosic biomass. **C)** Two-stage culture system where fungi are used to break down biomass and release sugar that can be fed to a production organism, such as *S. cerevisiae*, in a second step.

407   This analysis revealed a complete glycolytic pathway (Figure 4A), but EC annotations identified

408   two missing enzymes necessary for complete gluconeogenesis, fructose bisphosphatase (EC:3.1.3.11) and

409   glucose-6-phosphatase (EC:3.1.3.9). This result is consistent with previous observations in other gut fungi

410   that gluconeogenesis is incomplete (35). Complete fructose and xylose metabolism was also identified in

411   both fungi, while only *N. califonorniae* contained mannose and sucrose catabolism, and both fungi lacked

412   the enzymes required for galactose and arabinose catabolism. Xylose metabolism follows the xylose

413   isomerase pathway typical of prokaryotes (39); an observation consistent with previous findings for

414   *Piromyces* sp. E2 (31). This pathway may have arisen from horizontal gene transfer in the rumen

415   microbiome and lead to increased fitness over the eukaryotic oxido-reductase pathway. The anaerobic,

416   reducing environment of the gut is likely to upset the redox balance of the oxido-reductase pathway

417   resulting in accumulation of xylitol, while the xylose isomerase pathway is less affected by anaerobic

418   conditions (39, 40). While most of these observations are corroborated by growth experiments (Table S1),

419   the presence of xylose catabolism conflicts with growth experiments that revealed these gut fungi do not

420   thrive on the pentose sugar (Table S1). This discrepancy between transcriptomic and growth experiment

421   observations suggests another limitation is responsible for lack of xylose utilization, such as inefficient

422   transport or lack of environmental influences not present in these experiments.

423   Downstream, the enzymes required for lactate and ethanol production from pyruvate were

424   identified in both organisms (Figure 4A), yielding formate as a side product. Energy generation in gut fungi

425   primarily relies on the hydrogenosome organelle (41) also found in members of the *Trichomonas* genus

426   and some anaerobic protists (42). Here, we identified a complete hydrogenosomal pathway that takes

427   malate or pyruvate as inputs and produces ATP, acetate, formate, and molecular hydrogen as products

428   (Figure S5). This pathway contains soluble components of mitochondrial complex I, NADH:ubiquinone

429   oxidoreductase (EC:1.6.5.3) that were also strong homologs to the *Trichomonas vaginalis* enzymes NuoF

430   and Nuo E (E $\leq 10^{-150}$; Table S8). These enzymes regenerate NAD$^+$ for conversion of malate to pyruvate by

431   transferring electrons to ferredoxin (43). ATP is produced primarily in the regeneration of succinate and

432   CoA from succinyl CoA and molecular hydrogen is produced during the oxidation of ferredoxin by a

433   hydrogenase (Figure S5).

434

435   While glucose and fructose were the most abundant sugars remaining after biomass hydrolysis,

436   xylose and arabinose were also present in the hydrolysate (Figure 2 A-B). Glucose and fructose likely

437   accumulated due to an overabundance in the biomass, but xylose and arabinose likely accumulated due

438   to a lack of necessary enzymes (Figure 4A), or some other limitation that prevented assimilation by the

439   gut fungi. As these sugars primarily accumulated after fungal growth had ceased, we tested a two-stage

440   production system where fungi digest biomass in the first step and the hydrolysate supports the growth

441   of *S. cerevisiae* in the second (Figure 4C).

442

443   *3.5. Two-step co-culture reveals potential for gut fungi in bio-based production*

444   Following growth of anaerobic fungi on biomass, the "spent" fungal media was inoculated with

445   *Saccharomyces cerevisiae* (Figure 2C) to determine if the carbohydrate-rich fungal hydrolysate was

446   capable of supporting yeast proliferation. Spent media from growth on crystalline cellulose, containing 6-

447   7 g/L of glucose, supported growth of *S. cerevisiae* to saturation, with an OD_{600} of 14 while fresh media

448   containing no fungal hydrolysate grew to a negligible OD_{600} (Figure 2C). This demonstrates that the fungi

449   were capable of hydrolyzing enough excess sugar to support growth of *S. cerevisiae* and did not produce

450   any compounds that significantly inhibited yeast growth. *Escherichia coli* was also tested for growth on

451   media from fungal cultures grown on cellulose, resulting in a small increase in optical density compared

452   to the control case, again indicating no inhibitory compounds were produced by the fungi (Figure S6).

453   Biomass hydrolysate from fungal growth on crude reed canary grass was then tested for support of *S.*

454   *cerevisiae*. While the amount of glucose released from reed canary grass was much lower compared to

455   that released from cellulose (Figure 2A-B), the yeast reached a similar optical density (Figure 2C) when

456   grown on this media relative to a control. Measurements of sugar concentrations before and after yeast

457   growth (Figure 2D) revealed that the yeast consumed primarily glucose and fructose present in the fungal

458   media, as expected. There was a reduction in overall sugars of 79% and 73% after yeast growth in *N.*

459   *californiae* and *A. robustus* media, respectively, leaving primarily xylose and arabinose to accumulate in

460   the culture media.

461   The above results demonstrate the feasibility of a two-stage production process, with a wealth of

462   sugars released from biomass by anaerobic gut fungi that may be exchanged with another organism or

463   combination of organism specialists. Further, the extent to which the yeast can remove the excess glucose

464   and fructose suggests that carbon catabolite repression may be alleviated by the presence of another

465   organism during a simultaneous co-culture, increasing overall production of enzymes while improving

466   enzyme efficiency by removing sugar-based inhibition of enzymes. Previous studies on microbial co-

467   cultures and consortia for production have paired cellulolytic organisms, such as *Clostridium*

468   *phytofermentans* (44), with production organisms, requiring cellulose as an input rather than biomass.

469   *Trichoderma reesei* and *E. coli* have also been paired for production of isobutanol from biomass, but still

470   rely on the use of pretreated biomass (5). In contrast, gut fungi are capable of supplying sugars directly

471   from biomass without any pretreatment. Furthermore, pairing to growth of *T. reesei* limits production to

472   aerobic conditions, while the two-stage system proposed here is amenable to both anaerobic and aerobic

473   production conditions, tailoring the process to the desired product as needed.

474

475   **4. Conclusions**

476   Anaerobic gut fungi efficiently hydrolyze crude lignocellulose through a combination of

477   mechanical disruption and enzymatic activity from a wide array of biomass degrading enzymes that

478   release excess amounts of sugars into their environment. Reconstruction of metabolic pathways in

479   anaerobic fungi validated experimental phenotypes, and identified sugars that are likely to accumulate

480   from biomass hydrolysis alongside the most abundant component, glucose. While these breakout

481   products are available to exchange with other rumen microbes in their native system, here sugars can be

482   harnessed to support growth of industrially-relevant microbes not native to the rumen. Here, we have

483   demonstrated the ability of the fungal hydrolysate to support growth of the model organism, *S. cerevisiae*,

484   presenting a consolidated bioprocessing strategy that utilizes crude, rather than pretreated, biomass for

485   direct biochemical production. While additional work will be necessary to optimize the bi-phasic

486   fermentation scheme, these regulation studies provide a path forward for bolstering production of

487   biomass degrading enzymes, as well as identifying potential repressors of their production. The two-stage

488   fermentation approach presented here allows for the consolidation of biomass pretreatment and

489   hydrolysis into a single step to supply a monosaccharide-rich hydrolysate that can be donated to a model

490   organism for growth and production. In this way, the second fermentation step allows for the precise

491   control of the production bioreactor such that conditions can be optimized for the desired product rather

492   than fungal growth.

493

511

512
513

References

514

515   1.   Sanderson K (2006) US biofuels a field in ferment. *Nature* 444(December):673-676.

516   2.   Agbor VB, Cicek N, Sparling R, Berlin A, & Levin DB (2011) Biomass pretreatment: fundamentals
517   toward application. *Biotechnology advances* 29:675-685.

518   3.   Balan V, Chiaramonti D, & Kumar S (2013) Review of US and EU initiatives toward development,
519   demonstration, and commercialization of lignocellulosic biofuels. *Biofuels, Bioproducts and*
520   *Biorefining* 7(6):732-759.

521   4.   Olson DG, McBride JE, Shaw AJ, & Lynd LR (2012) Recent progress in consolidated bioprocessing.
522   *Curr Opin Biotechnol* 23(3):396-405.

523   5.   Minty JJ*, et al.* (2013) Design and characterization of synthetic fungal-bacterial consortia for direct
524   production of isobutanol from cellulosic biomass. *Proceedings of the National Academy of*
525   *Sciences of the United States of America* 110(36):14592-14597.

526   6.   Agapakis CM, Boyle PM, & Silver PA (2012) Natural strategies for the spatial optimization of
527   metabolism in synthetic biology. *Nat Chem Biol* 8(6):527-535.

528   7.   Zhou K, Qiao K, Edgar S, & Stephanopoulos G (2015) Distributing a metabolic pathway among a
529   microbial consortium enhances production of natural products. *Nat Biotechnol* 33(4):377-383.

530   8.   Zuroff TR & Curtis WR (2012) Developing symbiotic consortia for lignocellulosic biofuel
531   production. *Appl Microbiol Biotechnol* 93(4):1423-1435.

532   9.   Peng XN, Gilmore SP, & O'Malley MA (2016) Microbial communities for bioprocessing: lessons
533   learned from nature. *Current Opinion in Chemical Engineering* 14:103-109.

534   10.   Flint HJ (1997) The rumen microbial ecosystem-some recent developments. *Trends in*
535   *Microbiology* 5(12):483-488.

536   11.   Windham WRA, D E (1984) Rumen fungi and forage fiber degradation. *Applied and Environmental*
537   *Microbiology* 48(3):473-476.

538   12.   Nicholson MJ, Theodorou MK, & Brookman JL (2005) Molecular analysis of the anaerobic rumen
539   fungus Orpinomyces - insights into an AT-rich genome. *Microbiology* 151(Pt 1):121-133.

540   13.   Bootten TJ, Joblin KN, McArdle BH, & Harris PJ (2011) Degradation of lignified secondary cell walls
541   of lucerne (Medicago sativa L.) by rumen fungi growing in methanogenic co-culture. *J Appl*
542   *Microbiol* 111(5):1086-1096.

543   14.   Ljungdahl LG (2008) The cellulase/hemicellulase system of the anaerobic fungus Orpinomyces PC-
544   2 and aspects of its applied use. *Ann N Y Acad Sci* 1125:308-321.

545   15.   Haitjema CH, Solomon KV, Henske JK, Theodorou MK, & O'Malley MA (2014) Anaerobic gut fungi:
546   Advances in isolation, culture, and cellulolytic enzyme discovery for biofuel production.
547   *Biotechnol Bioeng* 111(8):1471-1482.

548   16.   Bauchop TM, Douglas O. (1981) Cellulose Ferementation by a Rumen Anaerobic Fungus in Both
549   the Absence and Presence of Rumen Methanogens. *Applied and Environmental Microbiology*
550   42(6):1103-1110.

551   17.   Li GJ*, et al.* (2016) Fungal diversity notes 253–366: taxonomic and phylogenetic contributions to
552   fungal taxa. *Fungal Divers.* 78.

553   18.   Theodorou MK, Brookman J, & Trinci APJ (2005) Anaerobic Fungi. *Methods in Gut Microbial*
554   *Ecology for Ruminants*, eds Makkar HPS & McSweeney CS (IAEA, Netherlands), pp 55-56.

555   19.   Theodorou MK, Davies DR, Nielsen BB, Lawrence MIG, & Trinci APJ (1995) Determination of
556   growth of anaerobic fungi on soluble and cellulosic substrates using a pressure transducer.
557   *Microbiology* 141(3):671-678.

558   20.   Solomon KV*, et al.* (2016) Early-branching gut fungi possess a large, comprehensive array of
559   biomass-degrading enzymes. *Science* 351.

560   21.   Bairoch A (2000) The ENZYME database in 2000. *Nucleic Acids Research* 28(1):304-305.

561   22.   Kanehisa M & Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids*
562   *Research* 28(1):27-30.
563   23.   Altschul SF, Gish W, Miller W, Myers EW, & Lipman DJ (1990) Basic local alignment search tool. *J*
564   *Mol Biol* 215.
565   24.   Finn RD*, et al.* (2017) InterPro in 2017-beyond protein family and domain annotations. *Nucleic*
566   *Acids Res* 45(D1):D190-D199.
567   25.   Grabherr MG*, et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a
568   reference genome. *Nat Biotechnol* 29(7):644-652.
569   26.   Love MI, Huber W, & Anders S (2014) Moderated estimation of fold change and dispersion for
570   RNA-seq data with DESeq2. *Genome Biology* 15(12):1-21.
571   27.   Li B & Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or
572   without a reference genome. *BMC Bioinformatics* 12:323.
573   28.   Trinci APJ*, et al.* (1994) Anaerobic fungi in herbivorous animals. *Mycological Research* 98(2):129-
574   152.
575   29.   Leis S*, et al.* (2014) Finding a robust strain for biomethanation: anaerobic fungi
576   (Neocallimastigomycota) from the Alpine ibex (Capra ibex) and their associated methanogens.
577   *Anaerobe* 29:34-43.
578   30.   Theodorou MK*, et al.* (1996) Anaerobic fungi in the digestive tract of mammalian herbivores and
579   their potential for exploitation. *Proc Nutr Soc* 55.
580   31.   Harhangi HR*, et al.* (2003) Xylose metabolism in the anaerobic fungus Piromyces sp. strain E2
581   follows the bacterial pathway. *Arch Microbiol* 180(2):134-141.
582   32.   Dien B*, et al.* (2006) Chemical composition and response to dilute-acid pretreatment and
583   enzymatic saccharification of alfalfa, reed canarygrass, and switchgrass. *Biomass and Bioenergy*
584   30(10):880-891.
585   33.   Seppälä S, Solomon KV, Gilmore SP, Henske JK, & O'Malley MA (2016) Mapping the membrane
586   proteome of anaerobic gut fungi identifies a wealth of carbohydrate binding proteins and
587   transporters. *Microbial Cell Factories* 15(1):212.
588   34.   Ali BR*, et al.* (1995) Cellulases and hemicellulases of the anaerobic fungus Piromyces constitute a
589   multiprotein cellulose-binding complex and are encoded by multigene families. *FEMS*
590   *Microbioogy Letters* 125(1):15-21.
591   35.   Youssef NH*, et al.* (2013) The genome of the anaerobic fungus Orpinomyces sp. strain C1A reveals
592   the unique evolutionary history of a remarkable plant biomass degrader. *Appl Environ Microbiol*
593   79(15):4620-4634.
594   36.   Couger MB, Youssef NH, Struchtemeyer CG, Liggenstoffer AS, & Elshahed MS (2015)
595   Transcriptomic analysis of lignocellulosic biomass degradation by the anaerobic fungal isolate
596   Orpinomyces sp. strain C1A. *Biotechnol Biofuels* 8:208.
597   37.   Amore A, Giacobbe S, & Faraco V (2013) Regulation of Cellulase and Hemicellulase Gene
598   Expression in Fungi. *Current Genomics* 14(4):230-249.
599   38.   Kanehisa M*, et al.* (2014) Data, information, knowledge and principle: back to metabolism in
600   KEGG. *Nucleic Acids Res* 42(Database issue):D199-205.
601   39.   Jeffries TW (1983) Utilization of xylose by bacteria, yeasts, and fungi. *Advances in Biochemical*
602   *Engineering/Biotechnology* 27:1-32.
603   40.   Bruinenberg PM, de Bot PHM, van Dijken JP, & Scheffers WA (1983) The role of redox balances in
604   the anaerobic fermentation of xylose by yeasts. *European Journal of Applied Microbiology and*
605   *Biotechnology* 18(5):287-292.
606   41.   Yarlett N, Orpin CG, Munn EA, Yarlett NC, & Greenwood CA (1986) Hydrogenosomes in the rumen
607   fungus Neocallimastix patriciarum. *The Biochemical Journal* 236(3):729-739.

608   42.   Hackstein JHP, Akhmanova A, Boxma B, Harhangi HR, & Voncken FGJ (1999) Hydrogenosomes:
609   Eukaryotic adaptations to anaerobic environments. *Trends in Microbiology* 7(11):441-447.
610   43.   Hrdý I*, et al.* (2004) Trichomonas hydrogenosomes contain the NADH dehydrogenase module of
611   mitochondrial complex I. *Nature* 432(7017):618-622.
612   44.   Zuroff TR, Xiques SB, & Curtis WR (2013) Consortia-mediated bioprocessing of cellulose to ethanol
613   with a symbiotic Clostridium phytofermentans/yeast co-culture. *Biotechnology for Biofuels*
614   6(1):59-70.
615