

# UC Santa Barbara

## UC Santa Barbara Electronic Theses and Dissertations

### Title

Model-driven design of anaerobic consortia for lignocellulose valorization

### Permalink

<https://escholarship.org/uc/item/612708j9>

### Author

Leggieri, Patrick A

### Publication Date

2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Santa Barbara

Model-driven design of anaerobic consortia for lignocellulose valorization

A dissertation submitted in partial satisfaction of the  
requirements for the degree Doctor of Philosophy  
in Chemical Engineering

by

Patrick A. Leggieri

Committee in charge:

Professor Michelle A. O'Malley, Chair

Professor Megan T. Valentine

Professor M. Scott Shell

Professor Arnab Mukherjee

Professor David L. Valentine

June 2023

The dissertation of Patrick A. Leggieri is approved.

---

David L. Valentine

---

Arnab Mukherjee

---

M. Scott Shell

---

Megan T. Valentine

---

Michelle O'Malley, Committee Chair

May 2023

Model-driven design of anaerobic consortia for lignocellulose valorization

Copyright © 2023

by

Patrick A. Leggieri

## ACKNOWLEDGEMENTS

I want to thank the people in my life who contributed to this work by giving me so many things to enjoy and look forward to. It's easier to work hard and get things done when there's a hockey game or a GDL tourney at the end of the day. This dissertation is dedicated to them.

To my girlfriend (for now), Candy, the most likable person I've ever met. Thanks for competing and having fun with me, and being the only person I want to see every day.

To my brother, Tom, you called that shot about the "beginning." I definitely could've done this without you, but I'm glad I didn't have to. Now let's get to work for real.

To my family: the Leggieris, Jacksons, and Macabuhays. Nan, Gram, AJ, and Kev, thanks for helping to form me into a person who could come out here on my own and create a life for myself. Quijadases, Milinkoviches, Mikey and the Macabuhays, thanks for making southern California feel like home for the kid in the Steelers trunks. We'll be back.

To Ben, Cory, Mrs. Skeldon, Professors Soh and Senra, and everyone else who helped me learn to walk in science. To Michelle for the freedom to carve this thing out in my own directions. To Megan for the advice along the way. To St. Elmo for helping me get started, and Elaina for helping me finish up.

To Jackie, Niko, the Duffields, and the Yetis, the single coolest thing I have ever been a part of. Nothing has ever meant more to me than being your captain. Thank you for buying in, skating hard, and making me a better player, leader, and person. You aren't like family to me; you are family. This is for you.

## VITA OF PATRICK A. LEGGIERI

May 2023

### EDUCATION

Bachelor of Science in Chemical Engineering, Lafayette College, May 2018 (Summa Cum Laude, Honors in Chemical Engineering)

Doctor of Philosophy in Chemical Engineering, University of California, Santa Barbara, May 2023 (expected)

### PROFESSIONAL EMPLOYMENT

2015: Fireworks Sales, Odd Lot Outlet

2015-2018: Peer Tutor, Lafayette College

2016-2017: EXCEL Scholar Undergraduate Researcher, Lafayette College

2017: NSF Research Experience for Undergraduates, Johns Hopkins University

2019-2021: Teaching Assistant, Department of Chemical Engineering, University of California, Santa Barbara

2019-2023: Graduate Student Researcher, Department of Chemical Engineering, University of California, Santa Barbara

2023: Fermentation Scientist, Allonnia, Boston, MA

### PUBLICATIONS

**P.A. Leggieri**, E.M. Blair, T.S. Lankiewicz, M.A. O'Malley, *Continuous culture of anaerobic fungi enables growth and metabolic flux tuning without the use of genetic tools*, Submitted.

T.S. Lankiewicz, H. Choudhary, Y. Gao, B. Amer, S.P. Lillington, **P.A. Leggieri**, J.L. Brown, C.L. Swift, A. Lipzen, H. Na, M.K. Theodorou, E.K. Baidoo, K. Barry, I.V. Grigoriev, V.I. Tomhokin, J. Gladden, S. Singh, J.C. Mortimer, J. Ralph, B.A. Simmons, S.W. Singer, M.A. O'Malley, *Lignin deconstruction by anaerobic fungi*, *Nature Microbiology*, 8, 2023

**P.A. Leggieri**, M.T. Valentine, M.A. O'Malley, *Biofilm disruption enhances growth rate and CAZyme production in anaerobic fungi*, *Bioresource Technology*, 358, 127361 (2022)

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

**P. A. Leggieri**, Y. Liu, M. Hayes, B.A. Connors, S. Seppälä, M.A. O'Malley, O.S. Venturelli, *Integrating systems and synthetic biology to understand and engineer microbiomes*, *Annual Review of Biomedical Engineering*, Vol. 23 (2021)

S.E. Wilken, J.M. Monk, **P.A. Leggieri**, C. Lawson, T.S. Lankiewicz, S. Seppala, C.G. Daum, J. Jenkins, A.M. Lipzen, S.J. Mondo, K.W. Barry, I.V. Grigoriev, J.K. Henske, M.K. Theodorou, B.O. Palsson, L.R. Petzold, M.A. O'Malley, *Experimentally validated reconstruction and analysis of a genome-scale metabolic model of the Neocallimastigomycota fungus Neocallimastix lanati*, mSystems, 6, 1 (2021)

S.E. Wilken, **P.A. Leggieri**, C. Kerdman-Andrade, M. Reilly, M.K. Theodorou, M.A. O'Malley, *Design and construction of an Arduino-based Pressure Evaluation System (A-APES) to quantify growth of non-model anaerobes in culture*, AIChE Journal, 66, e16540 (2020)

S.P. Lillington, **P.A. Leggieri**, K.A. Heom, M.A. O'Malley, *Nature's recyclers: anaerobic microbial communities drive crude biomass deconstruction*, Current Opinion in Biotechnology, 62, 38-47 (2020)

**P.A. Leggieri**, M. Senra, and L. Soh, *Cloud point and crystallization in fatty acid ethyl ester biodiesel mixtures with and without additives*, Fuel, 222, 243-249 (2018)

## AWARDS

Schlinger Fellowship for Excellence in Chemical Engineering Research, UCSB 2022  
Best Oral Presentation- 15<sup>th</sup> Amgen-Clorox Graduate Student Symposium, UCSB 2022  
Graduate Division Dissertation Fellowship, UCSB 2022  
Graduate Research Mentorship Program Fellowship, UCSB 2020  
Tau Beta Pi Fellowship 2018  
NSF GRFP Honorable Mention 2018  
National Academy of Engineering Grand Challenges Scholar 2018  
Carl Staska Prize, Lafayette College 2018  
Donald Noblett Prize, Lafayette College 2018  
Goldwater Scholar 2017  
Tau Beta Pi Scholarship 2017  
Dr. E.L. McMillen-K.K. Malhotra '49 Chemical Engineering Prize 2017  
S-STEM National Science Foundation Grant 2015-2017  
ACS Organic Chemistry Scholarship 2015  
Salvation Army Community Service Scholarship 2015

## ABSTRACT

Model-driven design of anaerobic consortia for lignocellulose valorization

by

Patrick A. Leggieri

It is urgent that we develop and implement bio-based alternatives to our existing petroleum chemical and energy infrastructure. Lignocellulose is the most abundant renewable carbon resource on the planet, making it a promising biofeedstock. However, its recalcitrance to degradation *via* chemicals and model microbes like *E. coli* and yeast precludes its utilization for bioenergy or production of commodity and specialty chemicals. So-called “non-model” anaerobic microbial consortia found in the rumen of large herbivores have evolved as specialized biomass degraders and have potential for lignocellulose-based bioproduction if they can be onboarded, characterized, and deployed at scale.

Anaerobic gut fungi (AGF) in rumen consortia produce nature’s greatest known variety and abundance of lignocellulose-degrading carbohydrate-active enzymes (CAZymes). This, combined with their mixed-acid fermentation profile, makes them interesting candidates for industrial CAZyme production and/or biomass deconstruction and conversion. However, AGF are not genetically tractable, and their physiology and primary metabolism are poorly understood, which limits our ability to predict and manipulate phenotypes for user-specified culture outcomes. AGF could therefore be deployed in communities alongside

genetically tractable workhorse strains, wherein AGF specialize in degradation of lignocellulose to sugars and conversion to bioproduct precursors.

Microbial communities are, in principle, capable of virtually limitless chemical transformations. In practice, designing consortia with predictable, prescribed functions is challenging, especially with a largely uncharacterized constituent species such as an AGF. Before we can deploy anaerobic consortia industrially, we must understand AGF physiology and metabolism. Specifically, we must know the entire space of achievable AGF phenotypes and how to accentuate the functions that we desire (fast growth, production of CAZymes, high flux of certain metabolites, etc.).

Without genetic tools, we require creative and multifaceted approaches to characterize and tune AGF growth and metabolism. Toward this goal, we synthesized multi-omic and biochemical data into the first AGF genome-scale metabolic model, offering the most complete description of AGF growth and metabolism available. The model established the theoretical AGF phenotype space; from there, we exposed AGF to myriad culture conditions (some resembling their natural habitat and some more artificial) to explore which phenotypes are both biotechnologically useful and achievable in practice.

Using a non-rhizoidal AGF, *Caecomyces churrovis*, we developed simple, yet vital methods for quantification of AGF growth and metabolic flux that are routine in model systems but have been unavailable to AGF. By stirring *C. churrovis* cultures, we elicited a suspended culture morphology that grows faster and expresses significantly more CAZymes per cell than typical biofilm cultures. We leveraged these well-mixed suspended cultures to develop methods for non-destructive quantification of AGF growth and flux in co-culture with prokaryotes, and showed that methanogens significantly increased AGF growth rate and

altered AGF metabolic flux to yield different fermentation product profiles. In a significant step towards industrial deployment of AGF, we demonstrated the first steady state continuous culture of AGF using a DIY Arduino-based continuous flow bioreactor. Turbidostat bioreactor operation uncovered relationships between setpoint titer and AGF growth rate and flux, enabling users to specify continuous production rates of target metabolites and enzymes and vary them depending on the application at hand.

Our understanding of AGF physiology remains far from comprehensive. However, the research presented in this dissertation has elucidated design rules for AGF cultures with measurable, predictable, and tunable growth and metabolite production rates, moving us closer to deployment of AGF and anaerobic consortia for industrial lignocellulose valorization.

## TABLE OF CONTENTS

I.	Characterization and modeling of microbial communities and anaerobic fungi...1	
A.	Introduction.....	2
B.	Tools for understanding microbiomes.....	4
C.	Computational models predict microbiome dynamics and functions...15	
D.	Deploying lignocellulose-valorizing microbial consortia with anaerobic fungi requires detailed understanding of AGF physiology and metabolism...23	
II.	Biofilm disruption enhances growth rate and carbohydrate-active enzyme production in anaerobic fungi.....	33
A.	Introduction.....	34
B.	Materials and Methods .....	37
C.	Results and Discussion .....	43
D.	Conclusions.....	53
III.	Non-destructive quantification of anaerobic gut fungi and methanogens in co-culture reveals increased fungal growth rate and changes in metabolic flux relative to mono-culture.....	54
A.	Introduction.....	55
B.	Materials and Methods .....	60
C.	Results and Discussion .....	64
D.	Conclusions.....	83
IV.	Continuous culture of anaerobic fungi enables growth and metabolic flux tuning without the use of genetic tools.....	84

A.	Introduction.....	85
B.	Materials and Methods .....	90
C.	Results and Discussion .....	93
D.	Conclusions.....	108
V.	Conclusions, perspectives, outlook.....	109
VI.	References .....	115
VII.	Appendices .....	148
A.	Appendix: Supplementary Materials for Chapter II.....	148
B.	Appendix: Supplementary Materials for Chapter III .....	151
C.	Appendix: Supplementary Materials for Chapter IV .....	157
D.	Appendix: Co-cultures of <i>C. churrovis</i> and <i>E. coli</i> ZSC113 .....	162
E.	Appendix: Proteomic analysis of <i>C. churrovis</i> cultures and hydrogenosome enrichments and <i>N. lanati</i> cultures and zoospore enrichments.....	166
F.	Appendix: Live-cell AGF Microscopy .....	170

## **I. Characterization and modeling of microbial communities and anaerobic fungi**

Reprinted with permission from (Leggieri et al., 2021b).

Microbiomes are complex and ubiquitous networks of microorganisms whose seemingly limitless chemical transformations could be harnessed to benefit agriculture, medicine, and biotechnology. The spatial and temporal changes in microbiome composition and function are influenced by a multitude of molecular and ecological factors. This complexity yields both versatility and challenges in designing synthetic microbiomes and perturbing natural microbiomes in controlled, predictable ways. In this section, we describe factors that give rise to emergent temporal and spatial microbiome properties and the meta-omics and computational modeling tools that can be used to understand microbiomes on the cellular and system-levels. We describe strategies for designing and engineering microbiomes to enhance or build novel functions. Throughout, we discuss key knowledge and technology gaps for elucidating the networks and deciphering key control points for microbiome engineering, and highlight examples where multiple omics and/or modeling approaches can be integrated to address these gaps.

A microbiome system with particular biotechnological promise is the lignocellulolytic microbial consortium native to the rumen of large herbivores. These communities have evolved to degrade crude lignocellulosic biomass, Earth's most abundant renewable carbon resource (Abdel-Hamid et al., 2013), which is highly recalcitrant to degradation *via* chemicals and model microbes. Anaerobic consortia could be leveraged to valorize lignocellulose industrially if they can be onboarded, characterized, and cultivated at scale. Within these consortia, alongside myriad prokaryotes, anaerobic gut fungi (AGF)

produce nature's greatest known variety and abundance of biomass-degrading carbohydrate-active enzymes (CAZymes), making them interesting candidates for lignocellulose-based bioproduction. The challenges of designing and deploying microbiomes are particularly daunting for systems that include AGF, as AGF physiology and metabolism are poorly understood, increasing the degrees of freedom in microbiome design space to intractable levels. To enable design and deployment of microbial communities with AGF for lignocellulose valorization, tools must be developed and implemented to characterize AGF growth and metabolism comprehensively.

## **A. Introduction**

Diverse communities of microorganisms inhabit every known environment, including oceans, soil, the surface and proximity of plants and intestines of humans, animals, and insects. Owing to the seemingly unbounded range of functions they perform, from biogeochemical cycling of nutrients to the transformation of dietary substrates into nutrients for multicellular hosts, microbiomes attract immense attention from industry and academic researchers alike. Efforts to understand and engineer microbiomes frequently require integrated approaches that blur the lines between microbiology, ecology, medicine, computer science, mathematics and engineering. Natural and synthetic microbiomes that robustly perform target functions could be exploited to address grand challenges facing society in human health, agriculture, bioremediation, and bioprocessing.

Target microbiome engineering goals include the ability to predictably modulate community composition, enhance existing functions or install novel capabilities. Harnessing the properties of microbiomes remains difficult, because we do not yet fully understand the molecular and ecological mechanisms that govern systems-level behaviors and therefore lack

the capability to predict their multi-functional properties. Microbiomes are immensely complex since they can consist of hundreds to thousands of organisms, exhibit temporal and spatial variability and can establish dynamic feedback loops with the environment. A detailed and quantitative understanding of microbiomes could ultimately inform the design of interventions to predictably modify system properties or guiding principles for how to construct desired community functions from the bottom up. Exploiting and understanding the full functional potential of microbiomes necessitates integration of multiple experimental and computational methodologies that bridges many different disciplines.

Microbiomes can be studied and engineered through two different and complementary approaches: top-down and bottom-up (Gilmore et al., 2019; Lawson et al., 2019; Peng et al., 2016). A top-down approach investigates natural communities by introducing them into highly controlled laboratory environments. Such top-down manipulations can be used to understand microbiome dynamics and functions in response to environmental inputs (e.g., nutrient availability or antibiotic stress) (Scarborough et al., 2018). By contrast, isolated species can also be assembled *in vitro* to form synthetic communities, which have reduced complexity compared to natural microbiomes and greater controllability via manipulation of initial community composition (Gutiérrez and Garrido, 2019; Venturelli et al., 2018). While molecular and ecological mechanisms of synthetic communities can be more easily dissected, these simplified systems can display reduced temporal stability in composition and/or function, limiting their deployment in real-world environments and for biotechnological applications (Gilmore et al., 2019). At the core, if we understand the temporal changes in “who is there” and “who can do what; when; and how?” we become better equipped to tailor microbiomes for e.g., medical and agricultural purposes.

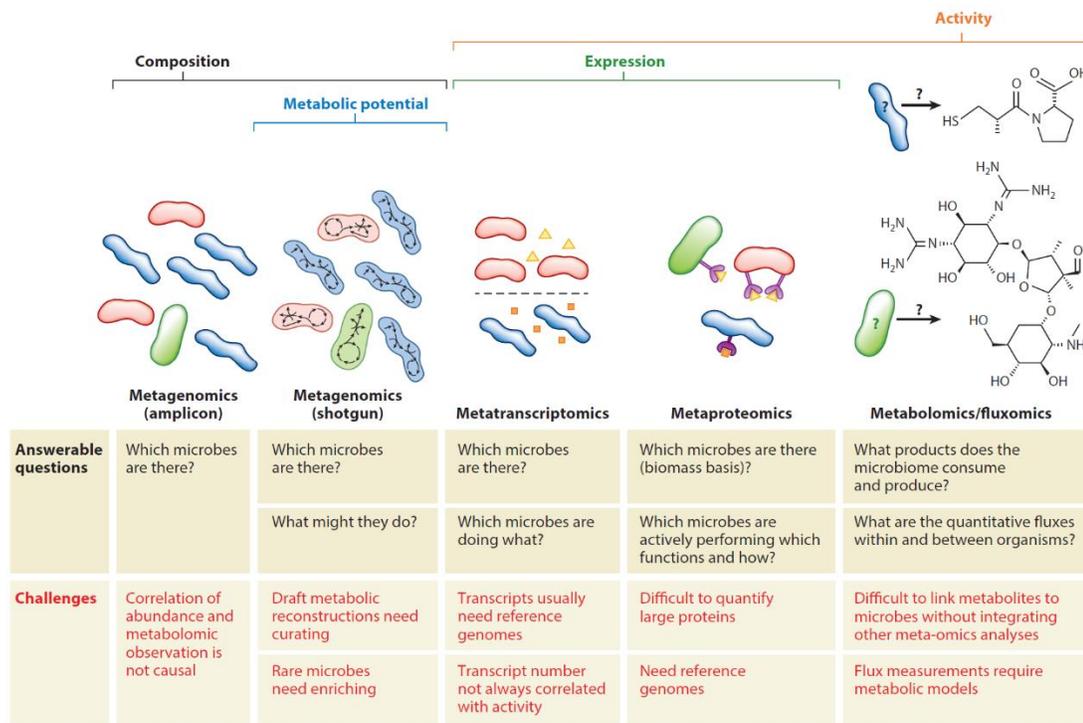
A promising approach is to combine ecological studies, quantitative measurements and computational modeling to map the functional potential of microbiomes with increasing resolution (Robinson et al., 2014).

Here, we describe tools that are currently used to understand microbiomes in diverse habitats and outline how “meta-omics” tools may be used and integrated to characterize microbiome composition and function. We describe representative case studies that showcase the various spatial and temporal scales that influence the composition and collective function of microbiomes, and how interactions between microorganisms can lead to emergent functions that cannot be predicted based on each community member’s behaviors in isolation. We discuss the relative advantages of several microbiome modeling approaches of varying degree of coarse-graining, and highlight recent efforts to integrate multi-omic data and multi-scale considerations into a single model. Throughout, we highlight opportunities to improve our understanding of the causal links between microbiome composition and function and our ability to engineer them for societal benefit.

## **B. Tools for understanding microbiomes**

To harness the properties of microbiomes, we must develop tools that decipher which microbes have the capability and flexibility to perform specific functions, quantify their functional activities across space and time, and decipher interactions between organisms and between organisms and the environment. High-throughput sequencing has significantly enhanced our ability to investigate microbiome composition and functional activities, as today’s next generation and emerging so-called third generation sequencing tools (van Dijk et al., 2018) can rapidly process billions of DNA base pairs (bp) with continuous read lengths

greater than 100 kilo-bp (Amarasinghe et al., 2020) (>2% of the average bacterial genome size). These technologies enable characterization of phylogeny (amplicon sequencing), functional potential (shotgun metagenomics), and gene expression (metatranscriptomics) in thousands of species or synthetic communities simultaneously. Beyond nucleic acid sequencing, metaproteomics and metabolomics can analyze the activities of microbiomes by quantifying the abundance of enzymes that perform key chemical transformations and the metabolites mediating interspecies interactions (**Figure 1**).



**Figure 1.** Different meta-omics tools are suited to answer different questions about microbiome composition and function. Amplicon metagenomics can reveal which organisms are present in a microbiome but not necessarily what each microbe’s role in the community is. Shotgun metagenomics elucidates which microbes are present in the community and what functions they have the capacity to perform. Metatranscriptomics and metaproteomics are necessary to uncover which functions are actually being performed in the community; assigning these transcripts and proteins to the microbes that produced them typically requires high-quality reference genomes or concurrent metagenomics analyses. Metabolomics and fluxomics quantify the chemical composition of the microbiome environment; however, linking metabolites to the microbes that produce or consume them is challenging, even with reference genomes. Linking microbiome composition and function is facilitated by integrating multiple meta-omics techniques, for example, concurrent shotgun metagenomics, metaproteomics, and metabolomics studies to assess which enzymes are producing an observed small molecule of interest and which microbes could produce those enzymes.

Such meta-omics tools can be applied to quantify the spatial distribution of organisms within microbiomes, characterize low-abundance members and assess cellular heterogeneity, identify the organisms that perform key chemical transformations, and elucidate the web of metabolic interactions that ultimately drive microbiome functions.

#### 1. Quantifying microbiome composition and functional potential via metagenomics

Quantitative measurements of microbial abundance are critical for understanding the spatial and temporal behaviors of microbiomes. Microbiome composition is frequently determined via highly conserved marker genes for ribosomal RNA (rRNA), usually the 16S rRNA gene in prokaryotes and the 18S, 28S, and internal transcribed spacer (ITS) regions in eukaryotes (Raja et al., 2017). Amplicon sequencing is particularly useful for community composition characterization in systems contaminated by host DNA or samples with low DNA template concentrations. However, these methods provide a genus-level resolution, and sequence-dependent variations in (nominally) universal primer affinity between clades can bias abundance results, along with extraction efficiencies of DNA and variation in gene copy numbers (Campanaro et al., 2018; Wilson et al., 2019).

In shotgun metagenomics, a library is constructed with all community DNA and the reads can be assembled into genomes of individual species called metagenome-assembled genomes (MAGs). These genome sequences provide simultaneous quantification of the microbiome's functional potential and its phylogenetic composition. Shotgun metagenomics has several advantages including finer phylogenetic resolution than amplicon-based sequencing down to the strain-level (Nayfach et al., 2016; Scholz et al., 2016; Shi et al., 2019; Truong et al., 2017), detection of viral DNA, with the tradeoff that more reads are needed to confidently quantify more genes. To probe rare microbes with potentially unique

functions, DNA extraction methodologies targeted for different species have been leveraged to assemble near-complete genomes of bacteria present below 1% relative abundance within a community by employing differential coverage binning (Albertsen et al., 2013). In addition, the variation in genome copy numbers in different regions of the chromosome have been used to infer the bacterial replication rates in natural environments (Brown et al., 2016), providing key insights into the distribution of metabolic activity states within a community.

Sequencing based methods provide relative abundance or compositional data, which presents challenges for statistical analyses and can lead to spurious correlations. Therefore, methods to quantify absolute abundance are critical to understanding microbiomes including determining correlations between organisms, growth rates, per-cell metabolic activities, or total microbial loads present in a host. Absolute DNA-based quantification of microbiome composition is a major challenge that has been approached with spike-in (Tkacz et al., 2018), quantitative polymerase chain reaction (qPCR) (Lou et al., 2018), flow cytometry (Vandeputte et al., 2017), and total DNA quantification (Contijoch et al., 2019) methods; however all of these have inherent biases and limitations. Recently, amplicon sequencing was coupled to digital PCR in a microfluidic format for absolute DNA-based quantification of microbiome composition with the advantage of concurrent evaluation of the limits of both clade detection and clade quantification (Barlow et al., 2020).

The spatial distribution of clades within a microbiome can be studied by sampling from different locations using sequencing methods. This approach has been used to map the biogeography of the mammalian gut (Sheth et al., 2019), and to demonstrate how anaerobic digestion communities self-assemble into distinct microbiomes when reactors are connected in series (Fontana et al., 2018). While the spatial resolution that can be achieved by sampling

different locations is limited (~20  $\mu\text{m}$ ), micron- and nanometer-level spatial variation can be elucidated using imaging approaches. To uncover how spatial clustering of clades influences community functions and interactions, it is necessary to elucidate both the spatial distribution of functions and the identity of organisms.

The information in MAGs can be used to predict the range of chemical transformations the system is capable of performing by evaluating putative metabolic pathways, as well as which metabolites may be taken up or secreted by certain microbes. For example, lignocellulose-degrading enzymes in the porcupine gut microbiome were identified via shotgun metagenomics and expressed in *E. coli*, leading to the discovery of an active endo-1,4- $\beta$ -xylanase even though the microbe encoding this gene was unknown (Thornbury et al., 2019). Accurate predictions of functional potential heavily rely on reference genomes and metagenomic datasets.

Our ability to collect metagenomic data has outpaced our ability to functionally annotate it for interpretation of biological context (i.e., “who has the capability to do what” in the microbiome). In genetically tractable organisms, a powerful approach to functionally annotate genes involves quantification of strain fitness within pooled genome-wide mutant libraries that can be grown in monoculture or co-culture (Ibberson et al., 2017) across many different environmental conditions (Price et al., 2018). However, since the majority of strains lack genetic tools, bioinformatic approaches including the Integrated Gut Genomes Database and associated IGGsearch tool (Nayfach et al., 2019), and Distilled and Refined Annotation of Metabolism (DRAM) tool (Shaffer et al., 2020) can be used to analyze large sequencing data sets containing potentially thousands of interacting species.

## 2. Metatranscriptomics maps organism identities to functional activities

Metatranscriptomics, in which total community mRNA is extracted, reverse transcribed to cDNA, and sequenced, provides insight into the potential functions performed by organisms within a community as well as which microbes may be performing them. For example, metagenomic and metatranscriptomics were combined to identify sugar fermenting and fatty-acid chain elongating microbes in an anaerobic bioreactor and propose routes of metabolite exchange between the clades (Scarborough et al., 2018). Gomez-Godinez et al. reported by quantifying nitrogen fixation transcripts that *Azospirillum brasilense* was the predominant nitrogen-fixing bacterium in a synthetic consortium of plant growth-promoting bacteria on maize roots (Gómez-Godínez et al., 2019). While metatranscriptomics also generates a large number of unannotated “hypothetical genes,” the ability to identify genes that change across different conditions greatly facilitates the downstream identification of genes involved in microbiome processes including the breakdown of lignocellulose or the biogeochemical cycling of elements. Further, investigating the genome-wide transcriptional activity of organisms within a community may guide the development of hypotheses about mechanisms involved in observed microbiome states (dysbiosis, steady state recovery after perturbation, etc.). For example, a combined metagenomic and metatranscriptomic study of the fecal microbiome of 308 adult men revealed that pathways that were encoded in the genomes of many members of the microbiome were actually transcribed by a small subset of species (Abu-Ali et al., 2018).

Notably, even when reference genomic data is lacking, metatranscriptomics can be used to mine microbiomes for enzymes with desired function. A transcriptomic survey of anaerobic gut fungi harvested from the intestinal tract of herbivores, revealed that these

unusual and understudied eukaryotes produce an unrivalled array of biomass-degrading enzymes (Solomon et al., 2016), marking these fungi as attractive targets for sourcing of valuable enzymes (Seppälä et al., 2017). Similarly, He et al. identified 125,252 putative CAZymes in a sheep gut microbiome, most of which had less than 75% identity to known proteins in the CAZy database or the NCBI database (He et al., 2019), but 19 out of 30 that showed cellulase activity when heterologously expressed.

In microbes, mRNA transcripts represent <10% of total RNA, therefore, rRNA should typically be removed prior to sequencing. Methods for prokaryotic rRNA depletion vary in efficacy based on microbiome type (biofilm vs. planktonic) and composition, and tool development is an active area of research (Petrova et al., 2017). Methods for single-cell prokaryotic rRNA depletion are only beginning to show some success (Fang and Akinci-Tolun, 2016; Wangsanuwat et al., 2020), and will be useful for interrogating the unique activities of low-abundance organisms as well as the cell-to-cell heterogeneity of gene expression within a given species that is not observable with bulk methods. Identifying the optimal spatial locations and timepoints to discover ecological driver organisms or novel biochemical pathways mediating microbiome functions remains unresolved. Further, since transcript number does not always correlate with protein abundance or activity, it is difficult to estimate the relative contribution of different metabolic reactions and pathways to the overall function of the microbiome from metatranscriptomics data alone.

### 3. Quantifying microbiome functional capabilities via metaproteomics

Metaproteomics, which studies all proteins recovered from a microbiome sample, can provide critical information about microbiome functional capabilities. Liquid chromatography can be coupled to mass spectrometry or to tandem mass spectrometry (LC-

MS or LC-MS/MS) for detection of tens of thousands of peptides in one sample (Kleiner, 2019; Mueller and Pan, 2013). In the future, nanopore-based devices hold promise to revolutionize proteomics and biotechnology by enabling amino acid sequencing of intact proteins, enabling structural characterization of larger proteins than LC-MS (Chinappi and Cecconi, 2018). In addition to differential enzyme expression, metaproteomics may be employed to quantify the abundance of individual organisms in a microbiome on a biomass basis (Kleiner et al., 2017), which offers an alternative method for microbiome composition. Indeed, proteinaceous biomass may be a better representation of composition for systems composed of eukaryotic and prokaryotic cells where the size and weight of organisms varies significantly, but perhaps more biased in cases where the intracellular protein content varies widely across species.

Beyond prospecting for genes with predicted functions, metaproteomics can be used to identify post-translational modifications as well as directly quantify the abundance of proteins in a community, which may not necessarily correlate with transcript abundances (Speda et al., 2017). This approach is greatly improved by integration with metagenomic analysis, due to the difficulties of mapping fragmented peptide sequences to genes. In particular, mining microbiomes for membrane proteins such as transporters is critical to elucidating molecular mechanisms involved in interspecies interactions, but remains challenging due to their low abundance compared to soluble proteins and technical/analytical challenges due to their hydrophobicity (Vit and Petrak, 2017).

#### 4. Metabolomics reveals the chemical repertoire of microbiomes

Microbes are exquisite chemists and these chemical mediators produced and utilized by constituent community members are a major driving force of microbiome functions.

Metabolomics can be used to detect small molecule metabolites, including intermediates and end products of cellular metabolism. Integration of metabolomics with metagenomics is particularly useful for formulating hypotheses about the role of measured metabolites in interspecies interactions and microbiome functions, as metabolites typically cannot be assigned to specific organisms. Therefore, it is necessary to integrate other meta-omics tools to determine which metabolic pathways are active in a community to hypothesize how different metabolites may be utilized, released, and exchanged to form an integrated community metabolic network.

Many metabolomic approaches employ gas chromatography (GC) to precede MS analysis instead of LC, offering greater chromatographic separation of metabolites. Nuclear magnetic resonance (NMR) spectroscopy offers an alternative, more quantitative measure of metabolites without the sample preparation and derivatization steps required in MS studies, but typically cannot detect metabolites below micromolar concentration (Emwas, 2015). Untargeted metabolomics seeks to characterize the structures of as many metabolites present in the sample as possible that can be identified. However, it is impossible to characterize all classes of metabolites with a single solvent and column chemistry, and many metabolites in databases remain unannotated (Zamboni et al., 2015). Therefore, strategies must be developed to predict unknown chemical structures and link them to the microbes and biosynthetic pathways. Recently, Henry et al. developed the Pickaxe tool (Jeffryes et al., 2015) for generating novel metabolites and predicting the enzymes and putative pathways based on Enzyme Commission (EC) numbers and the MINE database.

## 5. Investigating metabolic flux in microbial communities

Metabolic flux analysis (MFA) can quantify the distribution (flux) of carbon in cellular metabolism, providing a direct measurement of the activity of metabolic networks. Using this method, cells are typically exposed to  $^{13}\text{C}$ -labeled carbon and the degree of labeling of biomass components like glycogen and proteins is quantified via GC-MS (Long and Antoniewicz, 2019). Based on these data, software, such as METRAN, is used to estimate the flux through each pathway using a organism specific metabolic model. While MFA has been used to study well-characterized and simplified communities (Gebreselassie and Antoniewicz, 2015), the challenge associated with assigning metabolites to microbes in complex communities has stymied the broader application of MFA. However, analysis of isotopic labeled peptides, which can be mapped to individual microbes with reference genomes using metaproteomics may unlock MFA for microbiomes (Ghosh et al., 2014). Nevertheless, this method may perform best on communities composed of microbes with dissimilar metabolisms or that can be spatially separated. To quantify interspecies metabolic interactions and community-wide fluxes, development of “metafluxomic” protocols and software is necessary to translate isotopic labeling and multi-omic data into quantitative descriptions of community metabolic networks (Antoniewicz, 2020).

## 6. Stable isotope probing (SIP) in microbiomes

Stable isotope probing (SIP) has emerged as a promising technique to enrich rare microbes and link microbe identities to functions, and investigate interaction networks within microbiomes. In DNA or RNA SIP, isotopically labelled substrates are differentially taken up and incorporated into nucleic acids by community members. After extraction, nucleic acids are fractionated by density to simultaneously enrich for nucleic acids from rare microbes and

link the affinity for the labeled substrate to microbe identity (Coyotzi et al., 2016; Egert et al., 2018). For example, DNA-SIP was used to curate a complete genome of a *Saccharibacteria* with < 1X coverage in the bulk metagenome and decipher the metabolite exchange networks within the *Saccharibacteria*'s surrounding community (Starr et al., 2018).

A major limitation is that the SIP culturing procedure may not mimic a microbiome's natural microenvironment. While the vast majority of microbes remain uncultivated, meta-omics analyses may elucidate clues for isolating and culturing previously uncharacterized species (Nayfach et al., 2019). Nucleic acid-SIP requires that isotopes be incorporated directly into nucleic acids, and the incubation time with the isotope influences the degree of community labelling. For example, short incubation times exclude isotope uptake by slow-growing microbes, and long incubations may lead to non-specific cross-feeding of isotopes across the community, skewing which clades initially metabolized the substrate (Fortunato and Huber, 2016; Radajewski et al., 2003).

In principle, SIP can also be used to quantify incorporation of labels into proteins (Protein-SIP) (Seifert et al., 2012), metabolites (Metabolome-SIP), and phospholipid-derived fatty acids (PLFA-SIP), though linking these to microbe identity requires excellent reference genomes or concurrent metagenomic and/or metatranscriptomic analyses. When integrated with meta-omics analyses, SIP may link microbe identity to function, and uncover interspecies metabolite exchange mechanisms and community-wide metabolic networks. For example, Protein-SIP coupled to amplicon sequencing and shotgun metagenomics was used to map acetate metabolism to organism identity in anaerobic digester consortia (Mosbæk et al., 2016). In addition, RNA-SIP, metagenomics, and metatranscriptomics have been integrated to characterize the predominant CO<sub>2</sub> fixation pathways and their transcribing

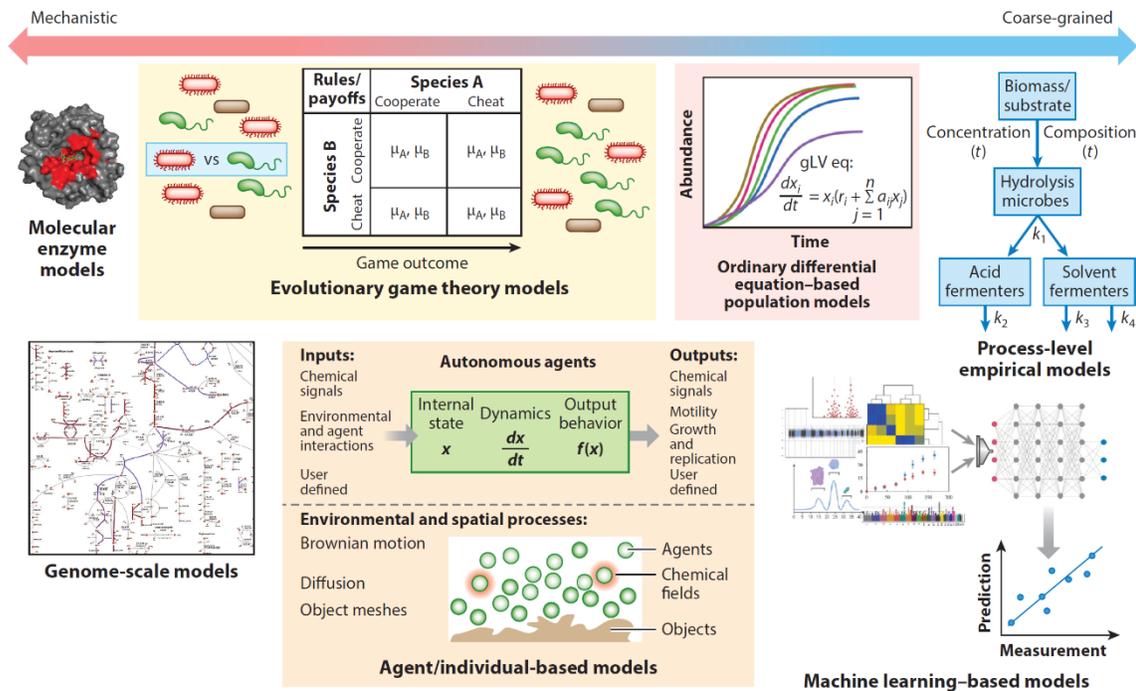
microbes in deep-sea hydrothermal vent microbiomes at a range of temperatures (Fortunato and Huber, 2016). Finally, DNA-SIP has been combined with differential coverage binning to enhance resolution of MAGs with specific activity in anaerobic digesters (Ziels et al., 2018). See (Berry and Loy, 2018) for a review of SIP applied to *in vivo* and *ex vivo* human and animal gut systems.

### C. Computational models predict microbiome dynamics and functions

Mathematical models can be used to simulate microbiome population dynamics and metabolic functions on many time and length scales, based on ecological, thermodynamic, and biochemical principles. These range from data-driven differential equation-based models of community composition and interactions to mechanistic genome-scale models of metabolic flux and interspecies metabolite exchange (**Figure 2**). Below, we discuss relative advantages and limitations of several modeling approaches, and how they may be used to enable microbiome engineering.

#### 1. ODE and EGT models of microbiome population dynamics and interactions

Ordinary differential equation (ODE) models, such as the generalized Lotka-Volterra (gLV) model (Venturelli et al., 2018), have been used to model microbiome population dynamics using time-series data of absolute organism abundance for parameter estimation and experimental validation (Kumar et al., 2019; Song et al., 2014). In the gLV model, the temporal changes in abundance of each species is a function of its growth rate, intra-species and inter-species interactions (Cao et al., 2019). The gLVs models can be used to analyze dynamic behaviors including the response to perturbations such as dilution rate, response to antibiotics and temperature fluctuations (Dam et al., 2016; Stein et al., 2013). In addition, the inferred parameters of the gLV model can be visualized as an interaction network to examine



**Figure 2.** Microbiomes can be modeled on many scales, and the choice of modeling technique depends on the question at hand. At the most mechanistic level, molecular simulations may be used to model the thermodynamics and kinetics of individual enzymes identified through metaproteomics; however, these are not scalable to encompass the entire microbiome. GEMs enable prediction of the metabolic fluxes and end-product profiles within a microbiome and can offer mechanistic insight into metabolomic observations given high-quality genomic reconstructions and sufficient experimental model validation. Evolutionary game theory models and differential equation-based models are particularly useful when microbiome population dynamics are of the greatest interest, because detailed metabolic reconstructions are not needed for each organism to be modeled. AbMs offer flexibility in that the user may define which inputs and outputs to include in the model, and are often the technique of choice when integrating both metabolic and physical interactions between microbes. Data-driven models, including emerging machine learning-based models, offer empirical predictions of microbiome behaviors under specified conditions given appropriate training data. Although less mechanistic than GEMs, machine learning-based models are a pragmatic approach to synthesizing large amounts of different data types into interpretable conclusions, for example, rate constant estimations for process-level models of microbiome function. The structure of the molecular enzyme model is from PDB ID 4QLK. The structure of the AbM is reproduced with permission from Reference 162. The microbes in the evolutionary game theory models panel and the entire machine learning-based models panel were adapted from images created with BioRender.com. Abbreviations: AbM, agent-based model; EGT, evolutionary game theory; GEM, genome-scale model; ODE, ordinary differential equation; PDB, Protein Data Bank.

the distribution of negative and positive interactions and identify ecological driver species (Venturelli et al., 2018). The gLV model could be used to design community cultivation strategies to achieve desired community compositions and stability properties.

Evolutionary game theory (EGT) can also be used to model microbiome population dynamics, as described in these reviews (Frey, 2010; Pusa et al., 2019; Song et al., 2014).

The fitness parameters that dictate the outcome of metabolic “games” in microbiomes can be difficult to estimate, as they are influenced by nonlinear environmental and intracellular conditions. For this purpose, EGT can be integrated with genome-scale models (GEMs, described below), enabling prediction of interspecies interactions and stable steady state fluxes and species abundances. The system states at which each microbe locally maximizes its own growth, but not necessarily the global maximum community growth, can be identified as Nash equilibria and evolutionarily steady solutions (a subset of Nash equilibria) (Nowak, 2006), or asymptotically stable solutions to dynamic replicator equations, all of which pose candidates for stable coexistence states that the microbiome could exhibit. Evolutionary stability is a key factor to consider when designing bottom-up communities or altering the composition of a native system. While gLV and EGT models capture context-dependent pairwise interactions, these models fail to capture higher-order or metabolite-based interactions in the community. Public goods games (PGGs) are, in principle, extendable to multispecies interactions and integrable with other modeling approaches like genome-scale models; some associated challenges are reviewed in (Pusa et al., 2019).

## 2. Predicting microbiome fluxes and interactions with genome-scale models (GEMs) and machine learning

Mechanistic genome-scale models (GEMs) offer prediction of microbial behaviors in untested conditions and serve as useful platforms for synthesizing multi-omic data into one comprehensible and interactive format. GEMs mathematically represent an organism’s metabolic pathways (with gene-protein-reaction associations (GPR) in metadata) as a stoichiometric matrix of reactions and metabolites (Gu et al., 2019; Thiele and Palsson, 2010). Through flux balance analysis (FBA) (Orth et al., 2010) and related techniques

including flux variability analysis (FVA) (Gottstein et al., 2016; Gudmundsson and Thiele, 2010), GEMs can be used to assess the effects of media and substrate changes and genetic edits on fluxes of target compounds; this approach has been used extensively to guide metabolic engineering (Hartmann et al., 2017; Kim et al., 2015; Zhang and Hua, 2016). As metabolic engineers look to co-cultures for specialty products (Sgobba and Wendisch, 2020; Wang et al.) and systems biology is applied to medicine (Thiele et al., 2020; Zhang and Hua, 2016), genome-scale modeling of microbiomes is becoming increasingly useful.

### 3. Automated genome-scale reconstructions: scaffolds for GEMs

Advances in meta-omics tools described previously have enabled semi-automated construction of metabolic networks for hundreds of species that encompass a microbiome (tools reviewed in (Mendoza et al., 2019)); however, manual curation is needed to accurately recapitulate metabolism in silico (Reimers et al., 2017). For example, Magnúsdóttir et al. published the AGORA (Assembly of Gut Organisms through Reconstruction and Analysis) database, complete with 773 genome-scale reconstructions of human gut microbes, and simulated pairwise microbial interactions when fed with different diets (Magnúsdóttir et al., 2017). Follow-ups to this report highlight the distinction between a reconstruction and a context-specific predictive genome-scale model, and the importance of identifying and applying appropriate constraints before attempting to simulate metabolism in silico (Babaei et al., 2018; Magnúsdóttir et al., 2018).

Databases exist for published genome-scale reconstructions such as BiGG (Norsigian et al., 2020), Model SEED (Seaver et al., 2020). Significant progress has been made in cataloguing reconstructions for the human gut microbiome and human body at large (Magnúsdóttir et al., 2017; Noronha et al., 2019; Robinson et al., 2020). These

reconstructions serve as templates for context-specific GEMs, such as a known viral infection of a macrophage cell (Aller et al., 2018) or dysbiosis of the gut (Kumar et al., 2018). In an important step towards universalizing GEMs over different sequence annotation styles, programming languages, and operating systems, Lieven et al. recently published MEMOTE, which scores GEMs for completeness and feasibility (Lieven et al., 2020). However, owing to the highly variable and environment-dependent nature of enzyme kinetics and transcriptional regulation, the individual constraints that make models predictive cannot yet be generally catalogued this way.

#### 4. Constraining and optimizing community GEMs

In curating a GEM from a GENRE, integration of multi-omic data and experimental metabolite concentrations (Kuang et al., 2020) facilitates filling in pathway gaps, constraining the solution space (Tian and Reed, 2018), and validating and improving simulations of microbiome function. To this end, Pandey et al. recently published REMI (Relative Expression and Metabolomics Integrations) for integrating transcriptomic, thermodynamic, and metabolomic data from differential expression analyses into GEMs (Pandey et al., 2019), which Hadadi et al. used to simulate the transition of *Pseudomonas veronii* from exponential to stationary phase, as well as from culture in liquid media to soil, demonstrating advancement in our ability to model microbial adaptation to environmental perturbations (Hadadi et al., 2020).

In addition to flux constraints, choosing and defining the objective function to be optimized is particularly challenging for community models (Gottstein et al., 2016). In single-organism FBA, the objective function is usually to maximize flux through a biomass-forming “reaction” based on the organism’s macromolecular composition (i.e., to grow as

fast as possible). For multispecies FBA systems, many optimization strategies exist. In so-called “supra-organism” approaches, the metabolic pathways in all organisms are combined into one stoichiometric matrix, and organisms are partitioned into separate “compartments” that exchange metabolites (Song et al., 2014; Stolyar et al., 2007). Optimization of the weighted community growth rate may require individual species to grow at sub-optimal growth rates, which is not always an accurate assumption (especially in competitively interacting systems). Bi-level optimization has been implemented to maximize individual and community growth rates together in algorithms such as OptCom (Zomorodi and Maranas, 2012) and CASINO (Shoaie et al., 2015).

To model microbiome fluxes at steady state, Community FBA (cFBA) (Khandelwal et al., 2013) and SteadyCom (Chan et al., 2017), impose a fixed community growth rate that is adopted by all organisms, however this is not always achievable in practice. Alternatively, EGT can be integrated with FBA to predict evolutionarily stable interactions and steady state flux distributions (Zomorodi and Segrè, 2017). In other bi-level optimization strategies, a microbe may be predicted to produce a metabolite that benefits the community but does not necessarily maximize its own growth rate. To avoid imposing this “forced altruism,” Cai et al. developed NECom, which predicts steady state community fluxes and pairwise interactions by identifying Nash Equilibria and removes any influence from the community optimization problem on a microbe’s incentive to secrete a metabolite in community GEMs (Cai et al., 2019).

While highly useful for predicting and describing microbial fluxes and interactions, GEMs require significant time and resources to construct and curate predictive models from automated reconstructions. Since a microbiome’s composition must be fully defined to

accurately apply FBA, GEMs are currently limited to bottom-up microbiomes with a few representative species, which often lack long-term stability. Further, experimental measurements of community growth rates and fluxes are critical to GEM validation improvement but remain a challenge to obtain for large (>3 member) communities.

#### 5. Machine learning can identify complex mappings between microbiome inputs and outputs

Machine learning can be employed to predict and link microbiome composition and functions using multi-omic data, as reviewed in (Namkung, 2020; Zhou and Gallins, 2019). Machine learning can create nonlinear mappings between microbiome inputs (cultivation conditions, species pairings, etc.) and outputs (metabolite concentrations, gene expression profiles, relative organism abundances, etc.) that may enable design of novel microbial consortia with target functions without the need to rigorously characterize each organism in isolation. A major limitation is that these models may not provide insight into the biological mechanisms that generate the observed microbiome states. To address this challenge, machine learning has been integrated with GEMs to extract information from multi-omic experiments and metabolic simulations that are relevant to the engineering objective at hand (Zampieri et al., 2019).

#### 6. Dynamic models of microbiome flux

Dynamic computational models are particularly useful when the time-varying microbiome fluxes or 3D structure are of interest (Øyås and Stelling, 2018). Dynamic FBA (dFBA) has been used extensively to model transient compositions and flux distributions in small (2-3 member) microbial communities, as reviewed in (Gottstein et al., 2016). In these cases, estimation of metabolite uptake and secretion kinetics is particularly important for

accurately simulating growth and fluxes. Characterization of transporter membrane protein specificity and influx/efflux kinetics promises to significantly improve both dynamic and steady state FBA simulations (Nilsson et al., 2017; Zeng and Yang, 2020), but remains a major challenge (Boyarskiy and Tullman-Ercek, 2015).

Spatial heterogeneity can be incorporated into community GEMs (Harcombe et al., 2014; Lillington et al., 2020; Øyås and Stelling, 2018), typically by using FBA to find each species' growth rate at each time step in numerical solutions of reaction-diffusion partial differential equations (Chen et al., 2016). Biofilms, in particular, are increasingly being modeled with spatiotemporal GEMs (Harcombe et al., 2014; Henson and Phalak, 2017; Patel et al., 2019; Phalak et al., 2016). To account for their moving boundary conditions brought on by film growth/expansion, it is useful to simulate them as collections of individual microbes in so-called agent-based models (AbMs).

#### 7. Agent-based models (AbMs) simulate predefined physical and metabolic interactions in microbiomes

In AbMs, microbes are treated as individuals with specified traits, rather than concentration state variables as in other methods. AbMs can capture gene regulation (Gorochowski et al., 2012), metabolic and mechanical interactions between microbes in a community, and are well-suited to model biofilm formation, deformation, and disruption (Jayathilake et al., 2017). Metabolism can be coarse-grained to allow only for reactions involving exchange with the environment; or AbMs can integrate genome-scale metabolism to compute fluxes with FBA. The latter approach is taken in BacArena, an R package that was demonstrated to model spatiotemporal metabolic interactions among seven human gut microbes (Bauer et al., 2017).

Although they are computationally demanding, AbMs are versatile models that suggest priority experiments for answering specific questions of microbiome function. For example, van Hoek et al. simulated metabolism in the human large intestine with a coupled dFBA-mass transport model of individuals from a supra-organism “metabacterium” (Hoek and Merks, 2017). Although they did not include experimental validation, their model generated hypotheses for the effects of diarrhea (a macro-scale system state) on the micro-scale spatial organization and relative abundance of microbes with various flux profiles, which can be explored experimentally. Similarly, Doloman et al. used a multispecies AbM (without genome-scale metabolism) to predict cultivation conditions that maximize methane productivity in spatially-structured anaerobic sludge granules (Doloman et al., 2017).

**D. Deploying lignocellulose-valorizing microbial consortia with anaerobic fungi requires detailed understanding of AGF physiology and metabolism**

Reprinted and adapted with permission from (Wilken et al., 2021).

Microbiome engineering should consider the complex interplay and feedback loops between the environment and the resident microbiome, which together drive community dynamics and multifunctional properties. Advancements in our understanding of microbiomes gleaned from ecology and systems biology must be leveraged to develop new strategies to program prescribed functions using tools from synthetic biology. Top-down design of microbiomes presents challenges in predictably modifying composition and function, whereas the engineering of stable and highly functional synthetic communities from the bottom-up also remains difficult. Development of microbiome engineering that can achieve high precision, robust and predictable outcomes holds promise for diverse applications in biotechnology, medicine, agriculture and the environment. A comprehensive

understanding of microbiomes necessitates an integration of multi-omic data as well as quantification of the spatial and ecological factors that contribute to community functions.

### 1. The anaerobic fungal genome scale model: iNlan20

Microbial communities including AGF have significant potential to valorize lignocellulose through biomass degradation and production of valuable enzymes and bioproducts. However, before AGF can be deployed in communities, AGF physiology and metabolism must be understood in detail. Toward this goal, we developed the first experimentally validated AGF GEM. The 3-compartment (extracellular, cytosolic, and hydrogenosomal compartments) model, named iNlan20, is composed of 1,018 genes, 1,023 reactions, and 816 metabolites, and captures the primary metabolism of *N. lanati*. The model is stoichiometrically consistent as well as mass and charge balanced.

Experimental, genomic, transcriptomic, and  $^{13}\text{C}$  metabolic flux analysis data were used to build and validate the model, which recapitulates extracellular metabolite production rates and accurately models the observed fungal growth rate. The model suggests possible pathways within the AGF hydrogenosome (an anaerobic mitochondrion-like organelle capable of energy production) and can be used as a platform upon which to base future experiments to understand AGF metabolism comprehensively.

### 2. iNlan 20 captures carbon, amino acid, vitamin, fatty acid, nucleotide, and lipid metabolism

The model focuses on the primary metabolism but includes CAZymes as generalized cellulase and hemicellulase reactions. Of the 791 metabolic genes included in the model, 216 do not have gene assignments, reflecting how superficial our understanding of AGF physiology is and motivating improvement of the existing AGF genomes. Energy-generating

pathways in AGF include the Embden-Meyerhof-Parnas variant of glycolysis, mixed-acid fermentation, and the hydrogenosome. It remains unclear which hydrogenosomal pathways carry significant flux *in vivo*, however the model described herein offers informed hypotheses based on both FBA simulations and experimental results and offers a starting point for further characterization of AGF metabolism in the hydrogenosome and at large.

The major components of anabolic metabolism (amino acids, nucleotides, vitamins, fatty acids, and lipids) are present in the *N. lanati* genome, which agrees with its ability to grow in defined M2 medium without addition of any of these components. Specifically, the genome contains complete biosynthesis pathways for all the proteogenic amino acids and the modeled fatty acids. Most of the canonical vitamin and cofactor (vitamin B5, vitamin B6, riboflavin, and thiamine) biosynthesis pathways are complete, except for folate, where no synthesis mechanism of 4-aminobenzoate was found. The heme and biotin biosynthesis pathways are incomplete in the genome; however, the latter vitamin is not required for growth in defined medium, suggesting that the pathway is either poorly annotated or that the fungus does not require it for growth. The model accurately predicts the experimentally observed growth rate (predicted:  $\mu = 0.044 \text{ h}^{-1}$ ; observed:  $\mu = 0.045 \pm 0.003 \text{ h}^{-1}$  using only the measured flux of glucose ( $1.5 \text{ mmol/gDW h}$ ) as an input constraint. However, it should be noted that, based on comparisons between stirred and unstirred AGF cultures (Leggieri et al., 2022), this observed *N. lanati* growth rate is likely based on substrate diffusion rate, rather than intrinsic growth rate, motivating extension of the model to well-mixed AGF systems.



metabolic flux measurements based on a statistically significant fit between measured and simulated proteinogenic amino acid labeling patterns (**Figure 3**).

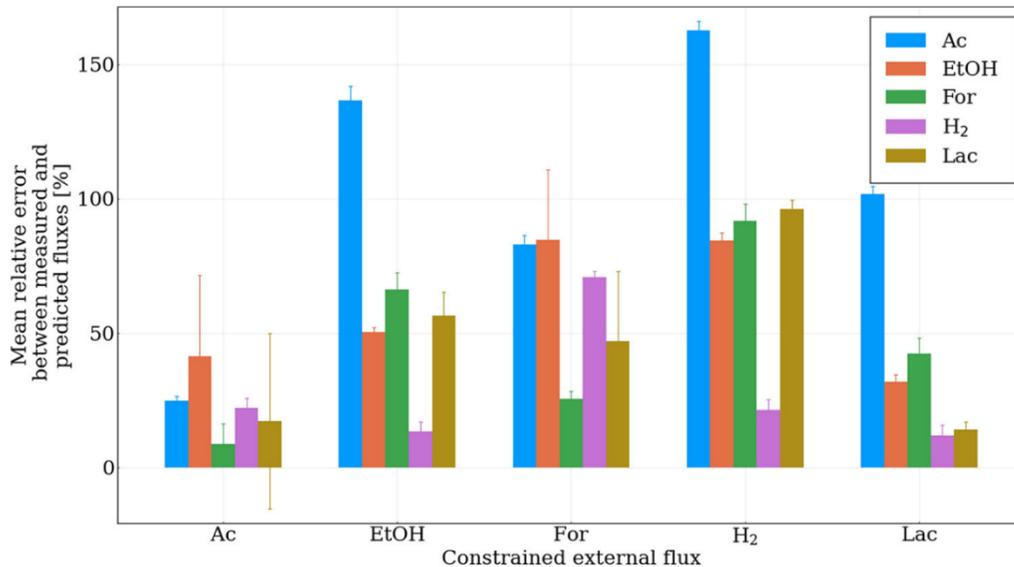
These  $^{13}\text{C}$  measured fluxes were then compared to the fluxes predicted using the GEM with independently measured metabolite flux constraints (**Table 1**). We then used parsimonious flux-based analysis (pFBA) to find unique flux predictions. Using these constraints, the coefficient of determination between the pFBA and MFA simulation was found to be 0.98 (linear regression fit  $p < 0.01$ ), suggesting that the constrained metabolic model accurately predicts the (assumed) steady state measured intracellular fluxes of *N. lanati*. However, this steady-state assumption is limited to a few hours of growth in this unmixed batch system, motivating method development to culture AGF continuously at true metabolic steady state.

**Table 1.** Experimentally measured external fluxes of metabolites produced by *N. lanati* grown on cellobiose in M2 medium

Metabolite	Flux (mmol/g [dry weight]/h)			
	Mean	SD	Lower bound	Upper bound
Succinate	0.03	0.01	0.02	0.05
Lactate	0.87	0.14	0.72	1.09
Ethanol	0.66	0.20	0.47	1.01
Formate	1.40	0.30	1.09	1.79
Acetate	0.56	0.12	0.42	0.71
H <sub>2</sub>	0.10	0.06	0.05	0.19

The aforementioned external flux constraints used in the MFA model are required for accurate FBA predictions, which represents a significant limitation of the model. As is typical of unconstrained GEMs, the modeled gut fungal metabolism displays significant degeneracy. This degeneracy is primarily due to the ability of *N. lanati* to regenerate NAD<sup>+</sup> in several different ways using its mixed acid fermentation pathways, i.e., through a combination of lactate dehydrogenase, acetaldehyde dehydrogenase, and alcohol

dehydrogenase. As shown in **Figure 4**, experimental fluxes of fermentation products can only be predicted with reasonable accuracy when the experimental acetate flux is used to constrain the model.

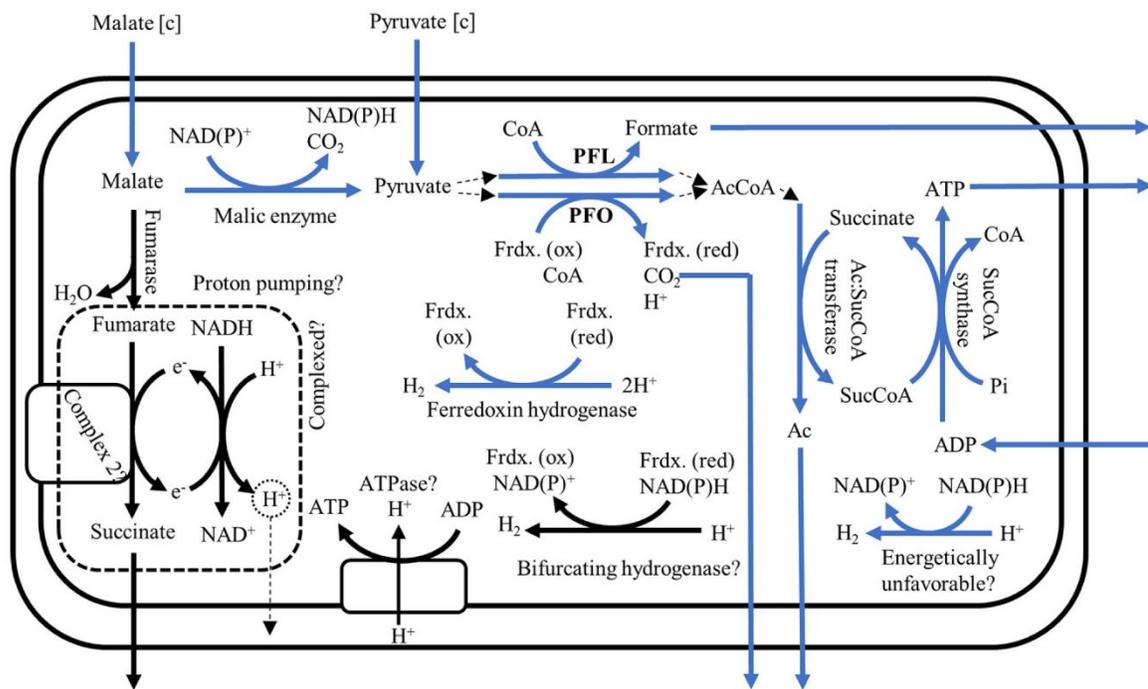


**Figure 4.** The absolute relative error between the model predictions and the experimentally measured values suggest that constraining the flux of acetate production has the biggest impact on the model’s accuracy. The flux of acetate (Ac), ethanol (EtOH), formate (For), H<sub>2</sub>, and lactate (Lac) was constrained, individually, to their observed ranges (variables on the x axis). The resultant predicted fluxes of these metabolites (generated by sampling 2,000 possible solutions where the biomass objective function was within 90% of its optimal value and subject to the respective additional constraints as shown in the figure) were then compared to the experimental observations as shown in the legend.

We must provide the model with a piece of the solution, and it fills in the blanks. This is useful in a descriptive sense, but not a predictive one. Acetate is only produced in the hydrogenosome, therefore **Figure 4** implies that the model cannot predict hydrogenosome flux without constraints. This is logical, as myriad pathways are transcribed in the hydrogenosome for both ATP production and NAD<sup>+</sup> consumption and generation. Stoichiometrically, we cannot predict which pathways the fungus would prefer to use, and biological regulation likely plays the deciding role. *iNlan20* includes all of these pathways; however, as described below, it is unclear which pathways carry significant flux (Boxma et al., 2004; Marvin-Sikkema et al., 1994).

#### 4. iNlan20 includes an expanded model of hydrogenosomal metabolism

Anaerobic gut fungi possess a variant of the hydrogenosome, with the core set of enzymes that catalyze the conversion of malate and pyruvate to acetate, hydrogen, and formate already identified, as shown in **Figure 5** (Boxma et al., 2004; Hackstein et al., 2019; Marvin-Sikkema et al., 1994; Yarlett et al., 1986). However, the metabolic pathways leading to hydrogen production are not resolved, with literature suggesting either pyruvate ferredoxin oxidoreductase (PFO) or pyruvate formate lyase (PFL) as possible routes (**Figure 5**). Both enzymes are present in the *N. lanati* genome and transcriptome and are thus included in the model of the hydrogenosome.



**Figure 5.** An expanded model of the hydrogenosome is included in the model based on genomic annotation, literature, and predicted localization data (14–16). Core hydrogenosome enzymes are colored in blue, while speculative enzymes are shown in black. PFL, pyruvate formate lyase; PFO, pyruvate ferredoxin oxidoreductase; Ac, acetate; SucCoA, succinyl coenzyme A; CoA, coenzyme A; AcCoA, acetyl coenzyme A; Frdx, ferredoxin.

Earlier enzymatic characterization of hydrogenosomal proteins in

Neocallimastigomycota suggested that PFO is the primary route for hydrogen production

through an associated ferredoxin hydrogenase, as found in the hydrogenosomes of other organisms (Marvin-Sikkema et al., 1994, 1993; Müller et al., 2012; Yarlett et al., 1986). However, more recent studies suggest that PFO is either absent or of only marginal importance in the gut fungal hydrogenosomal metabolism (Akhmanova et al., 1999; Boxma et al., 2004). These later studies suggest that PFL is significantly more active than PFO. It has been suggested that hydrogen evolution occurs through a hydrogen dehydrogenase working in an energetically infeasible reverse direction (Boxma et al., 2004; Youssef et al., 2013). PFL and proteins with some similarity to PFO were identified in all published gut fungal genomes as well as in *N. lanati*. The model was used to assess the role and relative importance of these two enzymes to hydrogenosome function in *N. lanati*.

Since there is no energetic cost associated with using PFO versus PFL (both produce one ATP molecule per pyruvate) (**Figure 5**), the model predicts that both could be used to maximize ATP production in the hydrogenosome. However, external metabolite flux measurements show only modest hydrogen production (**Table 1**), suggesting that cellular regulation may play a role in diverting flux to PFL instead of PFO. Since PFO is the only (known) energetically feasible way to produce hydrogen, the model's best guess is that PFL carries the most flux in the hydrogenosome, but PFO is used to produce hydrogen.

By combining literature sources, gene annotation, transcriptomic expression, and subcellular localization data, we have included additional pathways in the model of the hydrogenosome for *N. lanati* (**Figure 5**). We have no definitive experimental evidence to confirm the activity of the ATP synthase or the proton pumping mechanism, and FBA simulations are most accurate when their fluxes are constrained to zero. Therefore, it is not likely that they carry significant flux, at least during the brief window of “steady state” *N.*

*lanati* growth that simulations were based on. However, they are included in the model for future iterations of AGF GEMs. As for the bifurcating hydrogenase, we find high-homology sequences in the *N. lanati* genome to all three of the bifurcating hydrogenase subunits that have been enzymatically characterized in *Thermotoga maritima* (Schut and Adams, 2009), however FBA also suggests negligible activity of such an enzyme.

## 5. Conclusion and motivation of further work on AGF metabolism

iNlan20 represented the state of the art of AGF metabolism as of 2021, however it is significantly limited in its predictive capacity. We built the model through integration of multi-omic, wet lab, and literature data, and found reasonable FBA solutions for most of our experimental growth and flux data after we supplied the model with part of the solutions. It is likely that flux in the AGF hydrogenosome and AGF metabolism at large are governed by highly complicated biological regulation, in part due to the unorthodox AGF lifecycle. Such systems are difficult to capture in genome-scale models, as their accuracy depends on user-supplied (often empirical) constraints, as evidenced in **Figure 4**. The model deserves an update when it can accurately predict growth and flux in a well-mixed AGF system at true metabolic steady state without relying on user-supplied external flux constraints, because those fluxes are what we want the model to predict. This may require comprehensive understanding of the hydrogenosome, however if regulation is in fact as critical to the hydrogenosome as we expect, the system may simply be ill-suited for genome-scale modeling, at least until genetic tools become readily available for AGF.

iNlan20, despite not being predictive, is highly useful as a descriptive model. It is a self-contained, interactive format to explore and understand all of our collective knowledge of AGF metabolism as of 2021. We based the model on *N. lanati* because, at the time, it was

our only strain that grew in defined media. With simple modifications to our media protocol (adding hemin after autoclaving, rather than before), I uncovered that all of our strains grow in defined media, enabling comparison and extrapolation of the results from iNlan20 to metabolism in *Caecomyces churrovis*, *Anaeoromyces robustus*, and AGF at large. The model showed us the wealth of CAZymes encoded in AGF genomes and the different fermentation products AGF could produce. Based on this information, I designed studies to assess how we could **tune** AGF growth, CAZyme production, and flux without genetic tools. The model showed us the theoretical design space for AGF phenotypes, and I set out to uncover which phenotypes would be advantageous for eventual industrial deployment, and which of these phenotypes are attainable.

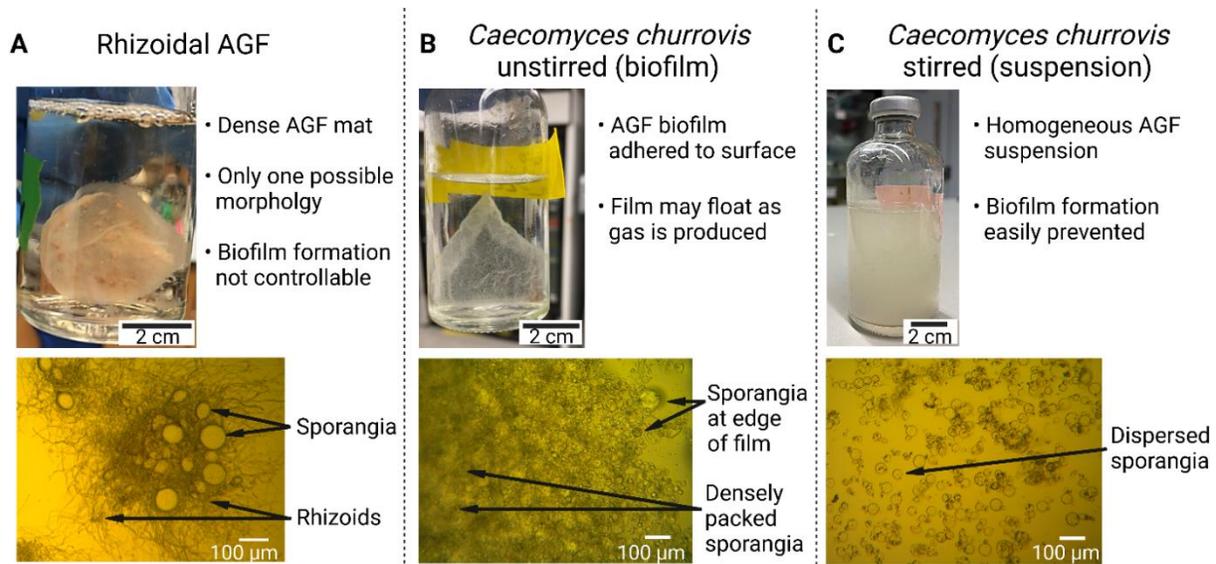
## **II. Biofilm disruption enhances growth rate and carbohydrate-active enzyme production in anaerobic fungi**

Reprinted with permission from (Leggieri et al., 2022).

Anaerobic gut fungi (AGF) are lignocellulose degraders that naturally form biofilms in the rumen of large herbivores and in standard culture techniques. While biofilm formation enhances biomass degradation and carbohydrate active enzyme (CAZyme) production in some bacteria and aerobic fungi, gene expression and metabolism in AGF biofilms have not been compared to non-biofilm cultures. Here, using the tunable morphology of the non-rhizoidal AGF, *Caecomyces churrovis*, the impacts of biofilm formation on AGF gene expression, metabolic flux, growth rate, and xylan degradation rate are quantified to inform future industrial scale-up efforts. Contrary to previous findings, *C. churrovis* upregulated catabolic CAZymes in stirred culture relative to biofilm culture. Using a *de novo* transcriptome, 197 new transcripts with predicted CAZyme function were identified. Stirred cultures grew and degraded xylan significantly faster than biofilm-forming cultures with negligible differences in primary metabolic flux, offering a way to accelerate AGF biomass valorization without altering the fermentation product profile. The rhizoidal AGF, *Neocallimastix lanati*, also grew faster with stirring on a solid plant substrate, suggesting that the advantages of stirred *C. churrovis* cultures may apply broadly to other AGF.

## A. Introduction

Anaerobic gut fungi (AGF) native to the rumen of large herbivores are potent lignocellulose degraders with potential for industrial valorization of crude biomass (Hooker et al., 2019). Their diverse repertoire of biomass-degrading carbohydrate-active enzymes (CAZymes) enables physical association with plant substrates to facilitate mass transfer of hydrolytic enzymes and metabolites (Lillington et al., 2020). Most characterized AGF are rhizoidal (Hanafy et al., 2020), such as *Neocallimastix lanati* pictured in **Figure 6A**, and form dense biofilm-like mats as their rhizoids entangle with solid substrates and other microbial biomass.



**Figure 6.** A) Rhizoidal AGF, which comprise the majority of AGF sequenced to date, only grow in mats as the rhizoids of multiple sporangia entangle with each other. This cannot be avoided, therefore biofilm formation in rhizoidal AGF cannot be controlled. Pictured: *N. lanati*. B) The non-rhizoidal morphology of *C. churrovis* enables control of biofilm formation. *C. churrovis* forms biofilms adhered to the culture surface when cultured without stirring. When mounted on a slide, films appear thick and dense at the interior, and with less tightly packed sporangia at the edge of the film. C) When continuously stirred, *C. churrovis* forms more homogeneous, nearly monodisperse cell suspensions, not biofilms. Created with BioRender.com.

Biofilm formation enhances the rate and extent of biomass degradation in many cellulolytic microbes (Brethauer et al., 2020) and has been linked to increased cellulase activity (Chahal, 1985; Gamarra et al., 2010; Xiros and Studer, 2017) and upregulation of

CAZymes in anaerobic bacteria (Dumitrache et al., 2017) and aerobic fungi (Gamarra et al., 2010; Zhao et al., 2019) relative to planktonic cultures. By contrast, the formation of biofilms by AGF cultures is not well described in the literature, and the biological advantage and function of biofilm formation in AGF remain unknown.

In particular, the effects of biofilm formation on AGF growth rates, substrate degradation rates, metabolic fluxes, and gene expression have not been studied. Gene expression and metabolism may vary substantially between biofilm and planktonic cultures of the same organism, therefore detailed understanding of the physiological effects of biofilm formation is required to design microbial systems for robust biomass conversion. This is especially important for industrial scale-up, as large-scale anaerobic cultures may benefit from agitation for accelerated mass transfer, which could disrupt naturally forming biofilm structures in both rhizoidal and non-rhizoidal AGF.

Typically, rhizoidal AGF grow exclusively in “mat” morphologies. The lack of well-mixed, suspension-based cultures precludes the use of non-biofilm control cultures and limits the ability to isolate the effects of biofilm formation using most standard AGF culture techniques. It is noted that rhizoidal AGF mats may not fit the technical definition of a bacterial-derived biofilm, as they are composed primarily of entangled extracellular rhizoids enriched in polysaccharides like chitin, rather than solely secreted extracellular polymeric substances (EPS) (Brethauer et al., 2020; Chandrasekar and Manavathu, 2008).

To allow for direct comparisons between biofilm-forming and suspended cultures, the focus here is on a non-rhizoidal AGF, *Caecomyces churrovis*. *C. churrovis* presents similar metabolism (97% of EC numbers shared with at least one other AGF) and CAZyme profiles to rhizoidal AGF (Henske et al., 2017; Leggieri et al., 2021a). However, *C. churrovis* forms

biofilms when cultured without stirring (**Figure 6B**), and forms nearly monodisperse cell suspensions in stirred culture (**Figure 6C**). Thus, *C. churrovis* allows for quantitative analysis of the direct effects of biofilm formation on AGF growth, gene expression, and metabolic flux. It is argued herein that understanding the effects of biofilm disruption and shear on gene expression and flux in *C. churrovis* can inform the understanding of the mechanisms and consequences of biofilm formation, as well as industrial scale-up efforts for all AGF (Gruninger et al., 2014; Vinzelj et al., 2020), even though the phenotypic response to shear may differ in rhizoidal strains.

Here, a simple pipeline to isolate the impact of biofilm formation on AGF growth and functional properties using *C. churrovis* as a model system is presented. Via differential gene expression analysis in unstirred biofilm-forming cultures vs. stirred cultures of suspended planktonic cells, the effects of biofilm formation on the expression of biomass-degrading CAZymes, primary metabolic enzymes, substrate binding proteins, and other gene types are quantified. Obtained RNA-Seq transcriptional data are aligned to the existing *C. churrovis* genome, and a *de novo* transcriptome of the unmapped transcripts is assembled to uncover previously unannotated CAZymes and search for proteins potentially involved in biofilm formation. Growth rates, xylan degradation rates, and metabolic fluxes are measured in the stirred and biofilm cultures. Interestingly, while growth and xylan degradation were significantly faster in stirred cultures, few primary metabolic enzymes were differentially expressed, and accordingly, primary metabolic fluxes were not significantly altered. A significant number of differentially expressed genes and *de novo* transcripts could not be annotated, although some showed homology to proteins in other biofilm-forming organisms. Several of the unannotatable genes and *de novo* transcripts clustered by sequence similarity,

potentially laying a foundation for the discovery of the genes responsible for AGF biofilm formation. Further, the rhizoidal AGF, *Neocallimastix lanati*, showed significantly faster growth and biogas production on a solid plant biomass substrate in stirred culture than in unstirred culture, suggesting that the results observed in *C. churrovis* may potentially apply to other AGF and motivating the design of agitated/continuous AGF culture systems for CAZyme production and biomass conversion.

## **B. Materials and Methods**

### **1. Culture of anaerobic gut fungi**

A fully defined, modified version of anaerobic Medium B (MB) (Leggieri et al., 2021a; Theodorou et al., 1995) was used for both routine culture and growth experiments. The headspace of AGF cultures was 5% H<sub>2</sub>, 35% CO<sub>2</sub>, balance N<sub>2</sub>. All cultures were grown at 39 °C in 75 mL anaerobic serum bottles with 45 mL of liquid medium. AGF were grown on xylan (from corn core, TCI America, Portland, OR) or dried, milled reed canary grass (milled with a 4 mm screen size, provided by the US Department of Agriculture, Agricultural Research Service, US Dairy Forage Research Center) at final concentrations of 5 g/L and 10 g/L, respectively.

*Caecomyces churrovis* was previously isolated from the feces of a large herbivore (Henske et al., 2017), and *Neocallimastix lanati* was isolated from the feces of a sheep (Wilken et al., 2021). For routine culture, AGF were transferred to new media every 2 days. In all growth experiments, starter cultures of AGF were grown for 48 hours and used for inoculation. All inoculums were 10% v/v. Growth of AGF was monitored using the pressure accumulation method described previously (Henske et al., 2018), which aligns with *C.*

*churrovis* concentrations measured *via* optical density and dry cell weight (Leggieri et al., 2021a) and *N. lanati* dry cell weight concentrations (Wilken et al., 2021).

Stirred cultures were grown with one 25 mm cylindrical magnetic stir bar in each bottle (autoclaved in the bottle) (Part No. SWN660, Globe Scientific, Mahwah, NJ) with constant stirring at 100 rpm on a VELP 15-position digital magnetic stir plate (Part No. F203A0180, VELP Scientifica, Usmate Velate, MB, Italy). Unstirred biofilm cultures were grown identically to stirred cultures, except with no stir bars in the bottles.

Alongside three blank media bottles, six biological replicate *C. churrovis* cultures were grown for each condition. Three were harvested for RNA extraction, and three were used to complete the growth and metabolite curves.

## 2. Microscopy

AGF were imaged using a Zeiss Primovert transmitted light microscope (Part No. 415510-1101-000, Carl Zeiss Microscopy GmbH, Gottingen, Germany). Samples were imaged with a 10× air objective using a slide and coverslip without fixation. Images were collected and analyzed using a SPOT Idea 28.2 5-MP camera and the SPOT 5.1 imaging software (SPOT Imaging, Sterling Heights, MI). Lateral magnification was 2.5 μm/pixel.

## 3. High performance liquid chromatography (HPLC) and gas chromatography (GC) analysis of metabolites

HPLC analysis was conducted as described previously (Leggieri et al., 2021a). Xylan and ethanol were detected using a refractive index detector; succinate, lactate, formate, fumarate, and acetate were detected using a variable wavelength detector set to 210 nm. Standards were created for all sugars and metabolites in deionized water at 1%, 0.1%, and

0.01% w/v concentrations and the above protocol was followed to run each standard.

Standard curves were created via linear regression;  $R^2$  values ranged from 0.9996 to 1.000.

Hydrogen was quantified *via* GC as described previously (Leggieri et al., 2021a). Hydrogen standards including 500 ppm H<sub>2</sub>, 2% H<sub>2</sub>, 5% H<sub>2</sub>, and 20% H<sub>2</sub> with balance helium (Douglas Fluid & Integration Technology, Prosperity, SC), were run at each measurement timepoint to account for the detector baseline that varied slightly each day. Standard curves were created via linear regression;  $R^2$  values ranged from 0.9890 to 0.9979.

Metabolic flux measurements for each metabolite were calculated based on measurements one and two days after inoculation as follows: the difference in amount (mmol) of that metabolite in each culture divided by the average calculated AGF mass (gram dry weight, GDW) present in the culture bottle during that time, divided by the elapsed time between the two measurements. Flux units reported here are mmol GDW<sup>-1</sup> h<sup>-1</sup>.

#### 4. RNA extraction, library preparation, and sequencing

The volume of each culture to harvest for RNA extraction was calculated to yield approximately the same mass of AGF cells. At the time of harvesting, a sample from each culture was added to RNAlater (Part No. R0901, Sigma-Aldrich, St. Louis, MO) in a 1:1 volume ratio and immediately centrifuged for 20 minutes at 4 °C and 12,000g. Following centrifugation, supernatants were removed, and each sample was flash-frozen in liquid nitrogen and stored at -80 °C for one week.

Cell pellet samples were removed from storage, thawed on ice, and lysed via liquid nitrogen grind using a pestle and mortar. Total RNA was then extracted from each lysed cell pellet with the RNeasy Mini Kit (Qiagen, Hilden, Germany) following the provided protocol for filamentous fungi, including the QIAshredder (Part No. 79656, Qiagen) and on-column

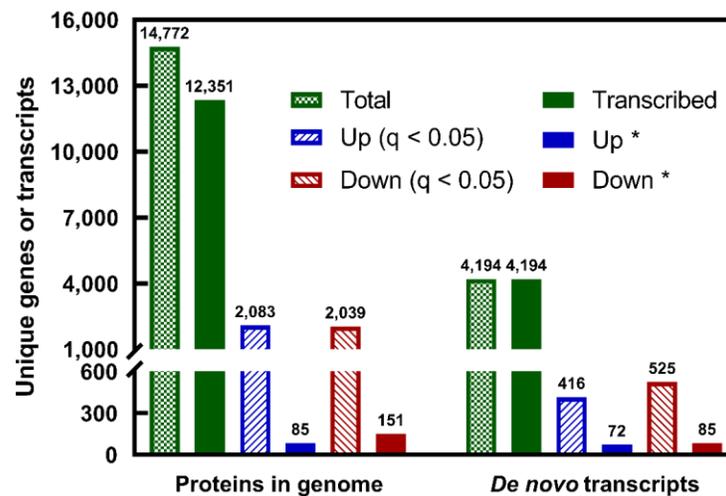
DNase digest (Part No. 79254, Qiagen). 3  $\mu$ L aliquots were taken from each purified RNA sample for quantification and quality analysis with a Qubit (Part No. Q33327, Thermo-Fisher Scientific) and TapeStation (Agilent), respectively. All RNA samples had RINe values above 9 and concentrations above 100 ng/ $\mu$ L.

Stranded RNA-Seq libraries were created with the NEBNext Ultra II RNA Library Prep Kit for Illumina (Part No. E7775, New England BioLabs, Ipswich, MA) and quantified via qPCR. Libraries were sequenced in the Biological Nanostructures Lab at the University of California, Santa Barbara using an Illumina NextSeq 500 with high output and 75 basepair single end reads (Illumina, San Diego, CA).

#### 5. RNA-Seq data analysis

Filtered reads from each library were checked for quality using fastQC [<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>] and trimmed with Trimmomatic (Bolger et al., 2014). The *C. churrovii* genome was downloaded from MycoCosm [<https://mycoCosm.jgi.doe.gov/>] and proteins with duplicate sequences within the genome were removed except for one representative protein (with the lowest protein ID). Trimmed reads were then aligned to this duplicate-checked *C. churrovii* genome using HISAT2 (Kim et al., 2019). StringTie2 (Kovaka et al., 2019) was used to assemble splice-aware transcripts from the aligned reads and to assemble *de novo* transcripts from the unaligned reads (6.1% of the total reads). *De novo* transcripts that overlapped with existing genes in the genome were removed; the reads initially attributed to those *de novo* transcripts were assigned to the existing genes with which they overlapped. DESeq2 (Love et al., 2014) was used to determine which genes showed statistically significant per-cell differential expression between the stirred and unstirred culture conditions.

To provide greater statistical specificity for phenotypic changes in stirred vs. biofilm cultures, stringent requirements beyond  $q < 0.05$  were imposed for a gene to be considered significantly differentially expressed. Namely, each protein or *de novo* transcript must also have a  $\log_2$ fold change  $> 1.5$  (magnitude) between conditions to select for large-magnitude differences in expression, and TPM  $> 2$  to omit genes that are extremely low in transcript number but technically meet the  $q < 0.05$  cutoff. While no universally accepted TPM cutoff threshold exists for differential expression analysis, one is included here to omit extremely low-abundance transcripts (or sequencing/assembly artifacts) from the differentially expressed gene sets (*e.g.*, 10 total transcript reads in stirred cultures and 0 in unstirred cultures). **Figure 7** shows number of detected transcripts that mapped to the genome or were assembled as *de novo* transcripts, the number of these genes/transcripts that were differentially expressed ( $q < 0.05$ ), and significantly differentially expressed according to the more stringent cutoffs. As seen in **Figure 7**, these cutoffs substantially reduce the number of genes/transcripts that are considered significantly differentially expressed.



**Figure 7.** Total proteins in *C. churrovii* genome, de novo transcripts assembled, and statistically significant up- and down-regulated genes ( $q < 0.05$ ) shown in dashed bars. Transcribed, and significantly differentially expressed with cutoffs (marked with \*,  $q < 0.05$ ,  $\log_2$ fold change  $> 1.5$ , max TPM  $> 2$ ) genes shown in solid bars.

CloudBLAST (Matsunaga et al., 2008) was used to annotate *de novo* transcripts assembled by StringTie2. Conserved protein domains within unannotated genes and *de novo* transcripts were identified using the Conserved Domain Search (CD-Search) tool (Lu et al., 2020; Marchler-Bauer and Bryant, 2004). CAZyme annotations were assigned to genes and *de novo* transcripts based on functional annotations in MycoCosm (Grigoriev et al., 2014) and dbCAN (Yin et al., 2012), using a cutoff of two tools to determine CAZyme classification. Gene Ontology (GO) numbers (The Gene Ontology Consortium, 2021; The Gene Ontology Consortium et al., 2000) and CAZyme annotations were assigned to detected genes and *de novo* transcripts using Python. DeepLoc (Almagro Armenteros et al., 2017) was used to localize transcripts to the cytosol, hydrogenosome, or other cellular compartments; here, it is assumed that localization to the mitochondria or plastid corresponds to the hydrogenosome in AGF. SignalP 5.0 was used to determine which genes and *de novo* transcripts are most likely secreted (Almagro Armenteros et al., 2019). UCLUST (Edgar, 2010) was used to cluster unannotated *de novo* transcripts and “hypothetical proteins” that aligned to the *C. churrovii* genome based on sequence similarity.

## 6. Statistical analysis

All statistical analyses (other than determination of differential expression with DESeq2) were conducted using the Prism 9.1.2 software (GraphPad, San Diego, CA). Prism 9.1.2 was used to (i) interpolate the concentrations of metabolites detected via HPLC and GC using standard curves, (ii) determine significant differences in growth rates and metabolite fluxes between growth conditions via *t*-tests, (iii) and conduct contingency tests (Fisher’s Exact Tests) to determine significant up/downregulation of gene classes (a “class” being a particular CAZyme type, a GO term, etc.). Although most gene classes showed some genes

significantly upregulated and some genes significantly downregulated in biofilm culture, Fisher's Exact Test determines whether a gene class is significantly overrepresented in the up/downregulated gene sets based on the number of genes in that class within the genome or transcriptome, the total number of genes in the genome or transcriptome, and the total number of genes up/downregulated. In all statistical tests,  $\alpha = 0.05$  was used.

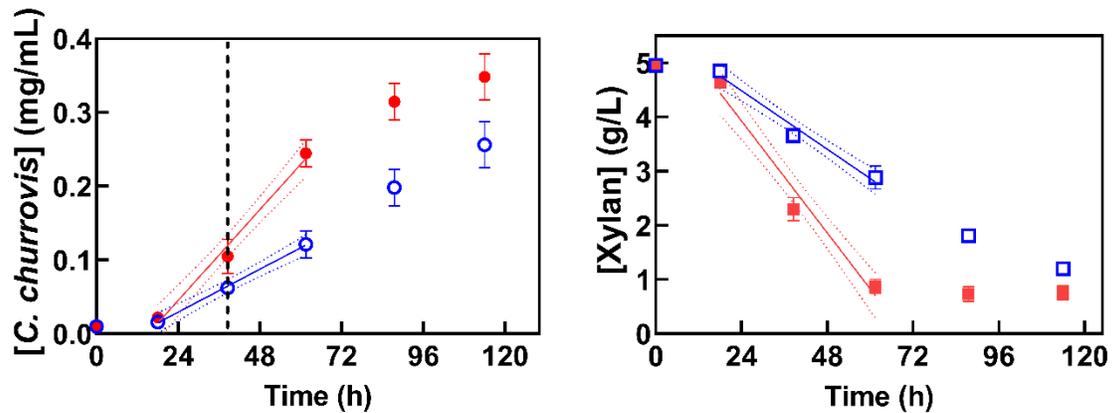
## C. Results and Discussion

### 1. Stirring prevents *C. churrovis* biofilm formation while enhancing xylan degradation and AGF growth

Stirring produces physical forces within microbial cultures that are expected to impact culture function. Stirring within AGF culture vessels introduces mixing and turbulence, thereby enhancing the mass transfer of substrates, enzymes, and metabolites to/from AGF cell surfaces, which is in turn expected to increase growth rates in AGF cultures which may be otherwise limited by diffusion. However, stirring also introduces shear stresses, which in *C. churrovis* disrupts biofilm formation. At present time, it is not understood if biofilm disruption by stirring impacts AGF phenotypes and culture outcomes.

To investigate the effects of stirring, the growth rate and xylan degradation rate of *C. churrovis* (**Figure 8A** and **8B**, respectively) were measured in stirred cultures and biofilm-forming cultures; both rates were significantly greater in stirred cultures ( $p < 0.0001$ ,  $p < 0.0001$  respectively). This suggests that growth of *C. churrovis* on xylan in biofilms is limited by diffusion, rather than its intrinsic growth rate. Further, the disruption of *C. churrovis* biofilms either does not elicit a phenotypic response that slows growth and xylan degradation, or these effects are outweighed by the mass transfer advantages of mixing.

Therefore, biofilm disruption and culture mixing via stirring yields an easily achieved, potentially scalable *C. churrovii* phenotype for polysaccharide degradation.



**Figure 8.** A) *C. churrovii* grew significantly faster on xylan when stirred (solid blue points) compared to unstirred (hollow red points),  $p < 0.0001$ . Dashed line represents time of harvest for RNA extraction during mid-exponential growth. B) *C. churrovii* degraded xylan significantly faster in stirred culture (solid blue points) than in unstirred culture (hollow red points),  $p < 0.0001$ . Dotted lines represent the 95% confidence intervals of each regression.

2. Catabolic CAZymes are significantly upregulated in stirred, faster-growing *C. churrovii* cultures relative to biofilm-forming cultures

AGF are promising candidates for bioprocessing primarily thanks to their abundant and diverse repertoire of CAZymes. However, no prior studies have compared CAZyme expression in biofilm vs. non-biofilm AGF cultures, and bacterial and aerobic fungal literature suggests significant differences likely exist (Dumitrache et al., 2017; Gutiérrez-

**Table 2.** Differentially expressed CAZymes labeled by CAZyme type. Fisher's Exact Tests show that the catabolic CAZyme classes (glycoside hydrolase, pectin lyase, carbohydrate esterase) are significantly downregulated in biofilm culture relative to stirred culture ( $p < 0.05$ ). No CAZyme classes are significantly upregulated in biofilm culture relative to stirred culture. 197 *de novo* CAZyme transcripts that did not map to the existing *C. churrovii* genome were identified, along with 47 transcripts with DOC2 domains (see **Supplementary Spreadsheets S1 and S2**).

CAZyme type	CAZyme count in genome	CAZyme count in <i>de novo</i> transcriptome	Genes upregulated	Genes downregulated	p-value Up	p-value Down
Glycoside hydrolase	304	81	5	15	0.2544	<b>0.0001</b>
Pectin lyase	9	11	1	3	0.1532	<b>0.0019</b>
Carbohydrate esterase	257	23	1	8	0.7329	<b>0.0243</b>
Glycosyltransferase	137	12	0	0	0.6385	0.2692
Carbohydrate binding module	504	70	5	11	0.8139	0.1747
Dockerin domains (DOC2)	392	47	6	9	0.1842	0.1243

Correa et al., 2012; Zhao et al., 2019). As shown in **Table 2**, per-cell expression of the catabolic CAZyme classes (glycoside hydrolase, pectin lyase, carbohydrate esterase) were significantly upregulated in stirred cultures relative to biofilm-forming AGF cultures, as determined by Fisher's Exact Test (see **Supplementary Spreadsheet S1** for a list of CAZymes identified in the *de novo* transcriptome with sequences).

While downregulation of catabolic CAZymes in AGF biofilms contrasts previous findings in cellulolytic bacteria and aerobic fungi (Dumitrache et al., 2017; Gutiérrez-Correa et al., 2012; Zhao et al., 2019), this may be due to slower, diffusion limited growth of the fungus in biofilms, rather than specific repression of CAZymes brought on by the biofilm morphology alone. It is likely that growth is enhanced in the well-mixed stirred cultures mostly due to enhanced mass transfer of enzymes and substrates, and that upregulation of catabolic CAZymes plays only a minor role in increasing growth rate.

In fact, the observed per-cell upregulation of catabolic CAZymes may be a consequence of faster growth in *C. churrovis*, rather than a cause. AGF downregulate CAZymes in response to increased concentration of monosaccharides released from polysaccharides (Henske et al., 2018; Solomon et al., 2016). They may also downregulate CAZymes under increased concentration of growth-inhibiting metabolic end products like formate and acetate. In co-culture with a methanogen, a system shown to grow faster than AGF mono-culture likely due to methanogen uptake of formate and hydrogen (Leggieri et al., 2021a), *C. churrovis* upregulated some CAZymes relative to mono-culture (Swift et al., 2019). In turbulent stirred cultures, these inhibitors and monosaccharides (secreted or produced at or near the cell surface) are in lower local concentration near the cells than in

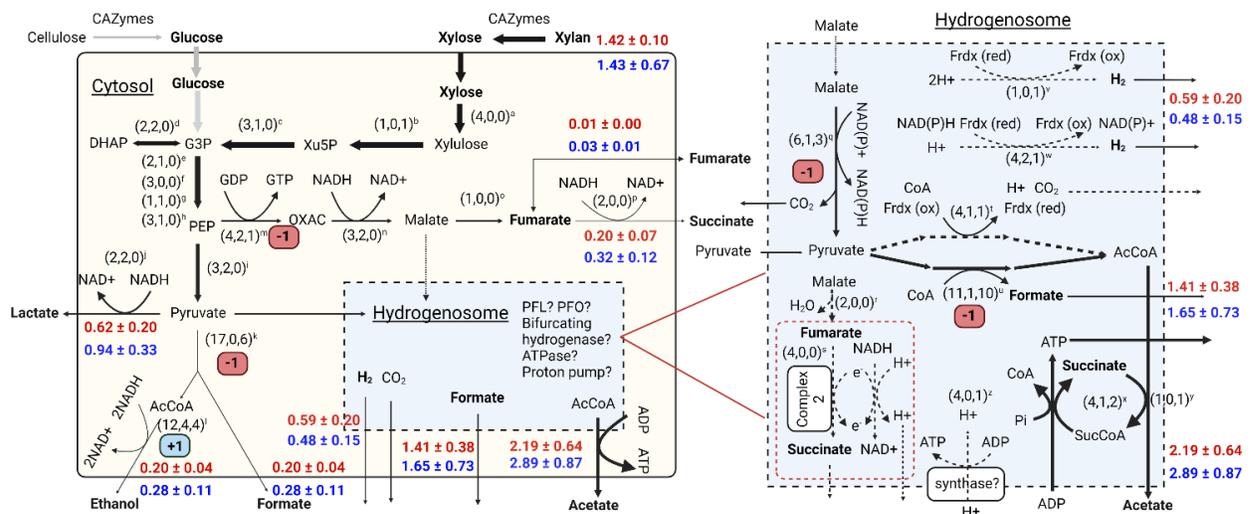
biofilm cultures in which they diffuse away from the cells on timescales that could influence gene expression (Shamir et al., 2016).

Taken together, the results presented here and previous findings show that in two distinct systems (stirred cultures and co-cultures with methanogens), upregulation of CAZymes in *C. churrovis* is correlated with faster fungal growth and decreased local concentration of metabolic inhibitors. While enhanced CAZyme production may be more of a consequence than a cause of faster AGF growth in turbulent culture, stirring is a simple approach to grow AGF more quickly and with more CAZymes produced per cell, offering an approach to scale up non-rhizoidal AGF cultures that may be uniformly sampled to assess and optimize culture conditions.

### 3.3) Metabolic fluxes and expression of primary metabolic enzymes are unchanged in stirred and biofilm-forming *C. churrovis* cultures

Primary metabolic flux determines the profile of fermentation products that can be made or used downstream for bioproduction. Therefore, it is critical to understand how biofilm disruption impacts AGF metabolic flux if stirring or agitation is to be used to control AGF growth. As shown in **Figure 9**, no primary metabolic fluxes differed significantly between the stirred and unstirred cultures. Similarly, few primary metabolic enzymes were differentially expressed between the two conditions. While stirred cultures degraded xylan significantly faster, xylan flux did not differ between stirred and unstirred cultures, suggesting similar growth yields in both conditions (gDW/mol<sub>xylan</sub>).

Many genes involved in primary metabolism showed q-values less than 0.05, but with the TPM and log<sub>2</sub>fold change cutoffs applied, only five primary metabolic enzymes were significantly differentially expressed. Pyruvate formate lyase (PFL) expression was



**Figure 9.** Fluxes and differential expression of central metabolism (glycolysis, TCA cycle, and hydrogenosome) of stirred and unstirred *C. churrovii* cultures. Fluxes in blue represent stirred cultures, red represent unstirred cultures. No fluxes were significantly different. Expression statistics are reported for each enzyme as (number of unique transcripts detected, number of upregulated transcripts, number of downregulated transcripts) using  $q < 0.05$  as the only threshold for significant regulation. Enzymes meeting the  $\log_2$ fold-change and TPM thresholds are denoted with numbers in colored boxes (blue significantly upregulated and red significantly downregulated). Enzymes [with EC numbers if available]: a) xylose isomerase [5.3.1.5] b) xylulokinase [2.7.1.17] c) glycoaldehydetransferase [2.2.1.1] d) triose-phosphate isomerase [5.3.1.1] e) glyceraldehyde 3-phosphate (G3P) dehydrogenase [1.2.1.12] f) phosphoglycerate kinase [2.7.2.3] g) G3P dehydrogenase (NADP<sup>+</sup>) [1.2.1.9] h) phosphopyruvate hydratase [4.2.1.11] i) PEP synthase [2.7.9.2] j) lactate dehydrogenase [1.1.1.28] k) pyruvate formate-lyase (PFL) cytosolic [2.3.1.54] l) alcohol dehydrogenase [1.1.1.1] m) PEP carboxykinase (GTP) [4.1.1.32] n) malate dehydrogenase [1.1.1.37] o) fumarase cytosolic [4.2.1.2] p) succinate dehydrogenase [1.3.5.1] q) malic enzyme [1.1.1.40] r) fumarase hydrogenosomal [4.2.1.2] s) complex 2 subunits A, B, C, D t) pyruvate ferredoxin oxidoreductase PFO u) PFL hydrogenosomal [2.3.1.54] v) iron-hydrogenase protein w) bifurcating hydrogenase subunits A, B, G x) succinyl coenzyme A (CoA) synthase y) acetate succinyl CoA transferase z) ATP synthase subunits A, B, G, D. Created with BioRender.com.

upregulated in both the cytosol (k) and hydrogenosome (u) in stirred cultures. Increased PFL expression could also be a consequence of faster AGF growth rather than a direct cause, as its upregulation here did not lead to increased PFL flux. Like catabolic CAZymes, PFL was upregulated in faster growing (Leggieri et al., 2021a) co-cultures of *C. churrovii* with a methanogen (Brown et al., 2021), although PFL flux also increased in co-culture systems.

Other significantly differentially regulated genes include an alcohol dehydrogenase (l), which was upregulated in unstirred biofilm cultures, and a phosphoenolpyruvate decarboxylase (m) and malic enzyme (q), which were upregulated in stirred cultures. It is possible that the upregulation (in biofilms) of an ethanol-producing enzyme in a strictly

cytosolic pathway and downregulation of two genes involved in malate production and eventual utilization in the hydrogenosome suggest a correlation between hydrogenosome activity and faster AGF growth (as shown in (Leggieri et al., 2021a)). However, no significant differences in hydrogenosomal fluxes were observed, and the uncertainty of the hydrogenosomal pathways in **Figure 9** precludes quantitative assessment of this hypothesis.

The GO Enrichment analysis in **Table 3** corroborates the phenotypic shift towards more plant cell wall catabolism in stirred cultures vs. biofilm cultures with negligible changes in primary metabolism. GO terms vary widely in specificity, however GO Enrichment analysis is useful to identify large-scale phenotypic differences between conditions if multiple related GO functions are overexpressed in one condition and not the other. The bold GO terms in **Table 3A** all relate to cell wall catabolism and/or hydrolysis of lignocellulose in plant cell walls, and they are all significantly overrepresented in the downregulated gene set and not the upregulated gene set, supporting the observed significant downregulation of the catabolic CAZyme classes in biofilm culture relative to stirred culture. In **Table 3B**, the overrepresented GO terms in biofilm cultures are not obviously connected to each other in a way that represents a phenotypic shift. This is expected, as AGF biofilm-forming machinery is unannotated and therefore lacks GO identification.

The combined gene expression and flux data suggest that primary metabolism is not significantly altered in AGF stirred vs. biofilm cultures. Interestingly, this implies that CAZyme expression and primary metabolic flux may not be correlated in AGF mono-cultures. Therefore, researchers may choose to stir AGF cultures for faster polysaccharide degradation and CAZyme production without altering the fermentation product profile.

**Table 3.** Enrichment analysis of significantly down-regulated and up-regulated GO IDs determined by Fisher’s Exact Test. Several GO IDs enriched in the down-regulated gene set (more expressed with stirring) relate to cell wall catabolism (bolded). GO IDs enriched in the up-regulated gene set (more expressed in biofilms) appear unrelated.

GO ID	GO Name	GO Type	p-value
(A) Down-regulated (i.e., more expressed with stirring)			
<b>GO:0004623</b>	<b>phospholipase A2 activity</b>	Molecular function	0.0026
<b>GO:0016042</b>	<b>lipid catabolic process</b>	Biological process	0.0026
GO:0055085	transmembrane transport	Biological process	0.03055
GO:0009332	glutamate-tRNA ligase complex	Cellular component	0.03274
GO:0004818	glutamate-tRNA ligase activity	Molecular function	0.03601
<b>GO:0009253</b>	<b>peptidoglycan catabolic process</b>	Biological process	0.0376
<b>GO:0016998</b>	<b>cell wall catabolic process</b>	Biological process	0.03822
GO:0031681	G-protein beta-subunit binding	Molecular function	0.04001
<b>GO:0008422</b>	<b>beta-glucosidase activity</b>	Molecular function	0.04494
GO:0000750	pheromone-dependent signal transduction involved in conjugation with cellular fusion	Biological process	0.04501
GO:0003796	lysozyme activity	Molecular function	0.04778
(B) Up-regulated (i.e., more expressed in biofilms)			
GO:0008519	ammonium transmembrane transporter activity	Molecular function	0.009
GO:0004252	serine-type endopeptidase activity	Molecular function	0.0063
GO:0000796	condensin complex	Cellular component	0.0276
GO:0006508	proteolysis	Biological process	0.03279
GO:0070534	protein K63-linked ubiquitination	Biological process	0.03312
GO:0000974	Prp19 complex	Cellular component	0.0365
GO:0090730	Las1 complex	Cellular component	0.04106
GO:0001584	rhodopsin-like receptor activity	Molecular function	0.0414

### 3.4) Differential expression identifies proteins that may be associated with biofilm formation

AGF biofilm-forming genes, or other genes directly involved in response to shear or biofilm disruption, have not yet been identified and annotated. However, they would likely be secreted and upregulated in biofilm-forming vs. stirred AGF cultures. To search for these

genes and spur future mechanistic research, the functions of the significantly differentially expressed unannotated genes and *de novo* transcripts, which comprise most of the dataset, were investigated. It is possible that biofilm-forming machinery is expressed in stirred cultures even though shear prevents biofilm development. Thus, unannotated genes and *de novo* transcripts downregulated in biofilm cultures were included in this search as well.

To annotate the significantly differentially expressed unannotated genes and *de novo* transcripts or identify homology with genes in known biofilm-forming organisms, the *de novo* transcripts and existing genes with “hypothetical protein” or “expressed protein” annotation were BLASTed via CloudBLAST (e-value threshold  $10^{-4}$ , % ID > 70%). This resulted in putative annotation of 100 of the 236 significantly downregulated genes and *de novo* transcripts (42%) and 69 of the 157 significantly upregulated genes and *de novo* transcripts (44%). This BLAST did not reveal any obvious biofilm-related functionality; however, some of these differentially expressed genes with secretion tags could possibly be involved in extracellular biofilm formation or modification. Therefore, all significantly differentially expressed *de novo* transcripts and secreted proteins were BLASTed using more lenient parameters (e-value threshold  $10^{-3}$ , % ID > 40%), identifying several proteins and *de novo* transcripts with homology to known biofilm and/or adhesion-related proteins, and even more proteins with no homology to any other proteins (see **Supplementary Table S1**).

The unannotated, significantly differentially expressed, *de novo* transcripts and secreted proteins are candidates for further investigation into their role in AGF biofilm function. Many show homology to subtilisin-like proteins, extracellular serine proteases produced by the most widely studied model biofilm-forming bacterium, *Bacillus subtilis*. While subtilisins typically degrade extracellular biofilms (Leroy et al., 2008; Liu et al., 2021;

Mitrofanova et al., 2017) rather than form them, their differential expression here suggests possible involvement in AGF biofilm growth and/or remodeling. As expected, more subtilisin-like secreted proteins and *de novo* transcripts were upregulated in biofilm cultures than downregulated. Protein 625875 could be influential to the mechanical properties of *C. churrovis* biofilms, as it is homologous with several eukaryotic titin-like proteins, large proteins which influence the elasticity of muscle tissues (Granzier and Labeit, 2004). Several differentially expressed *de novo* transcripts have chitin-binding GO annotations. While this GO category was not significantly overrepresented in the biofilm-upregulated gene set, these previously unannotated transcripts may participate in binding of AGF cell walls (predominantly made of chitin) to other cells and/or to substrates.

Some *C. churrovis* differentially expressed secreted “hypothetical proteins” and *de novo* transcripts showed no homology to any AGF proteins and could therefore contribute to *C. churrovis*’ unique morphology relative to the rest of the phylum (see **Supplementary Table S1**). Of these proteins present uniquely in *C. churrovis* (and no other AGF), some showed similarity with proteins involved extracellular cell adhesion or proteolysis, processes involved in biofilm formation and maturation. Others showed no appreciable homology to any known proteins and warrant further investigation to develop a better mechanistic understanding of whether or how they contribute to AGF biofilms.

Even after relaxing BLAST parameters, most significantly differentially expressed genes and *de novo* transcripts remained unannotated. To search these still-unannotated genes and transcripts for classes of proteins co-expressed to form biofilms or respond to shear, they were investigated for sequence similarity clustering. A greater proportion of the differentially expressed unannotated *de novo* transcripts clustered together (>25% sequence similarity)

than the unannotated differentially expressed proteins in the *C. churrovis* genome (see **Supplementary Figure S1**). These clusters of *de novo* transcripts were generally larger but showed lower sequence identity than the clusters of unannotated proteins in the genome.

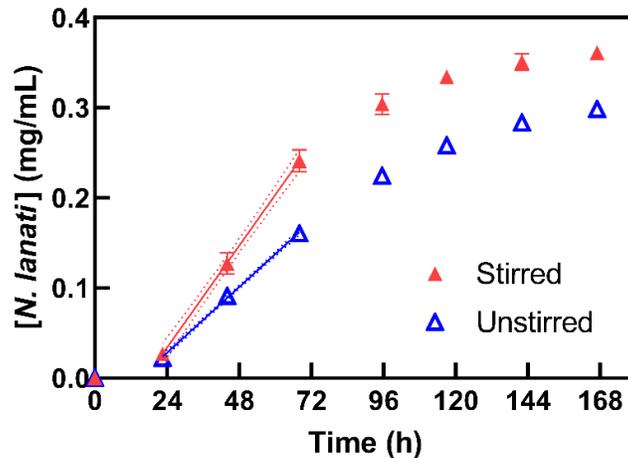
While these clusters may be performing coordinated biological functions, their conserved protein domains showed no obvious relationships within each cluster, complicating discovery of co-regulated biofilm-implicated genes. However, their co-expression is at least correlated with either biofilm formation or stirred cultivation under shear. Therefore, their structures may be of interest for future investigation of AGF biofilm-forming or shear response machinery. A master spreadsheet of all detected genes and *de novo* transcripts, their expression statistics, their amino acid or nucleotide sequence, whether they are most likely secreted, and their functional annotation (with any CAZyme prediction) is provided for further investigation (see **Supplementary Spreadsheet S2**).

3.5) Rhizoidal *N. lanati* grew significantly faster on solid milled grass with stirring

To determine whether the growth advantages observed in *C. churrovis* on soluble substrates also occur in suspended cultures of rhizoidal AGF, the properties of stirred cultures of rhizoidal *N. lanati* on a milled grass substrate were tested. As shown, non-rhizoidal AGF such as *C. churrovis* have promise for industrial scale up; however, rhizoidal AGF comprise most of the phylum, and should therefore be considered for scale-up as well.

**Figure 10** shows that stirred cultures of rhizoidal *N. lanati* on solid reed canary grass grew significantly faster than unstirred cultures ( $p < 0.0001$ ). *N. lanati* growth requires physical association with the plant substrate, i.e., formation of small biofilms around the grass. Therefore, biofilm vs. non-biofilm direct comparison experiments are not achievable with this strain. However, these results suggest that physical association occurs in the

turbulent stirred cultures, and the enhanced mass transfer introduced by stirring helps to accelerate AGF growth and biomass degradation. This is promising for use of rhizoidal AGF in mixed solid and liquid state cultures for future applications.



**Figure 10.** Rhizoidal *N. lanati* grew significantly faster in stirred culture on a solid milled plant substrate compared to unstirred culture ( $p < 0.0001$ ). Dotted lines represent the 95% confidence intervals of each regression.

#### D. Conclusions

AGF, promising lignocellulose degraders, form biofilm-like mats which putatively facilitate biomass breakdown but have precluded industrial-scale cultures. In non-rhizoidal *C. churrovis*, biofilms are easily disrupted with stirring, thereby increasing AGF growth rate, catabolic CAZyme expression, and xylan degradation rate, with similar metabolic fluxes to biofilm-forming cultures. Stirring also increases the growth rate of rhizoidal *N. lanati* on solid plant substrate, a system that could resemble prototypical scaled-up AGF anaerobic digestion. Biofilm disruption in non-rhizoidal AGF *via* stirring enhances CAZyme production and polysaccharide degradation without altering the fermentation product profile, offering another knob to turn to optimize scaled, biomass-valorizing AGF cultures.

### **III. Non-destructive quantification of anaerobic gut fungi and methanogens in co-culture reveals increased fungal growth rate and changes in metabolic flux relative to mono-culture**

Reprinted with permission from (Leggieri et al., 2021a).

Quantification of individual species in microbial co-cultures and consortia is critical to understanding and designing communities with prescribed functions. However, it is difficult to physically separate species or measure species-specific attributes in most multi-species systems. Anaerobic gut fungi (AGF) (Neocallimastigomycetes) are native to the rumen of large herbivores, where they exist as minority members among a wealth of prokaryotes. AGF have significant biotechnological potential owing to their diverse repertoire of potent lignocellulose-degrading carbohydrate-active enzymes (CAZymes), which indirectly bolsters activity of other rumen microbes through metabolic exchange. While decades of literature suggest that polysaccharide degradation and AGF growth are accelerated in co-culture with prokaryotes, particularly methanogens, methods have not been available to measure concentrations of individual species in co-culture. New methods to disentangle the contributions of AGF and rumen prokaryotes are sorely needed to calculate AGF growth rates and metabolic fluxes to prove this hypothesis and understand its causality for predictable co-culture design.

We present a simple, microplate-based method to measure AGF and methanogen concentrations in co-culture based on fluorescence and absorbance spectroscopies. Using samples of < 2% of the co-culture volume, we demonstrate significant increases in AGF growth rate and xylan and glucose degradation rates in co-culture with methanogens relative to mono-culture. Further, we calculate significant differences in AGF metabolic fluxes in co-

culture relative to mono-culture, namely increased flux through the energy-generating hydrogenosome organelle. While calculated fluxes highlight uncertainties in AGF primary metabolism that preclude definitive explanations for this shift, our method will enable steady-state fluxomic experiments to probe AGF metabolism in greater detail.

The method we present to measure AGF and methanogen concentrations enables direct growth measurements and calculation of metabolic fluxes in co-culture. These metrics are critical to develop a quantitative understanding of interwoven rumen metabolism, as well as the impact of co-culture on polysaccharide degradation and metabolite production. The framework presented here can inspire new methods to probe systems beyond AGF and methanogens. Simple modifications to the method will likely extend its utility to co-cultures with more than two organisms or those grown on solid substrates to facilitate the design and deployment of microbial communities for bioproduction and beyond.

## **A. Introduction**

Microbial communities continue to attract significant attention from researchers in microbiology, engineering, agriculture, medicine, and beyond owing to their ability to perform seemingly limitless chemical transformations (Leggieri et al., 2021b). Physical and metabolic interactions in microbial communities present challenges for quantifying population specific growth rates, metabolic fluxes, and other characteristic metrics. Developing easy, rapid, and non-invasive methods to characterize consortium membership is critical. However, microbes in natural consortia are difficult to physically separate and can even form biofilms (Brethauer et al., 2020), making colorimetric or spectroscopic methods difficult to deploy. Here, we describe the identification of microbe-specific spectroscopic signals that enable quantification of growth rates and fluxes in co-cultures of anaerobic gut

fungi (AGF) and methanogenic archaea (methanogens). Ultimately, these metrics enable testing of hypotheses related to their biomass valorization performances in co-culture relative to mono-culture.

AGF native to the rumen of large herbivores have promise for sustainable and economical degradation of lignocellulosic biomass and conversion to value-added products such as pharmaceuticals and commodity chemicals (Hooker et al., 2019), especially if they can be deployed in consortia with other rumen-native microbes. AGF possess nature's greatest quantity and variety of biomass-degrading carbohydrate-active enzymes (CAZymes) (Solomon et al., 2016), which are readily produced in laboratory culture to degrade a variety of lignocellulose, polysaccharide, oligosaccharide, and monosaccharide substrates for downstream conversion to value-added products (Henske et al., 2018b; Hooker et al., 2019). AGF physically associate and exchange metabolites with bacteria and methanogens in the rumen (Lee et al., 2000; Leng, 2017; Solden et al., 2018), leading some to suggest that interactions between AGF and prokaryotes significantly enhance both the rate and extent of biomass degradation relative to isolated AGF (Li et al., 2019; Marvin-Sikkema et al., 1990).

To leverage this effect for industrial bioproduction, researchers have formed "top-down" microbial consortia via laboratory culture of microbes enriched from herbivore fecal samples (Gilmore et al., 2019). While communities with AGF and prokaryotes outperform AGF mono-cultures in biogas production rate (Gilmore et al., 2019) and show increased biomass degradation relative to solely prokaryotic communities (Peng et al., 2021), the mechanisms (gene regulation, flux redirection, etc.) that yield these desired outcomes are difficult to probe (Leggieri et al., 2021b). Top-down rumen-derived consortia exhibit interwoven syntrophies that could inform model-based design of simpler, more tractable



microbial communities. However, even in two-member co-cultures including AGF, the biofilm-like morphology of AGF caused by their extensive extracellular rhizoidal network (**Figure 11A, Supplementary Figure S2**) and their physical associations with prokaryotes preclude non-destructive quantification of each species, obscuring how co-culturing with prokaryotes alters the growth rate, per-cell metabolic activity, and CAZyme secretion associated with AGF.

In co-cultures with AGF, methanogens remove putatively inhibitory AGF fermentation products such as formate and hydrogen (Joblin and Naylor, 1993; Wilken et al., 2020), which might accelerate AGF growth and biomass deconstruction. However, published results are mixed regarding whether these co-cultures show significantly different rates of biomass degradation relative to AGF mono-cultures (Bauchop and Mountfort, 1981; Gilmore et al., 2019; Hungate, 1969; Joblin et al., 1990; Joblin and Williams, 1991; Marvin-Sikkema et al., 1990; Mountfort et al., 1982; Teunissen et al., 1992), and have been unable to quantify individual species growth rates or cell mass-normalized metabolic fluxes (Li et al., 2017). These metrics directly affect the rate and extent of biomass degradation and the profile of metabolites produced; therefore, they are irreplaceable if anaerobic communities with AGF are to be deployed for bioproduction (Lillington et al., 2020), especially if predictive metabolic models are to be developed for co-cultures with AGF (Wilken et al., 2018). Quantitative polymerase chain reaction (qPCR)-based methods may be developed to estimate species concentrations in microbial communities (Junicke et al., 2014; Traversi et al., 2012), and researchers have used them to quantify AGF in mono-culture (Dollhofer et al., 2016); however, these methods require thorough optimization, and can be time consuming to run, motivating the development of simpler methods.

In this method, we leverage the non-rhizoidal morphology of the AGF species, *Caecomyces churrovis* (Henske et al., 2017), to form homogeneous shaken or stirred synthetic co-cultures of AGF with the autofluorescent rumen methanogen, *Methanobrevibacter thaueri* (Miller and Lin, 2002), that can be sampled for growth and flux quantification by harvesting as little as 0.5 mL of the culture at each timepoint (**Figure 11A**). Despite lacking extracellular rhizoids, *C. churrovis* produces a quantity and variety of biomass degrading enzymes comparable to rhizoidal AGF (Henske et al., 2017). *C. churrovis* is metabolically similar to other AGF, as 97% of its enzyme commission (EC) numbers are shared with at least one other AGF, making it a promising model AGF (**Supplementary Table S2**).

We utilize non-interfering fluorescence intensity (characteristic of all methanogens) and optical density measurements to resolve the concentrations of both species in co-culture simultaneously, unlocking growth curves and metabolic fluxes for AGF in co-culture with prokaryotes for the first time. To date, measurements of gas accumulation have been the standard for indirectly tracking AGF growth. Optical density measurements offer a simple, more direct measurement of AGF concentration. We test and validate our method by assessing whether co-culturing with *M. thaueri* significantly alters the growth rate or metabolic flux of *C. churrovis* in defined media on both mono- and polysaccharide substrates. Concentration-normalized AGF metabolic flux measurements highlight major discrepancies with widely-accepted models of the AGF hydrogenosome, an energy-generating organelle directly involved in metabolite exchange with methanogens and production of biorefinery precursors including formate and acetate. Further, limitations of the method are discussed, including how it may be extended to quantify individual species

populations in microbial communities with more or different species, as well as those without an autofluorescent organism. Ultimately, this method enables per-species measurements that are critical for the design and deployment of biotechnologically relevant microbial communities.

## **B. Materials and Methods**

### **1. Culture of anaerobic gut fungi and methanogens**

A modified version of anaerobic Medium B (MB) (Theodorou et al., 1995) was used for both routine culture and growth experiments of AGF and methanogens, available in the online version of the article (Leggieri et al., 2021a); methanogen mono-cultures received the yeast extract and casitone/peptone supplements, and AGF mono- and co-cultures were grown in the fully defined formulation. The headspace of AGF mono- and co-cultures was 5% H<sub>2</sub>, 35% CO<sub>2</sub>, balance N<sub>2</sub>, and the headspace of methanogen mono-cultures was 80% H<sub>2</sub>, balance CO<sub>2</sub>. All cultures were grown at 39 °C in 75 mL anaerobic serum bottles with 45 mL of liquid medium. AGF were grown on the soluble carbon sources glucose (anhydrous, Thermo Fisher Scientific, Waltham, MA) or xylan (from corn core, TCI America, Portland, OR) at final concentrations of 5 g/L.

*Caecomyces churrovis* was previously isolated from the feces of a large herbivore (Henske et al., 2017), and isolated *Methanobrevibacter thaueri* was purchased from DSMZ (DSM 11995). For routine culture, *C. churrovis* was transferred to new media every 2-4 days, and *M. thaueri* was transferred every 4-10 days. In all growth experiments, starter cultures of AGF and methanogens were grown for 48 hours and used for inoculation. All inoculums were 10% v/v. Growth of AGF and methanogens was monitored using the pressure accumulation method described previously (Henske et al., 2018b) and the combined

fluorescence intensity and optical density method outlined in **Figure 11**. Each condition (co-culture and mono-culture) was grown in biological triplicate, alongside three blank media bottles.

## 2. High performance liquid chromatography (HPLC) analysis of sugars and metabolites

Sulfuric acid (50 mM) was added (10% v/v) to AGF hydrolysate supernatant samples that were then vortexed, held at room temperature for 5 minutes, then centrifuged for 5 minutes at 21,000g. The supernatants were then dispensed into HPLC vials and run on an Agilent 1260 Infinity HPLC (Agilent, Santa Clara, CA) using a Bio-Rad Aminex HPX-87H column (Part No. 1250140, Bio-Rad, Hercules, CA) with an inline 0.22  $\mu\text{m}$  filter (Part No. 50671551, Agilent) followed by a Micro-Guard Cation H guard column (Part No. 1250129, Bio-Rad, Hercules, CA) before the analytical column. Samples were run with a 5 mM sulfuric acid mobile phase at a flow rate of 0.6 mL/min and a column temperature of 50 °C. Glucose, xylan, and ethanol were detected using a refractive index detector; succinate, lactate, formate, fumarate, and acetate were detected using a variable wavelength detector set to 210 nm. Standards were created for all sugars and metabolites in deionized water at 1%, 0.1%, and 0.01% w/v concentrations and the above protocol was followed to run each standard. Standard curve  $R^2$  values ranged from 0.9996 to 1.000.

Metabolic flux measurements for each metabolite were calculated based on measurements one and two days after inoculation as follows: the difference in amount (mmol) of that metabolite in each culture divided by the average calculated AGF mass (gram dry weight, GDW) present in the culture bottle during that time, divided by the elapsed time between the two measurements. Flux units reported here are  $\text{mmol GDW}^{-1} \text{h}^{-1}$ .

### 3. Gas chromatography analysis of hydrogen and methane

To analyze the headspace composition of each culture at each measurement timepoint, 100  $\mu$ L of headspace gas was collected and subsequently purged three times in a 100  $\mu$ L air-tight syringe and needle. Then, 20  $\mu$ L of headspace gas was collected and injected directly into a Thermo Fisher Scientific TRACE 1300 gas chromatograph (Thermo Fisher Scientific) with a TracePLOT™ TG-BOND Msieve 5 A (Part No. 26003-6100, Thermo Fisher Scientific) and an Instant Connect Pulsed Discharge Detector (PDD) (Part No. 19070014, Thermo Fisher Scientific). The oven temperature for each run was 30 °C and the PDD temperature was 150 °C. High-purity helium (Part No. HE 5.0UH-55, Praxair, Danbury, CT) was further purified with a heated helium purifier (Part No. HP2, VICI) and used as the carrier gas with a flow rate of 0.5 mL/min. The same flushing and analysis procedures were followed for methane and hydrogen standards (SPECS, respectively) including 500 ppm H<sub>2</sub>, 2% H<sub>2</sub>, 5% H<sub>2</sub>, 20% H<sub>2</sub>, 0.5% CH<sub>4</sub>, 1% CH<sub>4</sub>, 5% CH<sub>4</sub>, 10% CH<sub>4</sub>, and 20% CH<sub>4</sub> with balance helium (Douglas Fluid & Integration Technology, Prosperity, SC), which were run at each measurement timepoint to account for the PDD baseline that varied slightly each day. Standard curve R<sup>2</sup> values ranged from 0.7370 to 0.9979.

### 4. Quantification of anaerobic fungi and methanogens via plate reader and lyophilizer

AGF and methanogens were quantified with optical density at 450 nm and fluorescence intensity (excitation/emission: 425 nm/ 470 nm, bandwidth 10 nm) using a Tecan M1000 Infinite Microplate Reader (Tecan, Männedorf, Switzerland). Fluorescence intensity measurements were obtained with a manual gain setting of 77 for each sample and blank and were normalized by the fluorescence intensity of aliquots of Pacific Blue dye (100  $\mu$ g/L) (succinimidyl ester, Thermo Fisher Scientific, Canoga Park, CA). UV-transparent 96

well plates (Part No. 3635, Corning) were used for optical density measurements, and black MicroFluor2 96 well plates were used for fluorescence intensity measurements (Part No. 437111, Thermo Fisher Scientific).

The dry cell weights of AGF and methanogens in culture vessels at the end of growth were determined by harvesting and centrifuging the cultures (10,000g for 20 minutes) in tared centrifuge tubes, washing the cell pellets with deionized water and centrifuging again, lyophilizing for 48 hours in a FreeZone 4.5 Liter Benchtop Freeze Dry System (Part No. 77500200, Labconco Corp., Kansas City, MO), and weighing the dried samples in the centrifuge tubes.

## 5. Microscopy

Micrographs of AGF and methanogens were captured with a Leica SP8 resonant scanning confocal microscope (Leica Microsystems, Wetzlar, Germany) with photomultiplier tube (PMT) and HyD detectors and 405 nm, argon, and white light lasers. AGF were imaged using the white light laser and transmitted light PMT to collect brightfield images, and methanogens were imaged using the 405 nm excitation laser with HyD detector set to detect emission between 460 and 480 nm. Images were collected and analyzed using the LAS X Life Science Microscope Software Platform (Leica Microsystems). Samples were imaged without fixation using a slide and coverslip. A 20x water objective (numerical aperture = 0.75) was used to collect all images presented here. Lateral magnification was 284 nm/pixel.

## 6. Statistical analysis

All statistical analyses were conducted using the Prism 9.1.2 software (GraphPad, San Diego, CA). Prism 9.1.2 was used to (i) interpolate the concentrations of metabolites detected via HPLC and GC using standard curves, (ii) determine significant differences in growth

rates and metabolite fluxes between growth conditions via *t*-tests, (iii) compare calculated and measured co-culture concentrations via *t*-tests, and (iv) to determine the significance of the slopes and intercepts of linear regressions. In all statistical tests,  $\alpha = 0.05$  was used.

## C. Results and Discussion

### 1. Development of a non-destructive co-culture species quantification method

All AGF with PacBio-sequenced genomes to date, except for *C. churrovis*, grow in dense, biofilm-like mats that cannot be uniformly sampled, precluding direct measurement of cell concentrations without harvesting and weighing the entire culture (Peng et al., 2018; Theodorou et al., 1995). Therefore, growth of AGF in mono-culture is typically tracked via pressure accumulation in sealed culture vessels, as AGF produce hydrogen and likely carbon dioxide as they grow (Peng et al., 2018; Wilken et al., 2020). However, uncertainties in the regulation of and relative flux through gas-generating pathways coupled with the pH-dependent, gas-evolving bicarbonate buffer present in most AGF media (Peng et al., 2018; Vinzelj et al., 2020) make pressure accumulation an indirect measure of AGF growth. Using pressure as a proxy for AGF growth precludes analysis of per-cell hydrogen and carbon dioxide production and quantification of AGF concentration in continuous or even semi-batch cultivation systems, both of which are critical for the eventual deployment of AGF for industrial biotechnology. Further, because methanogens utilize hydrogen and carbon dioxide gasses as well as formate produced by AGF to synthesize methane, it is not possible to account for the total moles of gas produced by AGF in co-culture, preventing pressure-based tracking of the growth of either species in co-culture altogether.

Quantification of two species simultaneously in co-culture requires two independent signals that scale linearly with the concentration of either species or the concentration of the

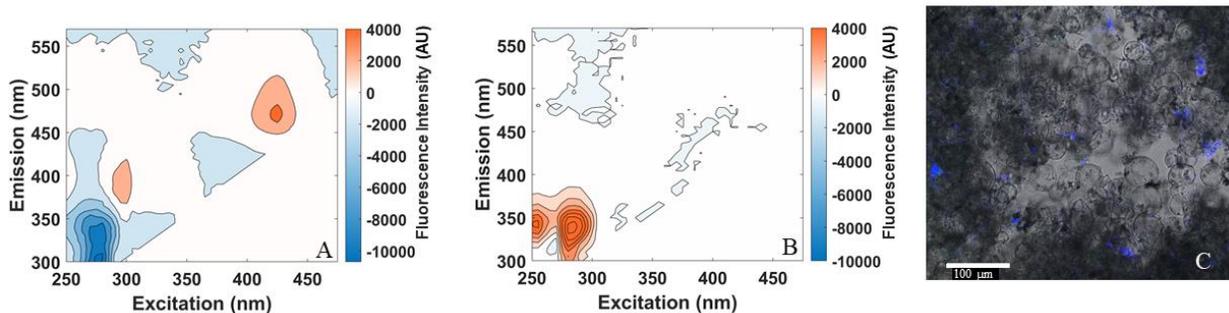
total co-culture. If a signal scales with the concentration of one species, but not the other, then the presence of the other species must interfere with the signal from the first. Further, the co-culture must be well-mixed enough to enable uniform sampling, and the signals must be measurable with a small enough sample of the culture that growth is not disturbed. In the well-mixed AGF-methanogen co-cultures studied here, we use fluorescence intensity to quantify the methanogen and optical density to quantify the total mass concentration of the co-culture. We use the linear relationship between fluorescence intensity and absorbance during exponential phase in the methanogen to calculate the contribution of the methanogen to the total optical density signal, enabling calculation of the population-specific optical density and therefore concentration of the AGF. The equations and propagation of uncertainty associated with the method are given in **Figure 11B** and **Supplementary Table S3**, respectively.

To track growth and metabolite production in mono- and co-cultures, the workflow outlined in **Figure 11B** was conducted at each timepoint. First, the accumulated pressure was measured for each culture vessel and blank media vessel. Next, the headspace gas of each culture vessel was sampled and analyzed via GC. Then, 1 mL of each well-mixed culture vessel and blank was sampled with a needle and syringe and transferred to a microcentrifuge tube. Each tube was vortexed briefly, then pipetted into two separate wells each on clear microplates for measurements of optical density and black microplates for measurements of fluorescence. The remaining volume in each microcentrifuge tube was centrifuged, and the supernatant removed and stored at -80 ° C for HPLC analysis. Finally, the culture and blank vessels were vented to 1 psig.

2. Normalized fluorescence intensity quantifies absolute methanogen concentration in co-culture

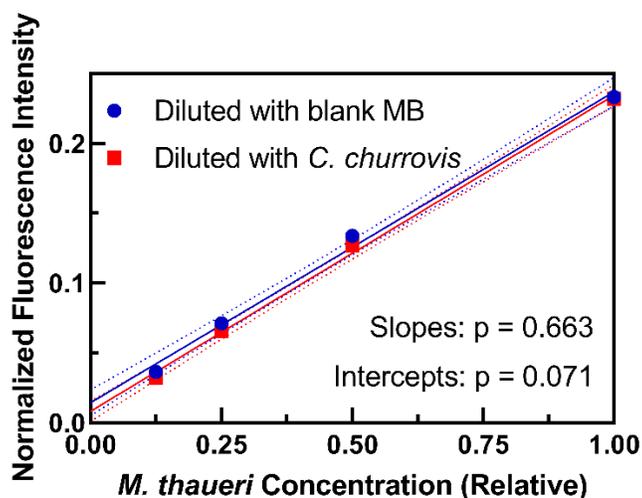
Methanogens are quantifiable with autofluorescence intensity due to the fluorescent coenzyme F<sub>420</sub>, which is present in all methanogens and is involved in all three major routes of anaerobic methanogenesis (hydrogenotrophic, acetoclastic, and methylotrophic) (Greening et al., 2016). The fluorescence spectrum of coenzyme F<sub>420</sub> is well-characterized, with expected peak excitation and emission wavelengths near 420 nm and 470 nm, respectively (Doddema and Vogels, 1978). The intracellular F<sub>420</sub> content has been shown to be constant in methanogens during exponential-phase growth (Heine-Dobbernack et al., 1988; Taya et al., 1986), supporting the use of fluorescence intensity as a direct methanogen concentration measurement. Because fluorescence intensity units are arbitrary and values are subject to vary with nuisance variables such as the lamp power in the microplate reader, ambient temperature, etc., we normalize the observed fluorescence intensity of all samples in a run by the observed fluorescence intensity of a freshly thawed aliquot of Pacific Blue dye. Pacific Blue dye has a similar fluorescence spectrum to methanogens (excitation/emission max: 410 nm/ 455 nm) and is subject to the same nuisance variables as the samples. Therefore, the normalized fluorescence intensity of a methanogen culture ( $F_{\text{culture}}/F_{\text{dye}}$ ) may be used as an absolute measurement of methanogen concentration, as long as all aliquots of the Pacific Blue dye are of identical concentration. We observed no change in the fluorescence intensity of Pacific Blue dye dissolved in dimethyl sulfoxide over 15 months of storage at -20 °C (**Supplementary Figure S3**).

As seen in **Figure 12A**, we observed the expected fluorescence spectrum for *M. thaueri*, with peak excitation and emission wavelengths at 425 and 470 nm, respectively; we observed negligible fluorescence in *C. churrovis* in this channel (**Figure 12B**). The micrograph shown in **Figure 12C** confirms that *M. thaueri* is visible in co-culture with *C. churrovis* using a 425/470 nm fluorescent filter, and that *C. churrovis* shows no fluorescence in this channel and is visible only in the brightfield overlay.



**Figure 12.** Illustrations of biofilm-like morphology of rhizoidal AGF which cannot be uniformly sampled to track growth in co-culture, and non-rhizoidal AGF such as *C. churrovis*, which form well-mixed co-cultures when shaken or stirred and enable tracking of both species' concentrations (A) via the method outlined in B. Created with BioRender.

As shown in **Figure 13**, the normalized fluorescence intensity of *M. thaueri* monocultures scaled linearly with cell concentration when cells were diluted with blank Medium B, and the slope and intercept of this regression was not significantly different when *M. thaueri* was diluted with concentrated *C. churrovis* instead of blank medium (slopes  $p = 0.663$ , intercepts  $p = 0.071$ ). This further verifies that there is no measurable fluorescence of *C. churrovis* in the 425/470 nm channel and demonstrates that the presence of *C. churrovis* does not interfere with the fluorescence signal of *M. thaueri*. These results therefore establish that normalized fluorescence intensity may be used to quantify the absolute concentration of methanogens in co-culture with *C. churrovis* without physically separating the cell populations.



**Figure 13.** *M. thaueri* (combined pellet and supernatant) normalized fluorescence intensity scales linearly with cell concentration when diluted with blank Medium B, and with a mature *C. churrovius* culture, indicating that *C. churrovius* does not interfere with the fluorescent signal from *M. thaueri*. Dotted lines represent the 95% confidence interval of each regression. The p-values represent a test for significant differences in the values of the slopes and intercepts of the two regressions.

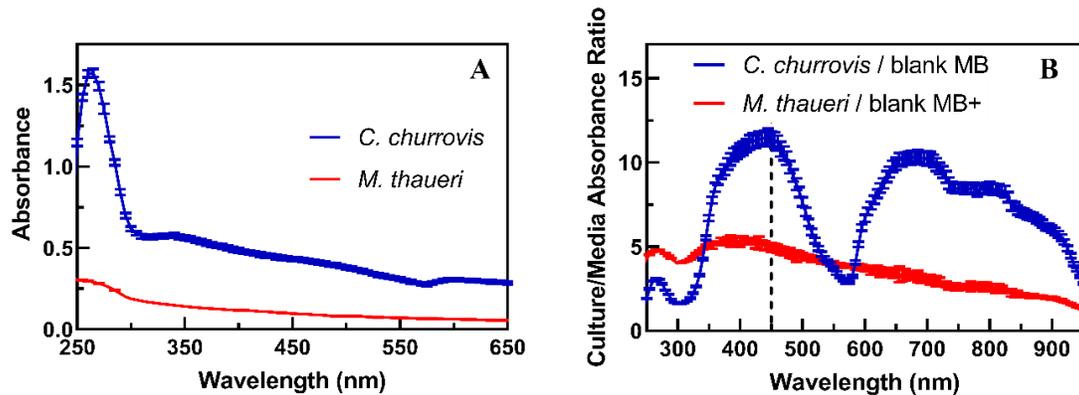
### 3. Optical density quantifies the concentration of AGF and methanogens in co-culture

Optical density, often at 600 nm, is a well-established measurement of cell concentration for model microbes (Beal et al., 2020). However, the dense biofilm-like rhizoidal morphology characteristic of almost all AGF isolated and routinely cultured to date precludes uniform sampling of the culture for quantification via optical density. *C. churrovius* lacks this rhizoidal phenotype and can therefore be uniformly sampled for quantification via optical density when grown with constant stirring or when shaken prior to sampling (see **Supplementary Figure S2** for a macroscopic visual comparison of rhizoidal AGF and *C. churrovius* morphologies).

As seen in **Figure 14A**, the peak absorbance values for *C. churrovius* and *M. thaueri* are both near 260 nm. However, the absorbance of blank Medium B is also large in this ultraviolet (UV) region, and the ratio of cell absorbance to media absorbance is at a minimum here (**Figure 14B**). Further, variable oxidation states of cofactors and other intracellular

metabolites yield appreciable variation in per-cell UV-range absorbances from day to day and batch to batch, making them an unreliable measure of absolute cell concentration.

Although the magnitude of absorbance at 450 nm is less than that at 260 nm for both *C. churrovii* and *M. thaueri*,  $A_{450 \text{ nm}}$  scales linearly with cell concentration for both species and offers the largest signal to background media ratio (**Figure 14B**), therefore 450 nm was used to determine the total cell concentration in co-cultures.



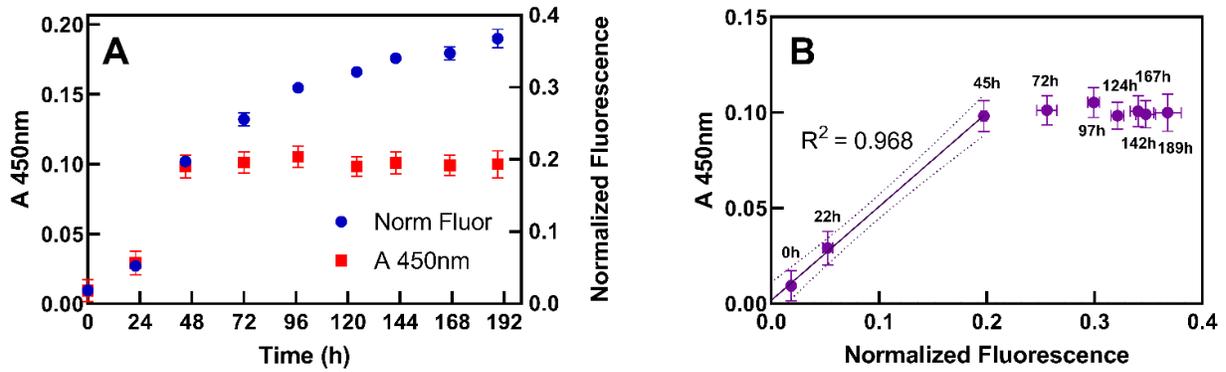
**Figure 14.** Absorbance spectra of late-exponential phase mono-cultures of *C. churrovii* and *M. thaueri* (with media blanks subtracted) (A) show peak values in UV range, but highest culture/media absorbance ratios at or near 450nm (B), highlighted with the dashed line. MB+ indicates the Medium B formulation with yeast extract and casitone added.

To calculate the optical density of *C. churrovii* in co-culture, we require an estimate of the optical density of *M. thaueri*, which is subtracted from the observed total co-culture optical density ( $A_{450 \text{ nm}}$  of the co-culture) to give the optical density of *C. churrovii* (mathematical steps outlined in **Figure 11B**). To estimate the optical density of *M. thaueri* using the normalized fluorescence intensity of the culture, the ratio of normalized fluorescence to absorbance at 450 nm must be constant for *M. thaueri* in mono-culture. As shown in **Figure 15B**, the relationship between absorbance and fluorescence intensity is linear ( $R^2 = 0.968$ ) during the exponential phase of growth (0 - 45 h in this case). Therefore, we assume that ratio of normalized fluorescence and absorbance is equal to the average value

of the ratio during this period (1.93) for *M. thaueri* in co-culture. As shown in **Figure 15A**, normalized fluorescence provides an accurate estimate of the optical density of *M. thaueri* in mono-culture up to the point where the methanogen reaches stationary phase (when absorbance stops increasing, after 45 h in this case). Beyond this point, the fluorescence of the culture continues to increase while the absorbance remains constant, yielding the nonlinear relationship between absorbance and fluorescence after 45 h shown in **Figure 15B**.

This divergent relationship between fluorescence and absorbance in stationary phase may be attributable to increased secretion of coenzyme F<sub>420</sub> by the methanogens in stationary phase relative to exponential phase, and a greater fluorescence intensity of secreted F<sub>420</sub> than intracellular F<sub>420</sub>. Some previous descriptions of fluorescence-based methanogen quantification recommend removing any culture supernatant and lysing the methanogens to measure only intracellular (and not extracellular) F<sub>420</sub> (Peck, 1989). This approach yielded only minor increases in fluorescence intensity compared to the unlysed methanogen pellets (**Supplementary Figure S4A**). Further, unlike the combined pellet and supernatant samples shown in **Figure 13**, methanogen pellet fluorescence did not scale linearly with methanogen concentration when diluted with concentrated *C. churrovis* (**Supplementary Figure S4B**). Therefore, extracellular fluorescence was included in quantification of the methanogen in all mono-cultures and co-cultures.

Most batch, semi-batch, and continuous co-cultures prioritize exponential-phase growth; therefore, the divergent ratio of fluorescence to absorbance in methanogens in stationary phase poses minimal practical drawbacks. To accurately calculate stationary phase methanogen concentrations in co-cultures, we assume both species are at stationary phase when the total absorbance of the co-culture stops increasing with time, and assume that the



**Figure 15.** The slope of the ratio of fluorescence to absorbance vs. time during exponential phase (48h and before) is not significantly different from zero (A,  $p = 0.9845$ ), demonstrating that fluorescence may be used to estimate the optical density of the methanogen in co-culture during exponential-phase growth. Normalized fluorescence and absorbance at 450 nm scale linearly during exponential phase growth (B), in agreement with previously observed constant intracellular coenzyme  $F_{420}$  content in methanogens. Dotted lines in (B) represent the 95% confidence interval of the linear regression of the ratio of fluorescence to absorbance vs. time during exponential phase.

absorbance (and concentration) of the methanogen remains constant at its initial stationary phase value even though fluorescence continues to increase. The ratio of normalized fluorescence to absorbance at 450 nm in *M. thaueri* is based on experimental values, therefore, we treat it as a random variate ( $1.93 \pm 0.13$ ) in all calculations to increase the sensitivity of statistical conclusions drawn based on the calculated concentrations of both species. As the observed fluorescence of a co-culture increases, so too does the uncertainty in the absolute methanogen concentration, and therefore in the AGF concentration as well (**Supplementary Table S3**).

**Table 4** shows the average ratios of optical density at 450 nm to cell concentration (determined *via* measurement of the culture dry weight after lyophilization) for *C. churrovii* and *M. thaueri* in mono-culture from six separate cultures of each species spread across two different batches with different inoculums.

**Table 4.** Individual species absorbance/concentration correlations for six replicate mono-cultures from two different inoculums for each species.

Organism	Cell concentration (mg/mL)	Absorbance (450 nm)	Abs/Conc (mL/mg)
<i>C. churrovis</i>	0.449 ± 0.043	0.235 ± 0.027	0.523 ± 0.078
<i>M. thaueri</i>	0.102 ± 0.015	0.100 ± 0.002	0.997 ± 0.148

Correlations for both species show % coefficients of variation of less than 15%

The coefficients of variation for both species are below 15%, and the predicted total mass concentrations for co-cultures based on these correlations are not significantly different from the observed values (**Table 5**), supporting the accuracy of the method.

Correlations are calculated using cell weight rather than cell number because the lytic lifecycle of AGF in which many zoospores develop inside a sporangium complicates the definition and detection of a single fungal “cell” via hemocytometry. While the concentrations of both species are calculable using only the fluorescence intensity and absorbance measurements outlined above, weighing the lyophilized co-cultures is a third, independent metric that may be used to validate the calculated concentrations.

**Table 5.** Total concentrations of co-cultures grown on xylan and glucose measured via lyophilization (left) and calculated using the individual species absorbance measurements (**Figure 16; Supplementary Figure S5**) and the absorbance/concentration correlations in **Table 4** (right).

Substrate	Measured co-culture concentration (mg/mL)	Calculated co-culture concentration (mg/mL)	p
Xylan	0.472 ± 0.006	0.442 ± 0.028	0.2005
Glucose	0.504 ± 0.016	0.505 ± 0.042	0.9721

On both substrates, the calculated concentration is not significantly different from the measured concentration

#### 4. Potential expansions of the method to other co-culture systems

The method described here may be extended to any co-cultures which can be grown in well-mixed systems and possess two linearly independent signals such as optical density at a given wavelength, fluorescence intensity in a particular excitation/emission channel,

fluorescence lifetime, fluorescence polarization, or any other signal that reproducibly scales linearly with the concentration of one species or the total concentration of the co-culture. We leverage methanogen autofluorescence as one of the two signals here; for genetically tractable organisms, fluorescence may be introduced via genetic engineering. However, constant expression of the fluorescent protein over the course of growth would be required, which is particularly difficult in anaerobic systems (Ozbakir et al., 2020).

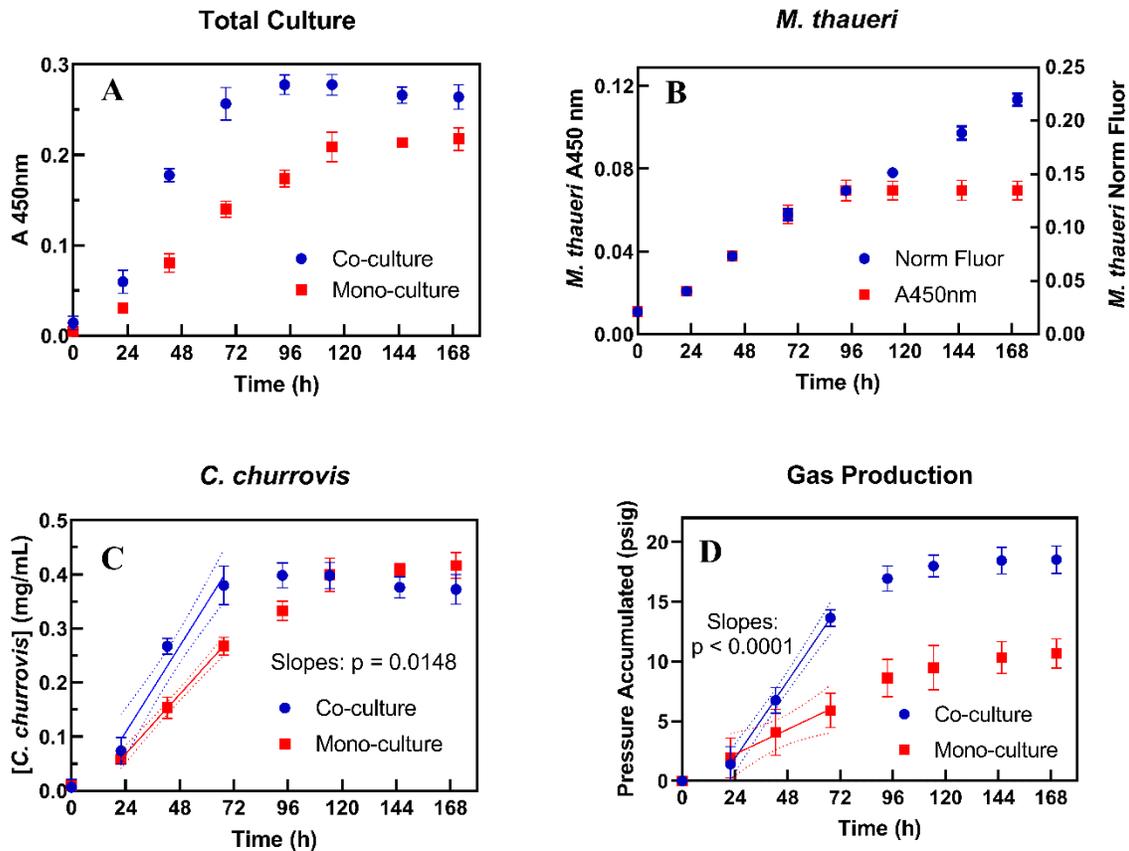
In a simpler case, individual species concentrations may be resolved in a co-culture using absorbance signals at two different wavelengths, provided that the ratio of per-cell absorbance between those two wavelengths is different in the two organisms and constant over the course of growth in both organisms. The absorbance profiles of microbes depend on many factors including their size and intracellular composition; therefore it is likely that two linearly independent wavelengths exist for most co-culture pairs, even for prokaryote-prokaryote systems. For tri-cultures, a third linearly independent absorbance wavelength must exist.

For systems grown on solid substrates such as lignocellulosic biomass, the method we present here may still be applied if the culture (with substrate) can be uniformly sampled, and the microorganisms can be subsequently removed from the substrate entirely, potentially with a detergent-based procedure similar to the one described in (Nettmann et al., 2013). Such systems will likely require more samples, and thus larger cultures, to capture the heterogeneity of the multiphase culture, as well as thorough controlling of the background autofluorescence of the substrate and any detergents.

5. Co-culturing *C. churrovis* with *M. thaueri* significantly increases AGF growth rate and xylan and glucose deconstruction rate relative to AGF mono-cultures

We used the method outlined above to determine whether co-culturing with a methanogen increases the growth rate, polysaccharide and monosaccharide deconstruction rate, and mass-normalized flux of key metabolites in *C. churrovis* relative to mono-culture. We use xylan and glucose as the substrates in separate experiments, as they are soluble in Medium B and quantifiable via HPLC, and therefore allow uniform sampling of the culture for quantification of both species and the chemical composition of the supernatant.

The growth curves of *C. churrovis* and *M. thaueri* as well as the pressure accumulation and total co-culture optical density curves during growth on xylan are shown in **Figure 16**. **Figure 16A** shows that, at all timepoints, the optical density of the co-culture was greater than that of the mono-culture, as expected. Using the measured relationship between fluorescence intensity and optical density for *M. thaueri*, we can estimate how much of the total optical density of the co-culture is attributable to *M. thaueri* (**Figure 16B**), and therefore determine the concentration of *C. churrovis* at each timepoint (**Figure 16C**). Note that the fluorescence of the co-culture increases throughout stationary-phase growth (after 96 h in this case), but the absorbance of the methanogen is assumed constant, as discussed previously. As seen in **Figure 16C**, the slopes of the *C. churrovis* concentration vs. time regressions during the period of approximately constant growth rate are significantly different between co-culture and mono-culture ( $p = 0.0148$ ), indicating that co-culturing with *M. thaueri* does increase the growth rate of *C. churrovis* on xylan. These data represent the first evidence of a significant difference in growth rates of AGF in synthetic co-culture vs. mono-culture using direct AGF concentration-based measurements. Like previous studies (Li et al., 2019), we observe significantly enhanced rates of gas accumulation in co-culture vs. mono-culture (**Figure 16D**,  $p < 0.0001$ ); however, this alone does not demonstrate faster growth of the



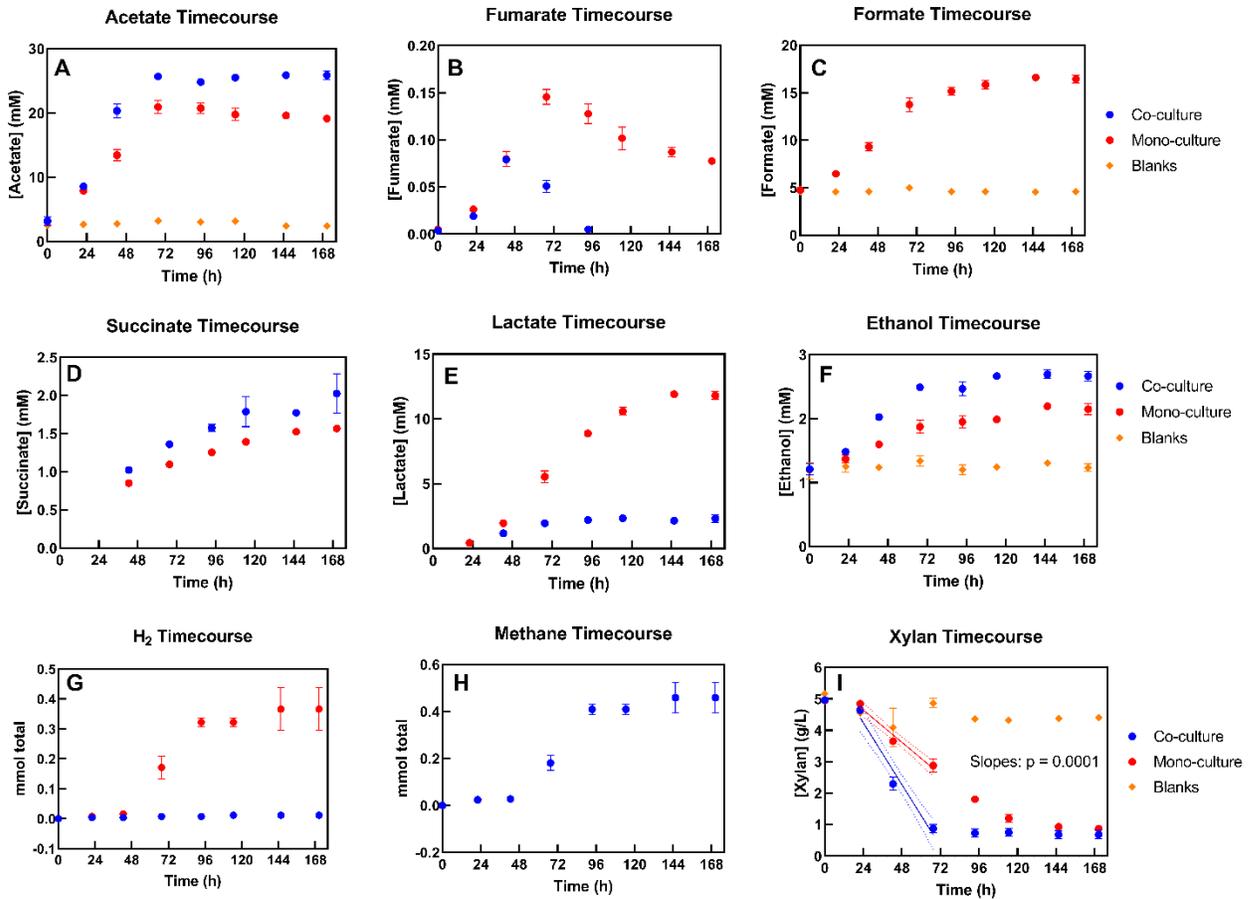
**Figure 16.** Total culture absorbance (A), *M. thaueri* fluorescence + absorbance (B), *C. churrovis* concentration (C) and accumulated pressure (D) curves show that both the growth rate of *C. churrovis* and the rate of gas production are significantly increased in co-cultures with *M. thaueri* grown on xylan, relative to monocultures. Panel B shows the divergence of *M. thaueri* fluorescence relative to absorbance in stationary phase also observed in mono-culture; the absorbance of the methanogen was assumed to remain constant after the absorbance of the co-culture stops increasing (96h and after). Dotted lines represent the 95% confidence interval of each regression. The p-values in panels C and D represents a test for significant difference in the values of the slopes of the two regressions.

fungus, even though pressure accumulation curves typically correlate with AGF

concentration in mono-culture (Theodorou et al., 1995).

As seen in **Figure 17I**, the rate of xylan degradation by AGF was significantly greater in co-culture than mono-culture ( $p = 0.0001$ ). To the best of our knowledge, this is the first evidence of a significantly greater polysaccharide degradation rate in a synthetic AGF-methanogen co-culture compared to an AGF mono-culture (see the supplement of (Gilmore et al., 2019) for a statistical analysis of previous studies related to this conclusion). These data support the previous finding that biomass-degrading CAZymes, including xylanases,

were upregulated in co-cultures of AGF and methanogens relative to AGF mono-cultures (Swift et al., 2019).



**Figure 17.** Metabolite profiles for mono- and co-cultures grown on xylan reveal significantly faster xylan degradation in co-culture (panel D), negligible hydrogen and formate accumulation in co-culture (panels C and G), greater lactate production in mono-culture (panel E), greater acetate, succinate, and ethanol production in co-culture (panels A, D, and F, respectively), and less accumulation of fumarate in co-culture (panel B). Metabolite concentrations at 24h and 43h shown here were used to calculate fluxes of each metabolite. Dotted lines represent the 95% confidence interval of each regression. The p-value in panel I represents a test for significant difference in the values of the slopes of the two regressions.

Previously, it was shown that neither the rate of sugar release from cellulosic filter paper (Gilmore et al., 2019) nor the rate of xylose utilization (Li et al., 2017) by AGF were significantly increased by co-culturing with a methanogen. Interestingly, we observed a similar result in an AGF-methanogen co-culture grown on glucose and inoculated with a seven-day-old methanogen culture that was in stationary phase (**Supplementary Figure S5**).

Methanogen growth in co-culture was confirmed by increasing fluorescence and production of methane, but neither the growth rate of *C. churrovis* nor the rate of glucose utilization differed from mono-cultures ( $p = 0.5509$ ,  $p = 0.1067$ , respectively). However, as seen in **Supplementary Figures S6 and S7**, when the glucose experiment was repeated with a 2 day-old methanogen inoculum, the growth rate of *C. churrovis* was significantly greater in co-culture than mono-culture ( $p = 0.0107$ ), the rate of glucose degradation was significantly greater in co-culture ( $p < 0.0001$ ), and gas productivity was greater in co-culture ( $p = 0.0025$ ). Some dependence of AGF growth rate on the growth phase of the methanogen inoculum may partially explain the variable results of AGF-methanogen co-cultures in literature.

6. AGF-methanogen co-cultures grown on xylan and on glucose show significantly different mass-normalized metabolic fluxes compared to mono-cultures

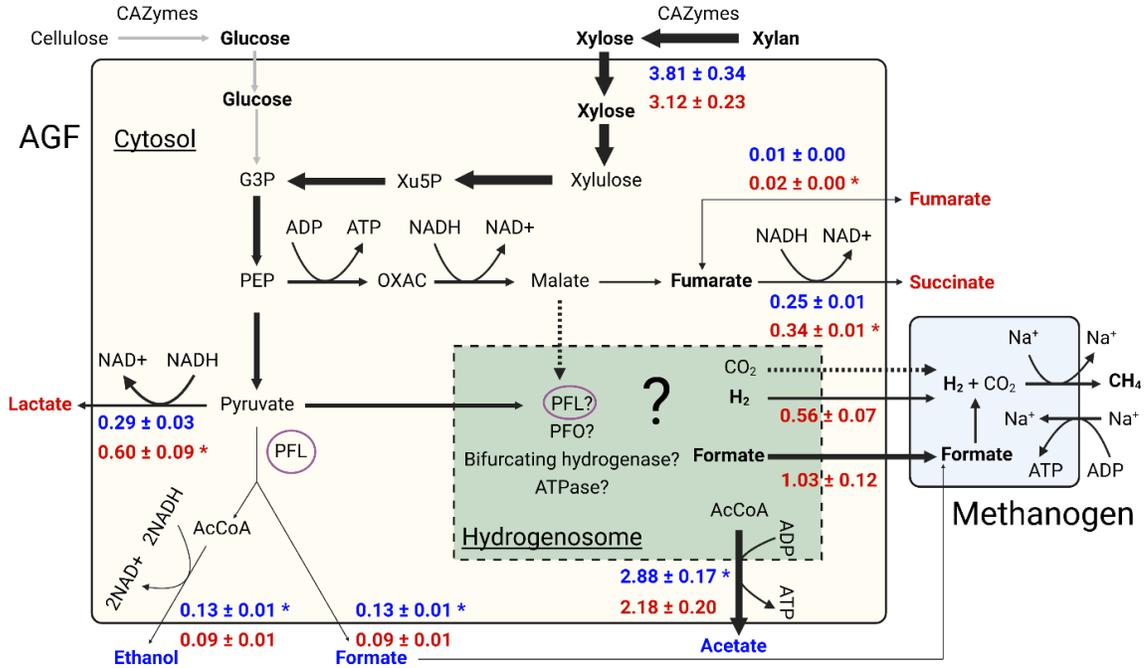
The metabolite concentrations in **Figure 17** combined with the *C. churrovis* concentrations in **Figure 16C** enable calculation of the flux of each metabolite in mono- and co-cultures grown on xylan. Because fluxes are typically most accurately analyzed in the context of predictive metabolic models during steady-state growth (Orth et al., 2010), we present fluxes at only one timepoint, 43h, the middle of the period of approximately constant growth rate. As seen in **Figure 18**, significant differences exist between mono- and co-culture for the fluxes of all metabolites measured except for formate (which was utilized by *M. thaueri*<sup>1</sup>, precluding accurate calculation of formate flux in co-culture) and xylan. The lack of difference in xylan flux between mono- and co-culture implies similar growth yields

---

<sup>1</sup> Formate utilization was observed in *M. thaueri* mono-cultures (not shown) and co-cultures, despite a previous contrary description (Miller and Lin, 2002). Formate dehydrogenase subunits are present in the only published *M. thaueri* genome on NCBI.

(gDW/mol<sub>xylan</sub>) and per-cell xylan deconstruction activities between the two conditions.

Internal fluxes are often compared between conditions by normalizing by the influx of carbon substrate for each condition (Long and Antoniewicz, 2019); in this case, although we are concerned primarily with external fluxes, we may directly compare the absolute flux values of each metabolite in mono- and co-culture because the xylan influxes are similar.



**Figure 18.** AGF mass-normalized fluxes reveal significant upregulation of acetate (via the hydrogenosome) and ethanol (via cytosolic PFL) fluxes, and significant downregulation of lactate and succinate flux in co-cultures. Formate and hydrogen are consumed by *M. thaueri* and therefore do not accumulate in co-cultures. While xylan is consumed more quickly in co-culture, the flux of xylan into *C. churrovii* is equal in mono- and co-cultures. Fluxes for succinate in mono- and co-culture and lactate in co-culture assume metabolite concentrations of 0 mM at 24h, as observed values were below the detection limit. Bolded metabolites are detectable via our HPLC method. Metabolites in blue or red (also starred) showed significantly greater flux in co-culture or mono-culture, respectively. Arrow thickness correlates qualitatively with mono-culture flux values. Xu5P: xylulose-5-phosphate; G3P: glyceraldehyde-3-phosphate; PEP: phosphoenolpyruvic acid; OXAC: oxaloacetic acid; ADP: adenosine diphosphate; ATP: adenosine triphosphate; NAD<sup>+</sup>: nicotinamide adenine dinucleotide (oxidized); NADH: nicotinamide adenine dinucleotide (reduced); AcCoA: acetyl coenzyme A; PFL: pyruvate formate lyase; PFO: pyruvate:ferredoxin oxidoreductase. Created with BioRender.

Although the pathways within the AGF hydrogenosome and their relative utilization remain uncertain (Wilken et al., 2021), the fluxes presented in **Figure 18** support the conclusion by Li et al. that co-culturing with methanogens causes AGF to direct more flux

through the hydrogenosome (Li et al., 2017). While the hydrogenosome model used in that work, proposed previously by Boxma et al. (Boxma et al., 2004), relies on an energetically unfavorable route of hydrogen production (reduction of protons to hydrogen coupled to regeneration of NAD(P)<sup>+</sup> from NAD(P)H), the yield of one mole of acetate per mole of malate or pyruvate that enters the hydrogenosome in that model was supported by a recent genomic and transcriptomic characterization of the AGF hydrogenosome (Wilken et al., 2021). Because acetate is putatively only produced in AGF in the hydrogenosome, and not in the cytosol *via* acetaldehyde dehydrogenase which would reduce additional NAD<sup>+</sup> instead of regenerating it from NADH, we estimate acetate flux as a proxy for hydrogenosome flux (note the uncertainty of hydrogenosomal pathways in **Figure 18**). See (Wilken et al., 2021) for a description of all observed hydrogenosomal transcripts in *Neocallimastix lanati*, an AGF that is metabolically similar to *C. churrovis* (84% of EC numbers shared between both species, **Supplementary Table S2**). While significantly more work is needed to characterize the AGF hydrogenosome, the significantly greater acetate flux in co-culture than mono-culture ( $p = 0.0320$ ) implies increased flux through the hydrogenosome in co-culture than in mono-culture.

In support of the increased hydrogenosome vs. cytosol flux in co-culture, as shown in **Figure 18**, the fluxes of lactate and succinate, metabolites produced to regenerate oxidized NAD<sup>+</sup> from NADH in the cytosol, were significantly lower in co-culture than mono-culture ( $p = 0.0177$ ,  $0.0012$ , respectively). The external flux of fumarate, an intermediate in cytosolic succinate production, was also lower in co-culture than mono-culture ( $p = 0.0197$ ). The only cytosolic flux that increased in co-culture vs. mono-culture was that of ethanol ( $p = 0.0156$ ), however the magnitude of this difference ( $0.03 \text{ mmol/gDW h}$ ) was lesser than the differences

between mono- and co-culture of lactate (0.33 mmol/gDW h) and succinate (0.10 mmol/gDW h).

It is possible that ethanol flux increases in co-culture because formate, a putative inhibitor of AGF growth which is produced during production of ethanol *via* the cytosolic pyruvate formate lyase (PFL) pathway, is taken up by methanogens. Production of ethanol regenerates two NAD<sup>+</sup> from NADH per pyruvate, whereas production of lactate generates only one, giving the AGF incentive to produce ethanol over lactate if accumulation of inhibitory formate is not an issue. This may also explain why AGF redirect more flux through the hydrogenosome in co-culture with methanogens; more ATP can be generated without accumulating inhibitory formate (Li et al., 2017), facilitating faster AGF growth and polysaccharide deconstruction. However, the uncertain and likely degenerate mechanisms of oxidized cofactor regeneration in the hydrogenosome (Wilken et al., 2021) preclude definitive explanation of increased hydrogenosomal fluxes in co-culture.

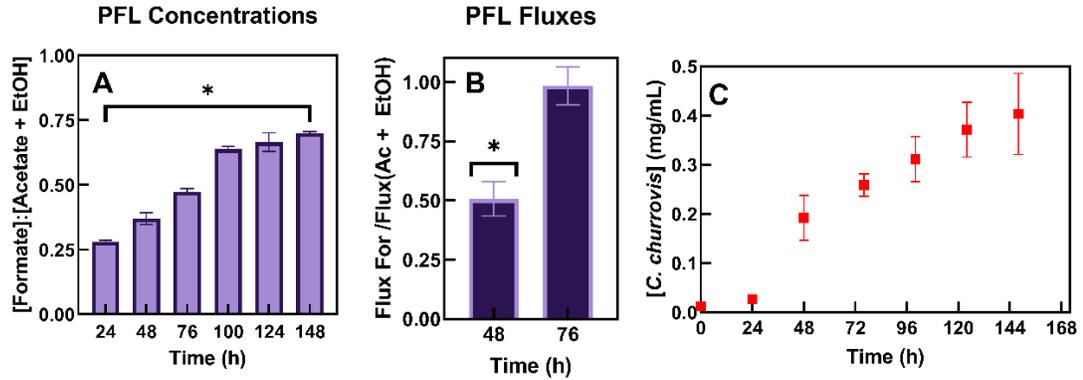
In co-cultures on glucose, significant differences in metabolite fluxes were the same as those described in **Figure 18**, with the exception that succinate fluxes did not differ significantly between mono- and co-culture on glucose (**Supplementary Figure S7**).

#### 7. Quantification of *C. churrovis* external metabolic fluxes highlights gaps in understanding of the AGF hydrogenosome

While a recent description of the AGF hydrogenosome in *N. lanati* showed transcription of several pathways for ATP and hydrogen production (Wilken et al., 2021), analysis of the fluxes of formate, acetate, and ethanol supported the hypothesis initially proposed by Boxma et al. (Boxma et al., 2004) that PFL is the dominant pathway in the hydrogenosome, and flux through other pathways is negligible. Because PFL is present in

both the cytosol and the hydrogenosome (**Figure 18**), the ratio of formate to ethanol plus acetate produced by the AGF will be unity only if acetate production in the hydrogenosome is always coupled to formate production, implying that only the PFL pathway carries significant flux in the hydrogenosome.

This has been observed in at least two AGF to date (Boxma et al., 2004; Wilken et al., 2021); however, as shown in **Figure 19B**, the ratio of formate to acetate plus ethanol fluxes was significantly different from unity during the phase of constant growth (48h) for *C. churrovis* ( $p = 0.0070$ ). Further, the ratio of formate to acetate plus ethanol concentrations was significantly different from unity throughout growth on both substrates (**Figure 19A**) ( $p < 0.0015$  for all timepoints), suggesting that PFL is not the sole dominant pathway in the hydrogenosome in *C. churrovis*. However, during late-exponential growth (76h, **Figure 19B**), the ratio of formate to acetate plus ethanol flux did not differ significantly from unity ( $p = 0.7628$ ), suggesting that hydrogenosome flux is dynamic and highly regulated, and PFL may dominate late in *C. churrovis* growth curves. In a separate study, all hydrogenosome components transcribed in *N. lanati* were also detected in *C. churrovis* (Brown et al., 2021), however, *N. lanati* showed PFL dominance throughout growth while *C. churrovis* did not.



**Figure 19.** Mono-culture fluxes in MB on glucose suggest that *C. churrovius* hydrogenosome pathways differ from widely accepted PFL model. At all timepoints, the ratio of formate to acetate plus ethanol concentrations produced during growth differs significantly from unity (A) ( $p < 0.0015$  for all timepoints). The ratio of formate flux to acetate plus ethanol flux differs significantly from unity during mid-exponential phase growth (48h,  $p = 0.0070$ ), but not during late-exponential phase growth (76h,  $p = 0.7628$ ) (B), suggesting that PFL may dominate hydrogenosome flux only after substrate is depleted and/or inhibitory metabolites have accumulated. Growth curve included (C) for reference of growth phases; 76h and before was considered exponential phase, and all timepoints after were considered stationary phase.

The regulation and relative utilization of other hydrogenosomal pathways present in AGF genomes and transcriptomes, such as pyruvate:ferredoxin oxidoreductase (PFO) (Marvin-Sikkema et al., 1994), a bifurcating hydrogenase, and a putative ATP synthase remains unclear. Previous flux balance analysis predictions suggest that observed hydrogen fluxes are much lower than would be expected if all transcribed hydrogenosomal pathways could carry flux unregulated (Wilken et al., 2021). To characterize the AGF hydrogenosome to the level required for predictable degradation and conversion of biomass in co-cultures, accurate measurement of steady-state fluxes during chemostat growth may be necessary. Existing pressure-based methods of AGF quantification, even in mono-culture with high temporal resolution (Wilken et al., 2020), are not suitable to quantify AGF concentration and fluxes in a chemostat; however, the method we present here is.

## **D. Conclusions**

Microbial communities can be leveraged to perform virtually infinite targeted chemical transformations, and new methods to track their performance in controlled systems are sorely needed. Individual species concentrations are critical metrics that enable many other analyses to understand communities but are challenging to obtain. Here, we have developed a method to obtain these metrics in biotechnologically promising co-cultures of biomass-degrading AGF and methanogens. While previous literature speculates that polysaccharide degradation and AGF growth is accelerated in co-culture with methanogens, methods have not been available to directly test this hypothesis.

With the method we have presented here for rapid quantification of non-rhizoidal AGF and methanogens in co-culture, we have demonstrated significant increases in AGF growth rate and xylan and glucose degradation rate in co-culture with a methanogen. Further, quantitative differences in AGF metabolic fluxes suggest a shift towards more energy-generating hydrogenosome flux in co-culture, however the highlighted uncertainties in the AGF hydrogenosome preclude definitive explanations for this shift. Detailed analyses of AGF that integrate flux measurements with transcriptomics and/or proteomics are likely necessary to characterize the AGF hydrogenosome and unlock their potential for predictable deployment in biotechnology applications; the method presented here is readily extendable to continuous or semi-batch systems for steady-state fluxomics to meet this need.

Importantly, this method may be modified and applied to co-cultures of other organisms with or without autofluorescence for detailed characterization of each organism's growth, flux, and other metrics that facilitate design and deployment of microbial communities with predictable, tunable functions.

#### **IV. Continuous culture of anaerobic fungi enables growth and metabolic flux tuning without the use of genetic tools**

Anaerobic gut fungi (AGF) have significant potential to valorize lignocellulosic biomass thanks to the diverse repertoire of carbohydrate-active enzymes (CAZymes) encoded in their genomes. However, AGF metabolism is poorly understood, and with no genetic tools available, there are few ways to manipulate growth rates and metabolic fluxes to enhance production of target products, e.g., cell biomass, CAZymes, or fermentation metabolites. Here, we present a method to achieve continuous culture of the non-rhizoidal AGF isolate, *Caecomyces churrovis*, to probe fungal metabolism and predictably alter AGF fluxes for the first time. A DIY, Arduino-based continuous flow bioreactor was constructed with online optical density control for operation under turbidostat or chemostat control in multiple reactor volume configurations (18 mL and 55 mL) under anaerobic conditions. Different reactor setpoints uncovered relationships between *C. churrovis* titer and growth rate (ranging from 0.04 to 0.20 h<sup>-1</sup>), metabolic flux, and production rates of acetate, formate, lactate, and ethanol, demonstrating how continuous culture can be used to tune *C. churrovis* flux predictably and reproducibly to bolster target products without the use of genetic tools to reroute flux. Lactate, in particular, showed a > 8x increase in flux, a > 38x increase in concentration, and a > 11x increase in molar production rate in continuous cultures in high vs. low titer setpoints. End-product spike-ins revealed that formate reproducibly altered AGF fluxes, but not growth, offering another simple tool to tailor fermentation products to application-specific profiles in both batch and continuous cultures. Continuous culture of *C. churrovis* represents a significant step towards understanding AGF metabolism and deploying AGF for industrial bioproduction and lignocellulose valorization.

## A. Introduction

Anaerobic gut fungi (AGF) from the phylum Neocallimastigomycota native to the rumen of large herbivores have significant biomass valorization potential thanks to the large, diverse repertoire of lignocellulose-degrading carbohydrate-active enzymes (CAZymes) encoded in their genomes (Hooker et al., 2019; Lillington et al., 2020; Solomon et al., 2016). Amplifying AGF CAZyme production to free sugars from plant biomass (Leggieri et al., 2022) is an attractive scheme to support more industrially tractable organisms in co-cultures (Jawed et al., 2019), or in multi-stage systems (Henske et al., 2018b). In addition to releasing sugars from biomass *via* CAZyme secretion, AGF produce several bioproduct precursors including lactate, acetate, formate, ethanol, succinate, and hydrogen. Depending on the target application, different profiles of these fermentation product are desirable. For example, hydrolysates rich in lactate or acetate are useful for medium-chain fatty acid (MCFA) production by chain elongating bacteria, (Scarborough et al., 2018), while high ethanol titers could be useful in bioenergy contexts (Toor et al., 2020).

Applications for lignocellulose-based bioproduction with AGF are numerous, and they all require a means to predictably, reproducibly control growth rates and metabolic fluxes. However, our control of AGF phenotypes is limited, and translation of AGF from academic research to industrial deployment for CAZyme production or crude biomass degradation and conversion to bioproduct precursors is precluded by several key challenges (Saye et al., 2021; Vinzelj et al., 2020). As non-model fungi, AGF physiology and primary metabolism are poorly understood. While the AGF are obligate anaerobes, many species are also filamentous and not easily amenable to standard process scaleup (Hanafy et al., 2020). Due to the uncertainty of the pathways in the AGF hydrogenosome (an anaerobic

mitochondrion-like organelle capable of ATP production) the existing genome-scale metabolic model is unable to accurately predict metabolic flux without experimental constraints (Wilken et al., 2021). Further, flux in AGF is highly regulated and changes over the course of batch growth (Leggieri et al., 2021a; Wilken et al., 2021). However, endpoint mixed acid metabolite profiles do not vary appreciably regardless of substrate, medium, or strain (Jin et al., 2011; Lankiewicz et al., 2023), and few strategies exist to elicit changes in AGF growth or metabolic flux to tailor fermentation profiles for different applications.

AGF genetic tools are nascent and have only recently shown success in transient transformation (Hooker et al., 2023), therefore alternative strategies to control AGF phenotypes are required. Co-culturing with methanogens has been shown to increase growth rate and alter flux and endpoint fermentation profiles in AGF, yielding more acetate and less lactate (Bauchop and Mountfort, 1981; Leggieri et al., 2021a; Li et al., 2017, 2019; Marvin-Sikkema et al., 1990). This is presumably due to removal of putatively inhibitory fermentation products, however co-cultures are notoriously difficult to maintain over long periods of time (Gilmore et al., 2019). Less precarious and more predictable means of tuning AGF flux are required if AGF are to be deployed for bioproduction at scale.

AGF scale-up strategies are virtually absent from literature, chiefly due to the rhizoidal morphology of most AGF strains (Hanafy et al., 2020). These long, root-like structures facilitate physical association with plant substrates, but complicate laboratory culture and analysis. On both solid and soluble substrates, rhizoidal AGF form dense mats which cannot be uniformly sampled to quantify growth. Even with stirring, well-mixed, nearly monodispersed suspensions of rhizoidal AGF are unachievable, precluding steady state continuous culture. However, the non-rhizoidal AGF, *Caecomyces churrovis*, forms

biofilm-like mats when cultured without agitation and well-mixed cell suspensions when cultured with stirring. We recently demonstrated that continuous stirring enhanced growth of *C. churrovis* and rhizoidal AGF on xylan and lignocellulose substrates, respectively (Leggieri et al., 2022). *C. churrovis* in stirred cultures also significantly upregulated expression of CAZymes with no changes in metabolic flux relative to biofilm cultures. These well-mixed cultures can be used to increase production rates of CAZymes and organic acids without altering the fermentation product profile and may be valuable for industrial scale-up to large systems in which mixing is critical. Further, well-mixed cultures *C. churrovis*, and other non-rhizoidal AGF can be uniformly sampled to quantify cells and metabolites and are amenable to continuous culture.

Continuous culture with turbidostat control can be used to tune AGF growth rates and fluxes to prescribed values to tailor AGF cultures to their many potential applications without genetic tools. Operators can take advantage of the regulated, dynamic flux profiles AGF demonstrate in batch culture to tune metabolic flux predictably. We hypothesize that by holding AGF cultures at constant, low titers, we will observe flux profiles that align with those from early batch growth, where cells experience a low-titer environment and mainly produce acetate and formate. Similar results are expected for high titers and late batch growth, where lactate production is expected to dominate. These various flux profiles enable biomass hydrolysates that are suited for different downstream applications, and the turbidostat control scheme presented herein allows operators to specify which products should be prioritized.

Continuous culture also enables controlled assessment of the effects of putatively inhibitory metabolic end-products on AGF growth and flux (Joblin and Naylor, 1993). Batch

AGF growth typically stops short of complete substrate utilization at higher loadings ( $> 10$  g/L) for reasons that are not yet known (Saye et al., 2021) and need to be uncovered to optimize biomass degradation. In batch cultures, all metabolite concentrations increase together with time, precluding isolation of the effects of each species on growth and metabolism. Metabolite spike-ins at the start of batch culture begin to address this point; however AGF growth rate varies significantly with titer over the course of batch growth (Wilken et al., 2020), and the resolution of batch flux measurements is limited. Continuous culture offers more controlled analysis of the effects of varying metabolite concentrations on AGF growth and flux.

Previous efforts have grown AGF with continuous liquid flow in systems with solid plant substrates using different liquid vs. solid retention times (Saye et al., 2021; Zhu et al., 1997, 1996). While these studies provide valuable insight into how continuous supernatant flow may enhance biomass degradation in AGF, these systems do not achieve steady state as fungi adhered to solids accumulate in the system over time. Further, motile zoospores (the immature phase of the AGF lifecycle) exist primarily in the liquid phase of AGF cultures, therefore differences in liquid vs. solid retention times may bias growth and flux unpredictably. Steady state culture is irreplaceable for detailed understanding of metabolism and improvements to the existing AGF genome-scale metabolic model (Wilken et al., 2021).

Here we present a simple, Arduino-based continuous flow bioreactor for continuous culture of anaerobic fungi in 18 mL or 55 mL configurations, assembled for under \$400. The bioreactor is equipped with an online optical density sensor that samples the live culture every 300 milliseconds and is used for turbidostat control to hold cultures at prescribed cell densities. Alternatively, a fixed dilution rate can be applied to the bioreactor for operation in

chemostat control. We demonstrate steady state culture over several days and greater than five residence times in both control modes and argue that turbidostat control is more reliable for AGF systems whose growth and substrate uptake rates may vary with setpoint titer in ways that are difficult to predict. With turbidostat control, the growth rate that corresponds to the setpoint titer does not need to be known *a priori*; the bioreactor identifies the required dilution rate (and therefore growth rate) and holds it constant. In chemostat control, if the growth and substrate uptake rates are not guessed exactly correctly for a given target titer, cells accumulate or wash out and fail to achieve steady state.

With this apparatus, faster AGF growth rates were achieved in continuous cultures compared to those previously reported for batch cultures, and they are based on direct optical density measurements rather than proxy measurements such as accumulated pressure (Theodorou et al., 1995) which are unreliable even at high resolution (Wilken et al., 2020) in systems with changing flux. By assessing steady state culture at several different prescribed titers, we uncover different flux profiles that maximize production of acetate, formate, and lactate. In batch culture, AGF flux profiles are difficult to resolve and only achievable for a few hours before regulation alters which metabolites are produced; with continuous culture we can produce each of these different fermentation product profiles indefinitely according to which products are desired.

Turbidostat control uncovers which setpoint titers can be used to maximize production rates in AGF systems depending on the target product (cells, CAZymes, lactate, formate, etc.). Because growth rate and flux both vary with titer, acetate and formate see their maximum production rate at low or intermediate titer, rather than high titer. Cell and CAZyme production rates are also maximized at intermediate titer. Lactate production, on

the other hand, is maximized at high titer. Further, metabolite spike-in experiments uncovered how formate concentration can be used to predictably control flux of formate, acetate, lactate, and ethanol in batch and continuous cultures.

## **B. Materials and Methods**

### **1. Culture of anaerobic gut fungi**

A fully defined, modified version of anaerobic Medium B (MB) (Leggieri et al., 2021a; Theodorou et al., 1995) was used for both routine culture and growth experiments; coenzyme M was not included in the formulation. The headspace of AGF cultures was 5% H<sub>2</sub>, 20% CO<sub>2</sub>, balance N<sub>2</sub>. All batch cultures were grown at 39 °C in 16 mL anaerobic Hungate tubes with 10 mL of anaerobic liquid medium. Continuous cultures grown using the bioreactor under turbidostat or chemostat control were grown in 75 mL anaerobic serum bottles or 20 mL scintillation vials at culture volumes of 55 mL and 18 mL, respectively. 55 mL cultures were stirred with one 25 mm cylindrical magnetic stir bar (Part No. SWN660, Globe Scientific, Mahwah, NJ); 18 mL cultures were stirred with one 13 mm cylindrical magnetic stir bar; stir bars were autoclaved in each culture bottle with the media. AGF were grown on glucose (anhydrous, Thermo Fisher Scientific, Waltham, MA) at final concentrations of either 5 g/L or 10 g/L. Continuous cultures contained chloramphenicol (Product No. C2255, TCI America, Portland, OR) at 100 µg/mL. In batch cultures, three biological replicate AGF cultures were grown for each condition alongside three blank media samples.

*Caecomycetes churrovis* and *Neocallimastix lanati* were both previously isolated from the feces of different sheep (Henske et al., 2017; Wilken et al., 2021). For routine culture, AGF were transferred to new media every day. All inoculums were 10% v/v. Growth of *N.*

*lanati* was monitored via pressure accumulation as described previously (Peng et al., 2018; Theodorou et al., 1995). Growth of *C. churrovis* was monitored via optical density (OD) at 450 nm, which has been calibrated with *C. churrovis* dry cell weight (Leggieri et al., 2021a). In continuous bioreactor trials, *C. churrovis* OD was monitored in real time using the online OD sensor which was calibrated with OD measurements on a plate reader (**Supplementary Figure S8**). Additionally, samples were taken from the reactor over the course of growth to validate the accuracy of the online OD sensor; optical density measurements reported herein were measured from these samples using a Tecan M1000 96-well plate reader.

## 2. High performance liquid chromatography (HPLC) analysis of metabolites

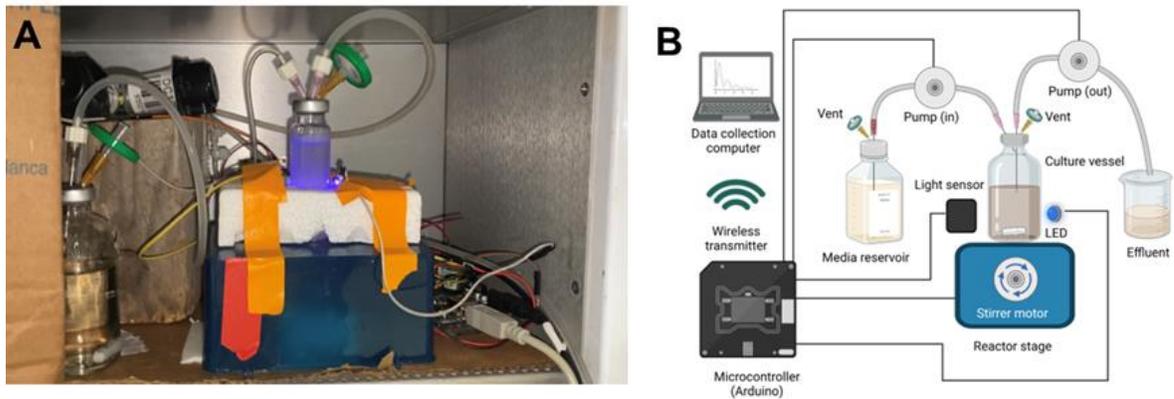
HPLC analysis was conducted as described previously (Leggieri et al., 2021a), with the exceptions that 60 °C was used instead of 50 °C and the refractive index detector (RID) was used to quantify all metabolites. Trace concentrations of fumarate impart a large signal on the variable wavelength detector at 210 nm which obstructs lactate and formate signals; the molar signal of fumarate is much smaller on the RID, therefore lactate and formate can be resolved cleanly. Standard curves were created for succinate, lactate, formate, acetate, and ethanol via linear regression;  $R^2$  values ranged from 0.9996 to 1.000.

Metabolic flux measurements for each metabolite were calculated as follows for continuous cultures in specified time windows: the average concentration of each metabolite (mM) in the bioreactor times the average volumetric flow rate ( $L h^{-1}$ ) divided by the average AGF mass (GDW) in the bioreactor (units  $mmol GDW^{-1} h^{-1}$ ). In batch systems, metabolic flux was calculated as follows: the difference in amount (mmol) of each metabolite in each culture divided by the elapsed time between the two measurements divided by the difference in AGF mass (difference in gram dry weight, dGDW) between the beginning and end of that

time window (units  $\text{mmol dGDW}^{-1} \text{h}^{-1}$ ). See **Supplementary Figure S9** for details regarding normalization of batch fluxes by the *change* in AGF mass rather than the *average* AGF mass. Where described, metabolic flux measurements are normalized by specific glucose uptake rates to enable direct comparison of fluxes between systems with different growth and/or glucose uptake rates.

### 3. Bioreactor design and operation

An image, diagram, and parts list of all reactor components with connections to/from the Arduino Mega microcontroller and part costs are provided in **Figure 20** and **Supplementary Table S4**. The Arduino IDE code varies depending on OD setpoint and reactor volume; representative scripts for turbidostat and chemostat control are provided in **Supplementary Scripts S1 and S2**, respectively.



**Figure 20.** Image (A) and diagram (B) of the 18 mL configuration of the Arduino-based bioreactor for continuous culture of anaerobic gut fungi. Culture vessel is held in place by a polystyrene mold to ensure uniform path length of OD sensor. Culture vessel sits atop a stage with a magnetic stirrer underneath, controlled by the Arduino. Media reservoir (left) and culture vessel are both vented to atmospheric pressure through  $0.22 \mu\text{m}$  filters. Reactor effluent drains to separate collection vessel (tubing runs out of frame at the bottom right). Arduino is powered by 9V DC, and motors for pumps and stirrer are powered by 12V DC. Data is wirelessly transmitted from the Arduino microcontroller to a computer outside of the incubator. (B) made with BioRender.com.

The bioreactor was run with both 18 mL and 55 mL volume configurations using different size vials as described above. All bioreactor runs were operated within an incubator at  $39 \text{ }^\circ\text{C}$  inside of an anaerobic chamber (5%  $\text{H}_2$ , 20%  $\text{CO}_2$ , balance  $\text{N}_2$ ). The incubator doors

were sealed to prevent any outside light from interfering with online OD measurements. Bioreactor culture vials and media reservoirs were constantly vented to the chamber atmosphere using a needle and 0.22  $\mu\text{m}$  filter to eliminate the effects of accumulated pressure on pump flow rates. A constant liquid level was maintained by fixing the outlet needle at the top of the liquid level in the culture vessel. All bioreactor runs were operated with turbidostat control except for the high-titer (OD setpoint 0.35) runs in the 55 mL configuration, which were operated at fixed dilution rates (chemostat control) due to loss of linearity in the OD signal at high cell concentrations in the 55 mL reactor. Each setpoint was run at least in duplicate.

Samples were taken from the bioreactor in two ways: timepoint samples in which the reactor effluent was immediately collected and analyzed, and combined effluent samples in which total reactor effluent in a given time window was collected and analyzed. Both sample types are reported and identified herein. Reactor flow rates were calculated by measuring the volume of effluent in a given timeframe (> 1 hour).

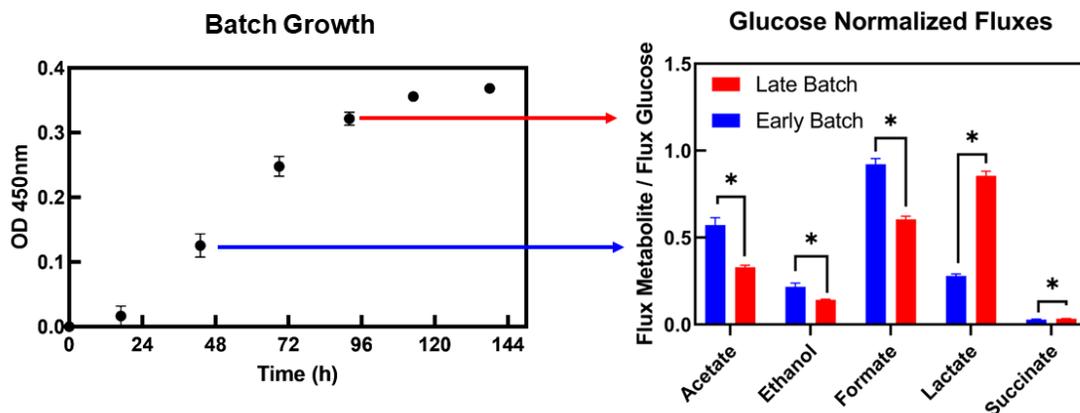
#### 4. Statistical analysis

All statistical analyses were conducted using the Prism 9.5.0 software (GraphPad, San Diego, CA). Prism 9.5.0 was used to (i) interpolate the concentrations of metabolites detected via HPLC using standard curves, and (ii) determine significant differences in growth rates and metabolite fluxes between growth conditions and timepoints via one-way ANOVA and multiple comparison t-tests. In all statistical tests,  $\alpha = 0.05$  was used.

### C. Results and Discussion

1. *C. churrovii* produces formate and acetate early in batch growth, lactate later in batch growth

As shown in **Figure 21**, *C. churrovis*, and many other AGF (Bauchop and Mountfort, 1981; Leggieri et al., 2021a; Srinivasan et al., 2001), produce more formate and acetate early in batch growth, and more lactate later in batch growth. We hypothesize that these different flux profiles are due to AGF sensing the culture microenvironment (titer) and that in continuous cultures held *via* turbidostat control at low titer (representative of early batch culture), *C. churrovis* will show greater formate and acetate flux and lesser lactate flux than continuous cultures held at high titer (representative of late batch culture). Operating continuous *C. churrovis* cultures at different, constant titers enables control and prescription of production rates of cells, CAZymes, and all primary metabolites.



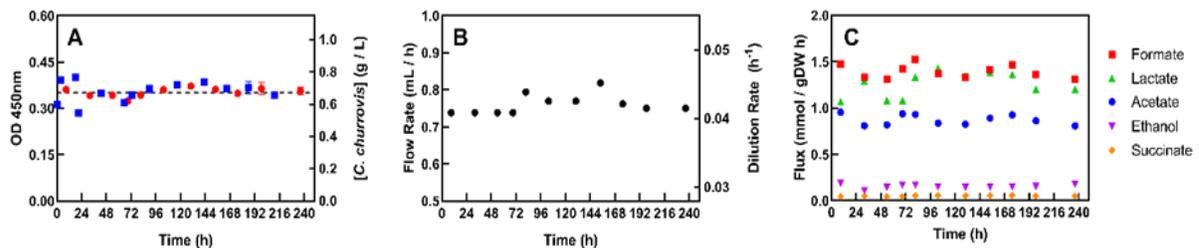
**Figure 21.** Fluxes change over the course of growth in *C. churrovis* batch culture. Acetate and formate fluxes are significantly greater in early batch vs. late batch (defined as shown in the growth curve in panel A), and lactate and succinate flux are significantly greater in late batch. This endpoint cell titer (OD 0.35 – 0.40) is the maximum that is typically achieved in *C. churrovis* batch culture regardless of initial substrate concentration.

2. Continuous, steady state growth and flux in *C. churrovis* is achievable with the DIY anaerobic bioreactor

To achieve steady state growth and flux to understand and control AGF metabolism, the bioreactor must be able to reach and maintain different setpoint titers. **Figure 22** demonstrates that the bioreactor achieves the specified OD setpoint (0.35 in the provided

example) in turbidostat control with a constant flow rate over 10 days and 10 residence times. Accordingly, metabolite concentrations, and therefore fluxes, were also constant over this time, indicating that *C. churrovis* achieved metabolic steady state. In **Figure 22**, the culture was held at high titer, near the maximum concentration of *C. churrovis* typically achieved in batch culture (see **Figure 21** for reference). At this high titer, lactate flux averaged 1.4 mmol/gDW h, which is comparable to late batch growth, and greater than early batch growth.

In addition to steady state continuous culture, the bioreactor presented herein produces batch OD-based growth curves for well-mixed *C. churrovis* cultures at far greater resolution than has been previously documented. These growth curves require no invasive sampling or destruction of the culture in any way and are more reliable than even high-resolution pressure-based methods. A representative high-resolution batch growth curve is shown in **Supplementary Figure S10**.



**Figure 22.** Constant OD (A), flow rate (B), and metabolite fluxes (C) in 18 mL reactor with high-titer turbidostat control at an OD setpoint of 0.35 (*C. churrovis* concentration 0.67 g/L). Blue squares in (A) represent OD measured with a single sample at that timepoint. Red circles in (A) represent OD measured using the combined reactor effluent over several hours. 240 hours corresponds to over 10 reactor residence times. Concentrations, flow rates, and fluxes vary depending on setpoint titer, but the steadiness shown here is representative of all reactor runs.

In turbidostat control, the OD is specified, and the flow rate required to maintain that setpoint OD is uncovered. When the flow rate is constant, metabolic steady state is implied, and that flow rate can be used to calculate the growth rate at that titer. In all cases, each bioreactor run showed constant OD, flow rate, and flux values over more than 5 residence

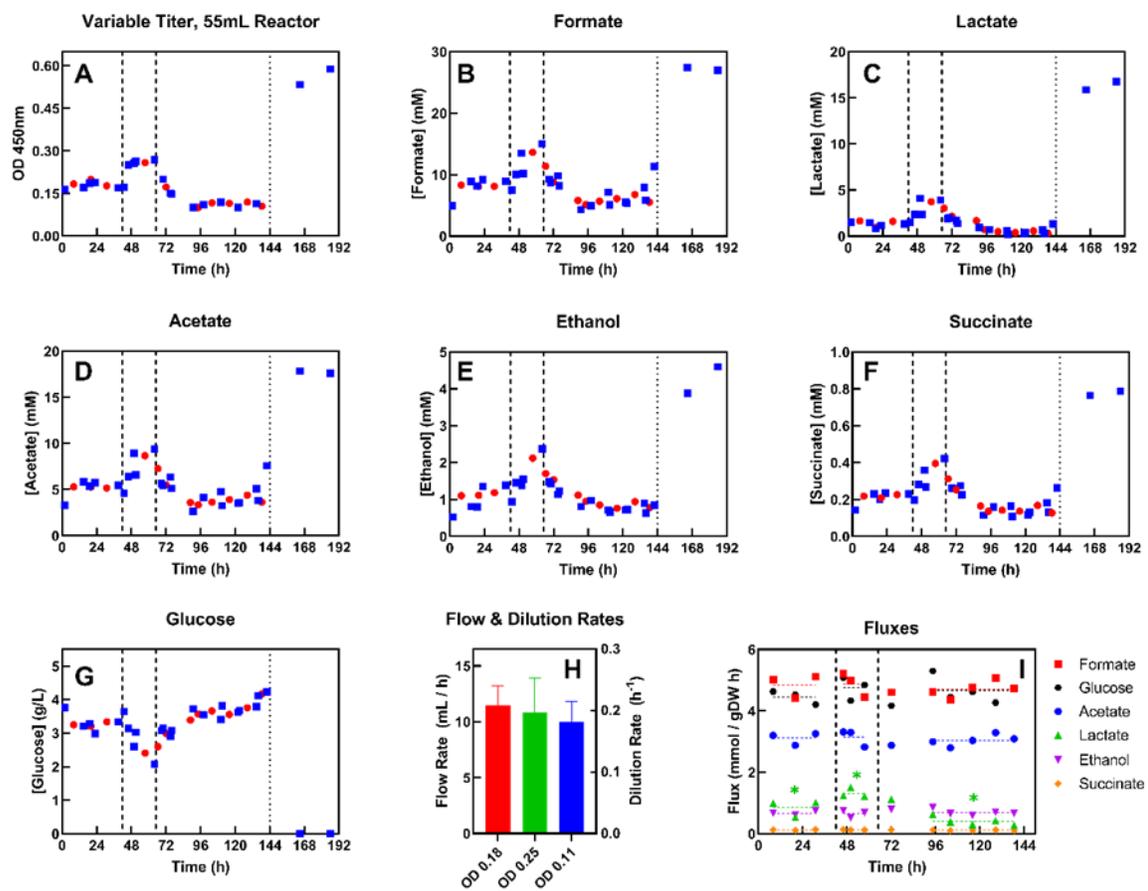
times, suggesting metabolic steady state. However, the specific values of cell concentration, growth rate, and metabolite fluxes change depending on the setpoint OD, which can be leveraged to predicably prioritize production of cells/CAZymes or different mixed acid products. This enables understanding and modeling of the available AGF phenotype space and elucidates design rules for how to culture AGF for specific applications.

Alternatively, chemostat control was employed to achieve high titer continuous culture in the larger 55 mL reactor configuration in which the light path length was too long to maintain a linear signal at high cell titers (**Supplementary Figure S8**). As shown in **Supplementary Figure S11**, metabolic steady state was achieved, and growth rates and fluxes were comparable between the 18mL and 55mL configurations operated at the same setpoint titer using turbidostat and chemostat control, respectively. Chemostat control was enabled by turbidostat control because a previous turbidostat control experiment elucidated the growth rate at that setpoint titer. Without that knowledge, one is forced to guess the growth rate at that high titer which is difficult to estimate from batch data and likely requires iteration to avoid cell accumulation or washout. Only the 55 mL 0.35 OD setpoint trials were run with chemostat control, all others reported herein were operated with turbidostat control.

As shown in **Figure 23**, the OD setpoint for the reactor was altered in real time to demonstrate how steady state AGF cultures respond to disturbances. Three different OD setpoints (OD 0.18, 0.25, and 0.11) were reached (A), with steady state growth and flux established at each of them. Small changes in setpoint titer effected no change in growth rate (H) or flux (I) of any metabolite except for lactate, which differed significantly between all three setpoints. Lactate flux appears to have the greatest sensitivity to titer of all AGF

metabolites. With turbidostat control, these different setpoints are easily reached and maintained without washout or accumulation, and culture viability is maintained.

After several days of continuous growth, flow was turned off, and *C. churrovis* was allowed to grow until all remaining substrate was depleted. As shown in **Figure 23 (A-G)**, after stopping flow at 144 h (dotted lines), the concentrations of *C. churrovis* and all metabolic end products reached high levels which are greater than those typically achieved in



**Figure 23.** Variable setpoint turbidostat control in the 55 mL reactor achieved three separate steady states in one reactor run (OD setpoints 0.18, 0.25, and 0.11, sequentially). Different setpoint regimes are marked off by dashed lines. Blue squares in (A-G) represent OD or [metabolite] measured with a single sample at that timepoint. Red circles in (A-G) represent measurements using the combined reactor effluent over several hours. Dotted line at 144h represents when flow wash shut off to allow consumption of remaining substrate and accumulation and cells and metabolites. Growth rates in (H) were not significantly different at these three OD setpoints. Fluxes (I) did not differ at between the three OD setpoints except for lactate, which was significantly differed at all three setpoints.

batch. The exact cause of this is not yet clear, however it is apparent that continuous culture may enable AGF phenotypes and titers that are not achievable in batch growth.

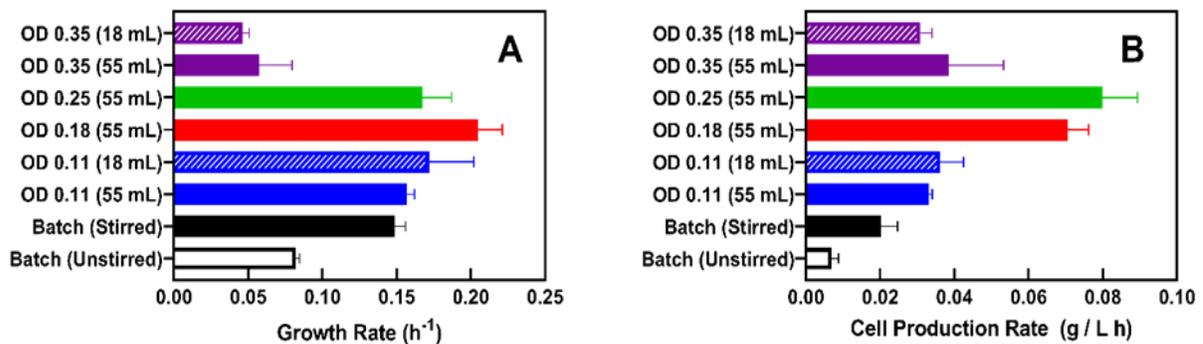
**Supplementary Figure S12** shows that growth stops in *C. churrovis* batch cultures near OD 0.40 even though almost half of the initial substrate remains. However, *C. churrovis* is able to grow in continuous culture near this OD setpoint, where all measured metabolites are at comparable concentration to the growth-arrested batch cultures. The unusual AGF lifecycle is poorly understood and is likely connected to what initiates and stops zoospore release and ultimately growth. While detailed characterization of the AGF lifecycle merits further investigation, pragmatically this lifecycle still allows *C. churrovis* to grow robustly in continuous culture at multiple different setpoint titers.

3. *C. churrovis* growth rate and cell production rate vary non-monotonically with titer

AGF growth rate changes the course of batch growth as the culture environment increases in titer. Therefore, the maximum achievable continuous AGF growth rate also likely depends on titer, and this relationship is difficult to quantify *a priori*. Turbidostat control was used to identify the maximum growth rate achievable at several different setpoint titers to find the global maximum AGF growth rate. **Figure 24A** shows that *C. churrovis* growth rate in continuous culture peaks at intermediate cell titer (OD 0.18) and decreases significantly at high titer (OD 0.35). This aligns with batch cultures in which *C. churrovis* growth slows substantially near this cell concentration even when excess sugar remains in culture (**Supplementary Figure S12**). Growth rates did not vary significantly between the 18 mL and 55 mL configurations at the same OD setpoints (0.35 and 0.11, dashed vs. solid bars of the same color,  $p > 0.05$ ). Growth rates did not vary significantly between OD setpoint 0.11, 0.18, or 0.25 ( $p > 0.05$ ), implying that the culture microenvironment is similar enough

at low and intermediate titer that growth is not significantly altered. OD setpoint 0.18 grew significantly faster ( $p = 0.0030$ ) than the fastest achievable stirred batch growth rate, showing promise for scale-up of *C. churrovii* growth for continuous production of CAZymes which are upregulated in stirred cultures (Leggieri et al., 2022).

When trying to maximize continuous production of *C. churrovii* biomass or CAZymes, the goal is not to identify the fastest-growing titer nor the titer of greatest *C. churrovii* concentration. Rather, the product of cell concentration and growth rate (equaling the cell production rate) should be maximized. As shown in **Figure 24**, the titer of maximum *C. churrovii* cell productivity, and likely CAZyme productivity as well (Leggieri et al., 2022), is OD 0.25, which is neither the highest cell titer nor the fastest-growing titer. Although cell titers are greater at OD 0.35 than OD 0.25, because growth is slower at high titer, OD 0.25 produces more *C. churrovii* biomass per unit time. Growth is slightly faster at OD 0.18 than OD 0.25 (although not significantly faster,  $p = 0.1698$ ), however OD 0.25 still produces more *C. churrovii* biomass per time because the titer of cells in the reactor is greater. OD setpoints 0.35 and 0.11 have vastly different growth rates and titers, but they

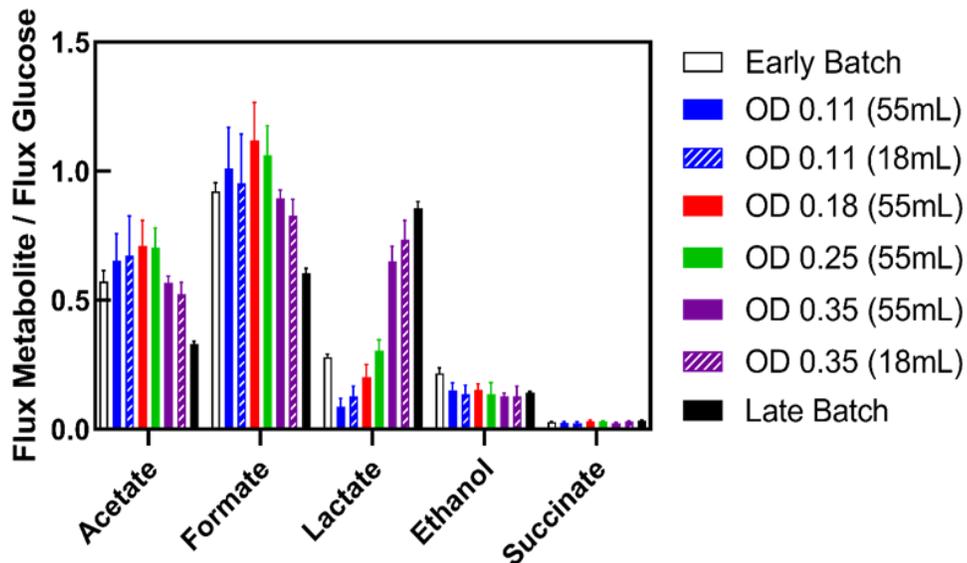


**Figure 24)** *C. churrovii* growth rate varies with titer non-monotonically (A). The maximum growth rate is at intermediate titer (OD 0.18); growth slows down significantly at higher titer (OD 0.35), whether chemostat or turbidostat control is used. OD setpoints 0.11, 0.18, and 0.25 all grew significantly faster than OD setpoint 0.35 ( $p < 0.05$ ). (B) The titer that yields the maximum cell production rate in *C. churrovii* is OD 0.25, which is neither the highest titer nor the fastest growing titer (B). All continuous cultures produced *C. churrovii* at higher rates than batch cultures.

produce *C. churrovis* biomass at the same rate ( $p > 0.05$ ). At all OD setpoints, the *C. churrovis* cell production rate was significantly greater in continuous culture than the fastest rate achievable in batch culture.

#### 4. Setpoint titer can be used to tune *C. churrovis* flux profiles and metabolite production rates

In many contexts, the target products from lignocellulose digestion are metabolites, which serve as precursors for bioproduction either downstream or in co-cultures (Patel and Shah, 2021; Zuroff and Curtis, 2012). Similar to growth rates, metabolic flux distributions, and therefore metabolite product profiles in *C. churrovis* continuous cultures change depending on the setpoint titer. As shown in **Figure 25**, low-to-intermediate titer (OD setpoints 0.11, 0.18, 0.25) fluxes resemble early-batch fluxes in which acetate and formate fluxes are dominant. At higher setpoint titer (OD 0.35), continuous flux resembles late-batch flux, in which acetate and formate fluxes decrease and lactate flux increases relative to early-

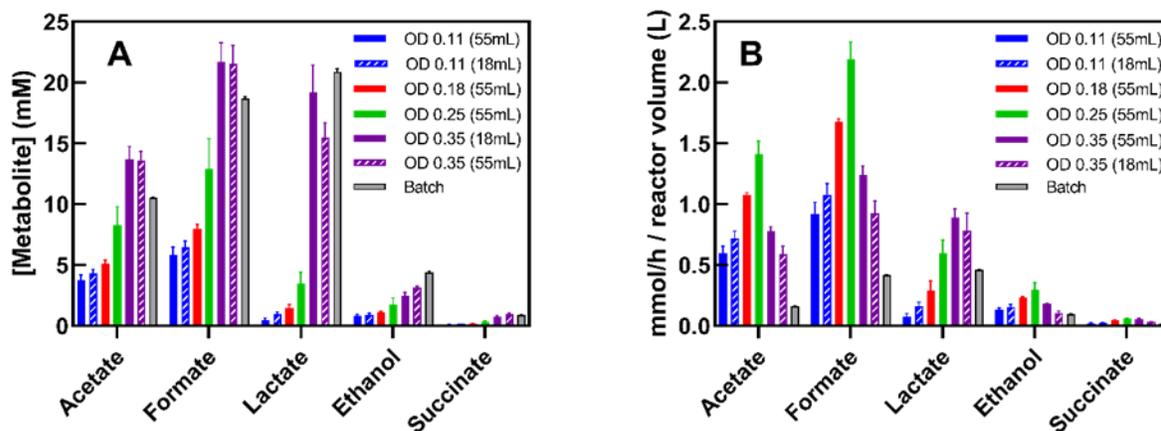


**Figure 25)** Metabolite fluxes normalized by glucose flux show greater lactate flux at high titer, and greater flux of acetate and formate at low titer. Similarly, formate and acetate fluxes are greater in early batch growth, and lactate flux is greater in late batch growth. No fluxes differed significantly between 18 mL and 55 mL configurations at the same OD setpoint (dashed vs. solid lines of the same color).

batch. All continuous flux distributions shown in **Figure 25** were maintained for several days and over 4 reactor residence times, representing the first direct tuning of AGF steady state metabolic flux. Depending on the target metabolite (acetate, formate, or lactate), different setpoint titers can be used to prioritize production of that product.

To maximize the continuous production rate of a given metabolite, rather than just the per-cell productivity represented by metabolic flux, the cell titer, growth rate, and flux must all be accounted for, as shown in **Figure 26**. **Figure 26** shows that, although all metabolite concentrations are greater at high titer (A), because growth rates are greater at intermediate titers, the maximum molar productivity (B) of most metabolites (acetate, formate, ethanol) is achieved at intermediate titer (OD setpoint 0.25). Lactate productivity, on the other hand, is maximized at high titer (OD setpoint 0.35). Even though growth is slower at this titer, the increased lactate flux at high vs. low titer leads to a greater molar production rate of lactate at high titer.

With this information, operators may easily tune the product profile of continuous AGF cultures using setpoint titers to maximize the production rate of the metabolite(s) of interest. For example, production of MCFAs *via* AGF co-cultures with chain-elongating bacteria is a burgeoning area of research. Depending on the specific chain-elongator and target MCFA product, operators may desire greater AGF fluxes of acetate or lactate (Stamatopoulou et al., 2020); the bioreactor presented here can predictably and reproducibly achieve those different desired flux profiles. As shown in **Figure 26**, all metabolites can be produced at significantly greater rates in continuous culture than in batch culture ( $p < 0.05$ ), indicating a promising strategy for AGF-based bioproduction.



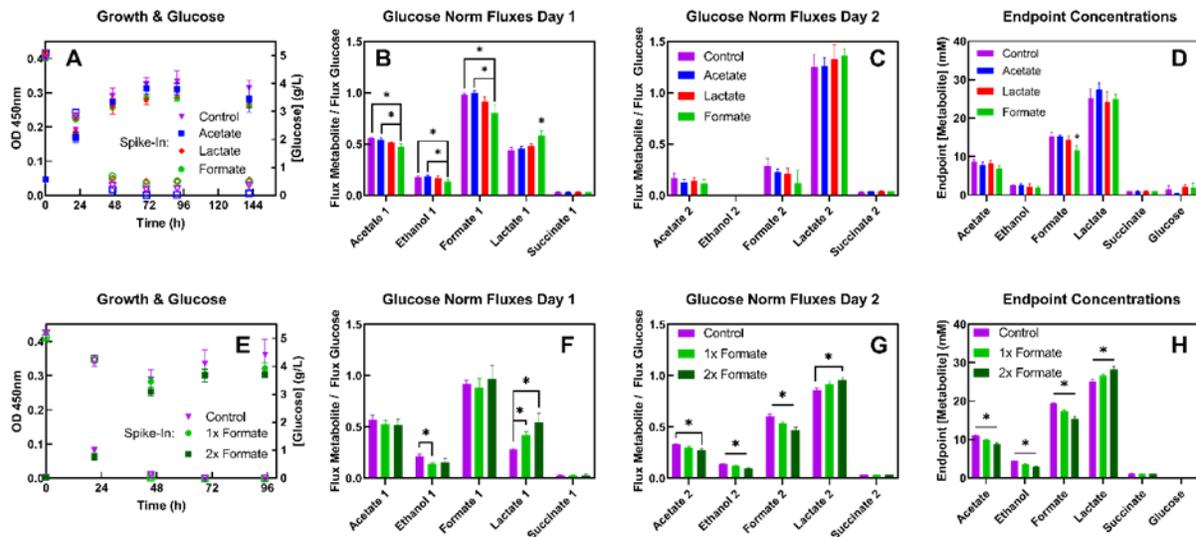
**Figure 26** All metabolite concentrations (A) are greater at high titer setpoints than low or intermediate titer setpoints. However, intermediate titer (OD 0.25) maximizes the molar productivity (B) of acetate, formate, and ethanol because growth rates and fluxes are greater at this titer. High titer (OD setpoint 0.35) continuous growth maximizes lactate productivity despite the slower growth rate because of the significantly greater lactate flux at high titer compared to intermediate and low titers. All metabolites can be produced at significantly faster rate in at least one continuous culture setpoint than in batch culture. Molar flow rates here are normalized by reactor volume.

5. Endproduct concentrations do not affect *C. churrovii* growth rate or yield, but formate concentration can be used to control flux of lactate, formate, acetate, or ethanol

**Figure 25** establishes a relationship between titer and flux in *C. churrovii*, however it is still not known exactly what *C. churrovii* senses and responds to to regulate its flux. To test whether acetate, lactate, or formate concentration directly affect *C. churrovii* growth rate, yield, or flux, each metabolite was spiked into separate batch cultures at the time of inoculation to yield concentrations typical of the end of batch culture (15 mM, 25 mM, 20 mM, respectively). AGF metabolism produces acetic acid, lactic acid, and formic acid which may decrease the culture pH as they accumulate, however most AGF media, including Medium B, is highly buffered and pH changes are small, but difficult to predict. Spiking in the conjugate bases, rather than the acids, removes any potential pH effects and assesses whether the anions themselves significantly alter AGF phenotypes *via* feedback on gene

expression or enzyme activity. As seen in **Figure 27A**, none of these metabolite spike-ins had any effect on *C. churrovis* growth rate, yield, or glucose consumption rate, contrary to previous results based on accumulated pressure or hydrogen as a growth proxy (Joblin and Naylor, 1993; Srinivasan et al., 2001).

**Figure 27B** shows that the only spike-in with any effect on flux was formate. As expected, batch lactate flux was significantly greater in all treatments later in growth (on day 2 of growth) than early in growth (on day 1). However, day 1 lactate flux was significantly greater with formate spike-in than in all other treatments, suggesting that higher formate concentrations may influence *C. churrovis* to produce more lactate even at lower cell titers. Formate spike-in also led to decreased formate, acetate, and ethanol flux on day 1, implying feedback inhibition of the acetate and ethanol production pathways which also produce formate. Accordingly, significantly less endpoint formate was produced in cultures that received the formate spike-in (**Figure 27D**). These fluxes align with previous observations in AGF-methanogen co-cultures in which removal of formate by methanogens increased acetate and ethanol flux and decreased lactate flux in AGF (Leggieri et al., 2021a).



**Figure 27)** Spike-ins of acetate, lactate, or formate during inoculation at typical batch endpoint concentrations had no effect on *C. churrovis* growth rate, yield, or glucose consumption (A); solid shapes represent culture OD, hollow shapes represent [glucose]. In panels B-D and F-H, different colors represent different spike-in treatments, while x-axis labels represent the fluxes or concentrations of the fermentation products produced. Spike-in of formate significantly increased day 1 lactate flux relative to all other conditions, and decreased day 1 formate, acetate, and ethanol fluxes relative to control and acetate spike-in groups (B). Spike-in of formate led to a significantly lower formate production over the course of batch growth relative to all other conditions (D). In follow-up experiment, spike-in of formate during inoculation at 1x and 2x typical batch endpoint concentration (25 mM and 50 mM, respectively) had no effect on growth rate, yield, or glucose consumption (E); solid shapes represent culture OD, hollow shapes represent [glucose]. Formate spike-in significantly increased day 1 lactate flux relative to control with an observable dose-response relationship; day 1 ethanol flux decreased with formate spike-in (F). Day 2 lactate flux increased with formate spike-in; day 2 ethanol acetate, ethanol, and formate flux decreased with formate spike-in, with an observable dose-response relationship for ethanol and formate (G). Formate spike-in yielded significantly greater lactate concentrations, and significantly lower acetate, ethanol, and formate concentrations with a dose-response relationship (H).

To test for a dose-response relationship between formate concentration and *C. churrovis* flux, formate was spiked into batch cultures at the time of inoculation at 1x typical end-batch concentration (20 mM), and 2x that concentration (40 mM). As shown in **Figure 27E**, formate concentration again had no effect on *C. churrovis* growth rate, yield, or glucose consumption rate. However, formate spike-ins again significantly increased lactate flux on day 1 of growth with an observable dose-response relationship, and decreased ethanol flux on day 1. On day 2, greater formate concentrations led to greater lactate flux and decreased acetate, ethanol, and formate fluxes, as expected. Accordingly, endpoint acetate, ethanol, and formate concentrations significantly decreased with increasing concentration of formate

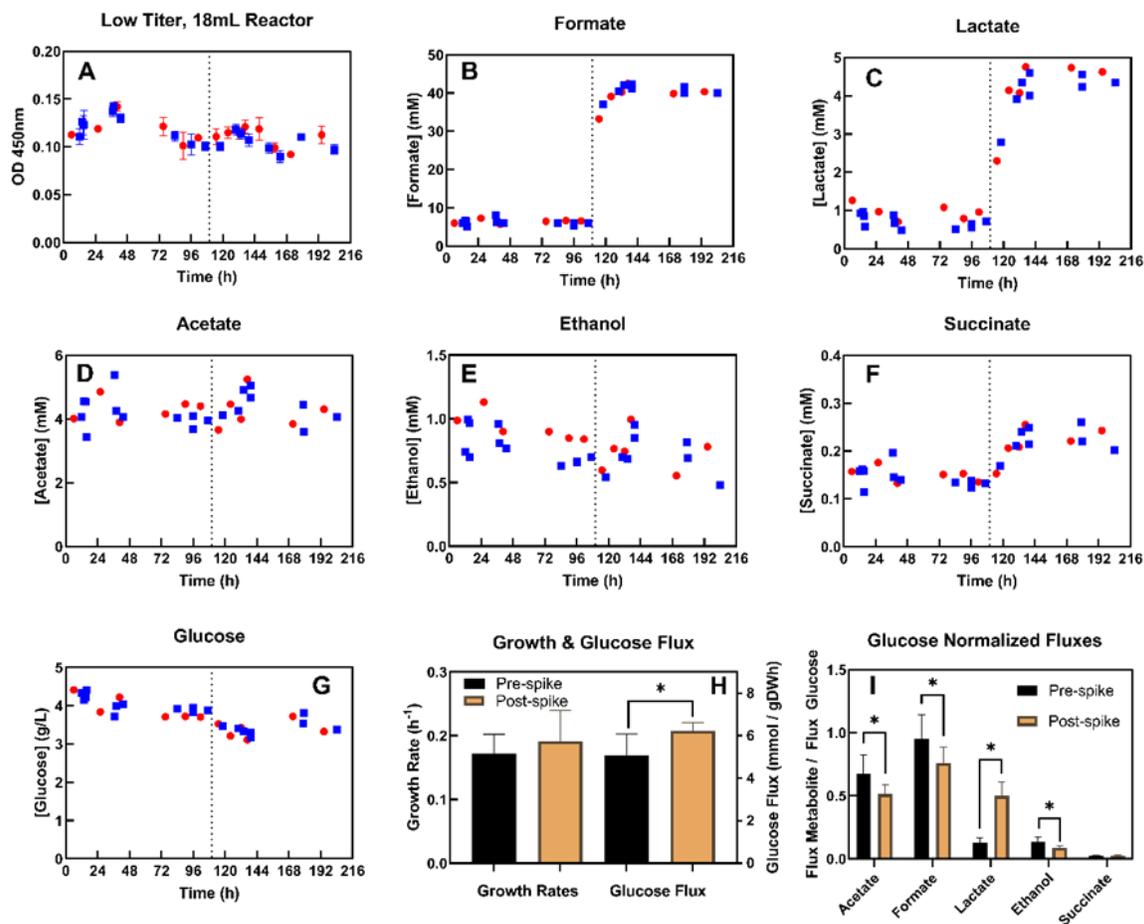
spike-in, and endpoint lactate concentration significantly increased with increasing formate concentration.

Decreased flux through formate-producing pathways at higher formate concentrations is logical. The only definite route to NAD(P)<sup>+</sup> regeneration without production of formate is lactate production, possibly explaining the observed flux adjustments. Other pathways could exist in the hydrogenosome, however this organelle remains largely uncharacterized in AGF (Boxma et al., 2004; Leggieri et al., 2021a; Marvin-Sikkema et al., 1994; Wilken et al., 2021).

As shown in **Supplementary Figure S13**, formate spike-ins yielded similar growth and flux results in a different, rhizoidal AGF, *Neocallimastix lanati*. Consistency of these results across AGF from different genera with different morphologies suggests that the effects of formate concentration on AGF growth and flux may extend across AGF more broadly. A mechanistic explanation for AGF flux redirection in response to increased formate concentration is not obvious from the data reported here, however it presents an interesting opportunity for future work. Formate could possibly affect flux *via* feedback on metabolic pathway expression or alteration of enzyme activity. Although AGF produce formate, given the heterogeneity of the AGF lifecycle, formate could also be assimilated (Mao et al., 2020); however, our data neither support nor refute that possibility. Future efforts could interrogate these hypotheses with transcriptomics/RT-qPCR, enzyme activity assays, and/or isotopic tracers.

6. Formate concentration can be used to tune flux of lactate, acetate, ethanol, and formate in *C. churrovis* continuous culture

To further verify that *C. churrovii* alters its flux profile in response to different formate concentrations, formate was spiked into a low-titer (OD setpoint 0.11) turbidostat-controlled continuous culture, and growth rates and metabolite fluxes were compared pre- and post-spike. First, the low-titer continuous culture was allowed to reach steady state for over 4 days (**Figure 28A**). At this low titer, low lactate fluxes were expected and observed. At 110 h of continuous culture, formate was spiked into the reactor and the media reservoir, effecting a step increase in formate concentration, as seen in **Figure 28B**. In response, lactate concentration (and flux) showed a step increase at this time (**Figure 28C**), supporting results from batch culture that formate concentration can be used to tune fluxes, most noticeably lactate. Specifically, formate supplementation can be used to suppress acetate and ethanol production and induce lactate production in low titer *C. churrovii* cultures which would otherwise produce little to no lactate. Acetate and ethanol concentrations did not decrease substantially after spiking in formate; however, because glucose flux increased (**Figure 28H**), the ratios of acetate and ethanol flux to glucose flux both decreased (**Figure 29I**) analogous to the batch cultures shown in **Figure 27**.



**Figure 28)** Spike-in of formate at 110h into an established, steady state culture of *C. churrovii* (OD setpoint 0.11) caused a step increase in lactate flux, and a decrease in flux of formate, acetate, and ethanol, analogous to batch culture. Growth rate was unaffected by the formate spike-in, however glucose flux significantly increased after the spike-in. Acetate, formate, and ethanol raw fluxes did not change significantly after the spike-in, however due to the increase in glucose flux, glucose-normalized acetate, formate, and ethanol fluxes decrease after the spike-in. Blue squares in (A-G) represent OD or [metabolite] measured with a single sample at that timepoint. Red circles in (A-G) represent measurements using the combined reactor effluent over several hours.

The lactate production rate achieved by artificially increasing formate concentration in otherwise low-titer (OD 0.11) *C. churrovii* continuous culture ( $0.75 \pm 0.06$  mmol / h L) is significantly greater than the low-titer lactate production rate with no formate supplementation ( $0.16 \pm 0.03$  mmol / h L), although it is lower than the high-titer (OD 0.35) lactate production rate ( $0.89 \pm 0.07$  mmol / h L). The low-titer continuous system was chosen to demonstrate the most dramatic increase in lactate flux after vs. before spiking in

formate. To maximize lactate continuous productivity in *C. churrovis* systems moving forward, formate could be spiked into faster-growing intermediate titer systems (OD setpoint 0.18, 0.25) to potentially produce more lactate using more cells. If genetic tools were available to AGF, lactate production could be increased by overexpressing the lactate dehydrogenase gene or possibly deleting pathways for production of succinate or ethanol. Without these tools, tuning formate concentrations and titer setpoints in continuous culture represent easily achievable strategies to predictably manipulate AGF metabolic flux.

#### **D. Conclusions**

We have presented a simple, cost-effective, Arduino-based continuous flow reactor for continuous, steady state culture of *C. churrovis* in defined media with both chemostat and turbidostat control modes. Turbidostat control uncovered relationships between titer and growth rate/metabolic flux profiles in *C. churrovis*, which may be used to prioritize production of different target products (cells, enzymes, or specific metabolites) at greater rates than have been achieved in batch culture. Formate concentration appears to have a significant effect on *C. churrovis*' metabolic flux profile, but not growth rate, and this effect can be leveraged to predictably tune product profiles in both batch and continuous culture for numerous bioproduction applications that require CAZymes, lactate, formate, and/or acetate. Continuous culture in *C. churrovis* offers a way to manipulate growth and flux without genetic tools and represents a significant step towards understanding AGF metabolism and deploying AGF for industrial bioproduction, enzyme synthesis and lignocellulose valorization.

## V. Conclusions, perspectives, outlook

The biotechnological potential of AGF remains untapped, but the work presented herein represents a small step toward determining whether AGF can feasibly be deployed for industrial lignocellulose valorization, either alone or in communities. The ability to degrade lignocellulose is evident in AGF genomes; however, we cannot capitalize on this genomic potential without knowing design rules to accentuate desired AGF phenotypes in laboratory culture, even at small scales.

Using a variety of modeling, bioinformatic, and wet lab approaches exposing AGF to myriad cultivation conditions (some resembling their natural environment and some more artificial), we have elucidated some of these design rules. We have developed and improved a fully defined culture medium for AGF to facilitate their growth and biochemical analysis. Although trivial for model microbes, we have a simple pipeline to quantify AGF *via* optical density, which uncovered the effects of co-culturing with methanogens on AGF growth and flux after decades of literature speculation. We can observe the AGF lifecycle in real time using live-cell microscopy, uncovering timescales and reproduction patterns that were previously unknown. We can culture AGF continuously at steady state and use the relationship between titer and AGF flux to tune metabolite production profiles predictably and reproducibly, thereby using chemical engineering principles to, essentially, metabolically engineer AGF without genetic tools.

Non-rhizoidal AGF, such as *Caecomyces churrovis*, are, in my opinion, the most useful AGF for laboratory characterization. Their ability to grow in both biofilms and well-mixed suspensions makes them amenable to uniform sampling in both mono- and co-cultures, continuous culture, and many of the routine analysis platforms enjoyed by model

microbes such as *E. coli* and yeast. These simple analyses were largely missing from AGF literature prior to this work. While bioprospecting for more AGF is an interesting effort in biology research, I believe it is most useful, in an engineering context, to devote efforts to comprehensively understanding the AGF we have already onboarded.

Of paramount importance in the near future is resequencing or at least better functional annotation of the genome of *C. churrovis*. *C. churrovis* is the model AGF thanks to its morphology and similarity to the rest of the Neocallimastigomycota phylum. However, its genome is of insufficient quality to enable comprehensive understanding of its metabolism, CAZyme productivity, lifecycle, or general behavior.

With more complete functional annotation of the *C. churrovis* genome, the existing *N. lanati* AGF GEM could be updated to represent *C. churrovis*. The current *C. churrovis* genome and observed biochemical data suggest that *C. churrovis* is very similar metabolically to AGF at large. *C. churrovis*, unlike *N. lanati* (upon which the current AGF GEM is based), can grow in continuous culture at steady state, which is a requirement for accurate FBA-based metabolic modeling. Further, *C. churrovis* offers an easier route to isolation and characterization of the AGF hydrogenosome, as its morphology is easier lysed than rhizoidal AGF mats and therefore does not require zoospore isolation.

The hydrogenosome is the single largest gap in our understanding of AGF metabolism. In its current state, the *N. lanati* GEM must be supplied with acetate flux (and therefore hydrogenosome flux) to accurately predict fluxes of any other metabolites. Lacking detailed hydrogenosome constraints, the GEM is currently descriptive, not predictive; this is still significantly beyond our understanding of AGF metabolism prior to the model, but it is

not sufficient to predict AGF behavior in communities or how AGF will respond to genetic edits once tools become available.

*C. churrovis* could provide the required hydrogenosome information to improve the GEM, but only after genome curation and annotation efforts enable mapping of proteomic results to genes of known function. Proteomic characterization combined with biochemical assays on isolated AGF hydrogenosomes could help quantify hydrogenosome components to logically constrain flux in FBA simulations. However, any proteomic results or peptide standards are based on sequences and annotations from the genome which are incomplete and not fully reliable. Most putative AGF hydrogenosome components are labeled as such based on sequence similarity to homologs in hydrogen-producing bacteria or protists, with some of these similarities as low as 40%. Further, observed flux ratios of formate: (acetate + ethanol) suggest that at least one pathway in the cytosol or hydrogenosome is missing, incomplete, or erroneously considered complete. Quantitative proteomics with peptide standards may uncover the third case, but not the first two. To discover new pathways or parts of pathways, we require a more complete annotation of relevant genes within the *C. churrovis* genome.

AGF are bizarre in their lifecycle, morphology, and behavior in laboratory culture, and as such warranted thorough fundamental research into best practices for isolation, cultivation, cryopreservation, and genomic and biochemical characterization prior to any scale-up consideration. The AGF hydrogenosome remains an open question; however, even without knowing specific internal fluxes, our work has uncovered how growth, gene expression, and external fluxes can be tuned predictably with mixing, spike-ins, and continuous culture. Now, it is time to sincerely explore scale-up and deployment options for

AGF to assess whether they can truly improve the current industrial lignocellulose valorization paradigm.

The DIY bioreactor presented in Chapter IV is a useful proof-of-concept that *C. churrovis* can grow continuously. However, industrial systems will likely operate using solid plant substrates, which demand an entirely different suite of practical engineering considerations, especially at scale. Fortunately, Matt Reilly at the University of York is well on his way to developing one such multi-phase continuous flow bioreactor for pilot-scale cultivation of AGF on grass substrates. As demonstrated in Chapter II, both rhizoidal and non-rhizoidal AGF can be grown with mixing on solid plant substrates. Rhizoidal AGF are, in fact, more amenable to mixing on solid substrates than soluble substrates, extending the design space of large-scale, well-mixed biomass-degrading AGF systems to both non-rhizoidal and rhizoidal AGF.

As for bioproduction, the mixed-acid fermentation pathways in AGF have industrial value, especially if their fluxes can be tuned and if they can reach high titers from crude biomass substrates. Whether product profiles are tuned *via* titer manipulation, metabolite spike-ins, or possibly genetic edits in the future, our newfound ability to predictably bias production of target compounds, including CAZymes or specific metabolites, makes AGF promising bioproduction candidates. Now, the work presented herein should be extended to assess whether the described strategies for predictable phenotype manipulation extend to systems grown on lignocellulose, potentially with comparisons between rhizoidal and non-rhizoidal strains to determine best practices for achieving industrially relevant production rates of CAZymes and bioproduct precursors.

Anaerobic communities have real potential to valorize recalcitrant lignocellulose and contribute to the imperative bio-based production of commodity, specialty, and energy products. Whether these communities can feasibly include AGF at industrial scales remains to be seen and should be considered an important area of research in the near future. As described in Chapter I, inclusion of prokaryotes or other co-culture partners with AGF significantly increases the degrees of freedom and decreases our ability to fully understand and reproduce the system over long periods of time.

As such, AGF should only be included in communities with specific goals in mind, not just to increase complexity for complexity's sake. AGF-methanogen co-cultures, for example, alter AGF flux and bring out different AGF fermentation product profiles that could be desirable in addition to producing methane instead of hydrogen; the desirability of this last feature is highly context-specific. An example of a well-designed AGF-bacterial system is a medium-chain fatty acid (MCFA)-producing community, in which AGF degrade lignocellulose and produce MCFA precursors that are converted to MCFAs (which have value in the makeup industry, for example) by chain-elongating bacteria. Another thoughtfully designed community is a co-culture of AGF with ZSC113, a strain of *E. coli* that consumes xylose but not glucose (Appendix D). In theory, in systems with both glucose and xylose available, such as lignocellulose hydrolysates, AGF should preferentially consume glucose, leaving xylose for the genetically tractable ZSC113 to consume and convert into target products. However, in *C. churrovis* – ZSC113 co-cultures, and to some extent in the MCFA-producing community, community membership is difficult to maintain, and death of one or more species results in loss of consortium function even though there is

no competition over substrates. This emphasizes the difficulty of designing microbial communities for even simple applications, especially when those communities contain AGF.

### **Reflections – Don't overcomplicate the problem**

While AGF are highly complex, non-model organisms, I have found that my most meaningful contributions to our understanding of them have been the simplest. We could not measure AGF concentrations because of their biofilm morphology, so I stirred them and identified a useful phenotype for both analysis and growth/CAZyme production. We wanted to quantify AGF in co-culture with methanogens; this is a system with two unknowns, so we needed two signals, and the easiest ones to measure were absorbance and fluorescence. We could not observe the AGF lifecycle in real time because their oxygen and temperature sensitivities complicated live-cell microscopy, so I grew them in spectroscopy cuvettes and stuck a light microscope inside an incubator. The bioreactor was sensitive to ambient light, temperature, and oxygen, so I scaled it down, operated it inside an incubator within an anaerobic chamber, and surrounded it with cardboard and duct tape.

Anaerobic life is all about being efficient with the resources that are available. We, as researchers, should consider this for inspiration. Be resourceful, think like an engineer, and ask direct, actionable questions. Good science comes out of that.

## VI. References

- Abdel-Hamid, A.M., Solbiati, J.O., Cann, I.K.O., 2013. Insights into Lignin Degradation and its Potential Industrial Applications, in: *Advances in Applied Microbiology*. Academic Press, pp. 1–28. <https://doi.org/10.1016/B978-0-12-407679-2.00001-6>
- Abu-Ali, G.S., Mehta, R.S., Lloyd-Price, J., Mallick, H., Branck, T., Ivey, K.L., Drew, D.A., DuLong, C., Rimm, E., Izard, J., Chan, A.T., Huttenhower, C., 2018. Metatranscriptome of human faecal microbial communities in a cohort of adult men. *Nat. Microbiol.* 3, 356–366. <https://doi.org/10.1038/s41564-017-0084-4>
- Akhmanova, A., Voncken, F.G.J., Hosea, K.M., Harhangi, H., Keltjens, J.T., Op Den Camp, H.J.M., Vogels, G.D., Hackstein, J.H.P., 1999. A hydrogenosome with pyruvate formate-lyase: anaerobic chytrid fungi use an alternative route for pyruvate catabolism. *Mol. Microbiol.* 32, 1103–1114. <https://doi.org/10.1046/J.1365-2958.1999.01434.X>
- Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K.L., Tyson, G.W., Nielsen, P.H., 2013. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat. Biotechnol.* 31, 533–8. <https://doi.org/10.1038/nbt.2579>
- Aller, S., Scott, A., Sarkar-Tyson, M., Soyer, O.S., 2018. Integrated human-virus metabolic stoichiometric modelling predicts host-based antiviral targets against Chikungunya, Dengue and Zika viruses. *J. R. Soc. Interface* 15. <https://doi.org/10.1098/rsif.2018.0125>
- Almagro Armenteros, J.J., Sønderby, C.K., Sønderby, S.K., Nielsen, H., Winther, O., 2017. DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics* 33, 3387–3395. <https://doi.org/10.1093/BIOINFORMATICS/BTX431>
- Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O.,

- Brunak, S., von Heijne, G., Nielsen, H., 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks 37, 420–423. <https://doi.org/10.1038/s41587-019-0036-z>
- Amarasinghe, S.L., Su, S., Dong, X., Zappia, L., Ritchie, M.E., Gouil, Q., 2020. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol.* 21. <https://doi.org/10.1186/s13059-020-1935-5>
- Antoniewicz, M.R., 2020. A guide to deciphering microbial interactions and metabolic fluxes in microbiome communities. *Curr. Opin. Biotechnol.* 64, 230–237. <https://doi.org/10.1016/j.copbio.2020.07.001>
- Babaei, P., Shoaie, S., Ji, B., Nielsen, J., 2018. Challenges in modeling the human gut microbiome. *Nat. Biotechnol.* 36, 682–686. <https://doi.org/10.1038/nbt.4213>
- Barlow, J.T., Bogatyrev, S.R., Ismagilov, R.F., 2020. A quantitative sequencing framework for absolute abundance measurements of mucosal and luminal microbial communities. *Nat. Commun.* 11, 1–13. <https://doi.org/10.1038/s41467-020-16224-6>
- Bauchop, T., Mountfort, D.O., 1981. Cellulose fermentation by a rumen anaerobic fungus in both the absence and the presence of rumen methanogens. *Appl. Environ. Microbiol.* 42, 1103–1110. <https://doi.org/10.1128/aem.42.6.1103-1110.1981>
- Bauer, E., Zimmermann, J., Baldini, F., Thiele, I., Kaleta, C., 2017. BacArena: Individual-based metabolic modeling of heterogeneous microbes in complex communities. *PLOS Comput. Biol.* 13, e1005544. <https://doi.org/10.1371/journal.pcbi.1005544>
- Beal, J., Farny, N.G., Haddock-Angelli, T., Selvarajah, V., Baldwin, G.S., Buckley-Taylor, R., Gershater, M., Kiga, D., Marken, J., Sanchania, V., Sison, A., Workman, C.T., 2020. Robust estimation of bacterial cell count from optical density. *Commun. Biol.* 3, 512.

<https://doi.org/10.1038/s42003-020-01127-5>

- Berry, D., Loy, A., 2018. Stable-Isotope Probing of Human and Animal Microbiome Function. *Trends Microbiol.* 26, 999–1007. <https://doi.org/10.1016/J.TIM.2018.06.004>
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Boxma, B., Voncken, F., Jannink, S., Van Alen, T., Akhmanova, A., Van Weelden, S.W.H., Van Hellemond, J.J., Ricard, G., Huynen, M., Tielens, A.G.M., Hackstein, J.H.P., 2004. The anaerobic chytridiomycete fungus *Piromyces* sp. E2 produces ethanol via pyruvate:formate lyase and an alcohol dehydrogenase E. *Mol. Microbiol.* 51, 1389–1399. <https://doi.org/10.1046/j.1365-2958.2003.03912.x>
- Boyarskiy, S., Tullman-Ercek, D., 2015. Getting pumped: Membrane efflux transporters for enhanced biomolecule production. *Curr. Opin. Chem. Biol.* 28, 15–19. <https://doi.org/10.1016/j.cbpa.2015.05.019>
- Brethauer, S., Shahab, R.L., Studer, M.H., 2020. Impacts of biofilms on the conversion of cellulose. *Appl. Microbiol. Biotechnol.* 104, 5201–5212. <https://doi.org/10.1007/s00253-020-10595-y>
- Brown, C.T., Olm, M.R., Thomas, B.C., Banfield, J.F., 2016. Measurement of bacterial replication rates in microbial communities. *Nat. Biotechnol.* 34, 1256–1263. <https://doi.org/10.1038/nbt.3704>
- Brown, J.L., Swift, C.L., Mondo, S.J., Seppala, S., Salamov, A., Singan, V., Henrissat, B., Drula, E., Henske, J.K., Lee, S., LaButti, K., He, G., Yan, M., Barry, K., Grigoriev, I. V., O'Malley, M.A., 2021. Co-cultivation of the anaerobic fungus *Caecomyces*

churrovis with *Methanobacterium bryantii* enhances transcription of carbohydrate binding modules, dockerins, and pyruvate formate lyases on specific substrates. *Biotechnol. Biofuels* 14, 1–16. <https://doi.org/10.1186/S13068-021-02083-W/FIGURES/4>

Cai, J., Tan, T., Chan, S.H.J., 2019. Bridging evolutionary game theory and metabolic models for predicting microbial metabolic interactions. *bioRxiv*. <https://doi.org/10.1101/623173>

Campanaro, S., Treu, L., Kougias, P.G., Zhu, X., Angelidaki, I., 2018. Taxonomy of anaerobic digestion microbiome reveals biases associated with the applied high throughput sequencing strategies. *Sci. Rep.* 8, 1926. <https://doi.org/10.1038/s41598-018-20414-0>

Cao, X., Hamilton, J.J., Venturelli, O.S., 2019. Understanding and Engineering Distributed Biochemical Pathways in Microbial Communities. *Biochemistry* 58, 94–107. <https://doi.org/10.1021/acs.biochem.8b01006>

Chahal, D.S., 1985. Solid-state fermentation with *Trichoderma reesei* for cellulase production. *Appl. Environ. Microbiol.* 49, 205–210. <https://doi.org/10.1128/aem.49.1.205-210.1985>

Chan, S.H.J., Simons, M.N., Maranas, C.D., 2017. SteadyCom: Predicting microbial abundances while ensuring community stability. *PLOS Comput. Biol.* 13, e1005539. <https://doi.org/10.1371/journal.pcbi.1005539>

Chandrasekar, P.H., Manavathu, E.K., 2008. Do *Aspergillus* species produce biofilm? *Future Microbiol.* 3, 19–21. <https://doi.org/10.2217/17460913.3.1.19>

Chen, J., Gomez, J.A., Höffner, K., Phalak, P., Barton, P.I., Henson, M.A., 2016.

Spatiotemporal modeling of microbial metabolism. *BMC Syst. Biol.* 10, 21.

<https://doi.org/10.1186/s12918-016-0259-2>

Chinappi, M., Cecconi, F., 2018. Protein sequencing via nanopore based devices: a nanofluidics perspective. *J. Phys. Condens. Matter* 30, 204002.

<https://doi.org/10.1088/1361-648X/aababe>

Contijoch, E.J., Britton, G.J., Yang, C., Mogno, I., Li, Z., Ng, R., Llewellyn, S.R., Hira, S., Johnson, C., Rabinowitz, K.M., Barkan, R., Dotan, I., Hirten, R.P., Fu, S.C., Luo, Y., Yang, N., Luong, T., Labrias, P.R., Lira, S., Peter, I., Grinspan, A., Clemente, J.C., Kosoy, R., Kim-Schulze, S., Qin, X., Castillo, A., Hurley, A., Atreja, A., Rogers, J., Fasihuddin, F., Saliyaj, M., Nolan, A., Reyes-Mercedes, P., Rodriguez, C., Aly, S., Santa-Cruz, K., Peters, L., Suárez-Fariñas, M., Huang, R., Hao, K., Zhu, J., Zhang, B., Losic, B., Irizar, H., Song, W.M., Di Narzo, A., Wang, W., Cohen, B.L., DiMaio, C., Greenwald, D., Itzkowitz, S., Lucas, A., Marion, J., Maser, E., Ungaro, R., Naymagon, S., Novak, J., Shah, B., Ullman, T., Rubin, P., George, J., Legnani, P., Telesco, S.E., Friedman, J.R., Brodmerkel, C., Plevy, S., Cho, J.H., Colombel, J.F., Schadt, E.E., Arghmann, C., Dubinsky, M., Kasarskis, A., Sands, B., Faith, J.J., 2019. Gut microbiota density influences host physiology and is shaped by host and microbial factors. *Elife* 8.

<https://doi.org/10.7554/eLife.40553>

Coyotzi, S., Pratscher, J., Murrell, J.C., Neufeld, J.D., 2016. Targeted metagenomics of active microbial populations with stable-isotope probing. *Curr. Opin. Biotechnol.* 41, 1–8. <https://doi.org/10.1016/J.COPBIO.2016.02.017>

Dam, P., Fonseca, L.L., Konstantinidis, K.T., Voit, E.O., 2016. Dynamic models of the complex microbial metapopulation of lake mendota. *npj Syst. Biol. Appl.* 2, 16007.

<https://doi.org/10.1038/npjbsba.2016.7>

Doddema, H.J., Vogels, G.D., 1978. Improved identification of methanogenic bacteria by fluorescence microscopy. *Appl. Environ. Microbiol.* 36, 752–754.

Dollhofer, V., Callaghan, T.M., Dorn-In, S., Bauer, J., Lebuhn, M., 2016. Development of three specific PCR-based tools to determine quantity, cellulolytic transcriptional activity and phylogeny of anaerobic fungi. *J. Microbiol. Methods* 127, 28–40.

<https://doi.org/10.1016/J.MIMET.2016.05.017>

Doloman, A., Varghese, H., Miller, C.D., Flann, N.S., 2017. Modeling de novo granulation of anaerobic sludge. *BMC Syst. Biol.* 11. <https://doi.org/10.1186/s12918-017-0443-z>

Dumitrache, A., Klingeman, D.M., Natzke, J., Rodriguez Jr, M., Giannone, R.J., Hettich, R.L., Davison, B.H., Brown, S.D., 2017. Specialized activities and expression differences for *Clostridium thermocellum* biofilm and planktonic cells. *Sci. Rep.* 7, 43583. <https://doi.org/10.1038/srep43583>

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST.

*Bioinformatics* 26, 2460–2461. <https://doi.org/10.1093/BIOINFORMATICS/BTQ461>

Egert, M., Weis, S., Schnell, S., 2018. RNA-based stable isotope probing (RNA-SIP) to unravel intestinal host-microbe interactions. *Methods* 149, 25–30.

<https://doi.org/10.1016/J.YMETH.2018.05.022>

Eiteman, M.A., Lee, S.A., Altman, E., 2008. A co-fermentation strategy to consume sugar mixtures effectively. *J. Biol. Eng.* 2, 1–8. <https://doi.org/10.1186/1754-1611-2-3/FIGURES/7>

Emwas, A.H.M., 2015. The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. *Methods Mol. Biol.* 1277,

161–193. [https://doi.org/10.1007/978-1-4939-2377-9\\_13](https://doi.org/10.1007/978-1-4939-2377-9_13)

Fang, N., Akinci-Tolun, R., 2016. Depletion of Ribosomal RNA Sequences from Single-Cell RNA-Sequencing Library. *Curr. Protoc. Mol. Biol.* 115, 7.27.1-7.27.20.

<https://doi.org/10.1002/cpmb.11>

Fontana, A., Campanaro, S., Treu, L., Kougias, P.G., Cappa, F., Morelli, L., Angelidaki, I., 2018. Performance and genome-centric metagenomics of thermophilic single and two-stage anaerobic digesters treating cheese wastes. *Water Res.* 134, 181–191.

<https://doi.org/10.1016/J.WATRES.2018.02.001>

Fortunato, C.S., Huber, J.A., 2016. Coupled RNA-SIP and metatranscriptomics of active chemolithoautotrophic communities at a deep-sea hydrothermal vent. *ISME J.* 10, 1925–1938.

<https://doi.org/10.1038/ismej.2015.258>

Frey, E., 2010. Evolutionary game theory: Theoretical concepts and applications to microbial communities. *Physica A* 389, 4265–4298.

<https://doi.org/10.1016/j.physa.2010.02.047>

Gamarra, N.N., Villena, G.K., Gutiérrez-Correa, M., 2010. Cellulase production by *Aspergillus niger* in biofilm, solid-state, and submerged fermentations. *Appl. Microbiol. Biotechnol.* 87, 545–551.

<https://doi.org/10.1007/s00253-010-2540-4>

Gebreselassie, N.A., Antoniewicz, M.R., 2015. <sup>13</sup>C-metabolic flux analysis of co-cultures: A novel approach. *Metab. Eng.* 31, 132–139.

<https://doi.org/10.1016/j.ymben.2015.07.005>

Ghosh, A., Nilmeier, J., Weaver, D., Adams, P.D., Keasling, J.D., Mukhopadhyay, A.,

Petzold, C.J., Martín, H.G., 2014. A Peptide-Based Method for <sup>13</sup>C Metabolic Flux

Analysis in Microbial Communities. *PLoS Comput. Biol.* 10, e1003827.

<https://doi.org/10.1371/journal.pcbi.1003827>

- Gilmore, S.P., Lankiewicz, T.S., Wilken, S.E., Brown, J.L., Sexton, J.A., Henske, J.K., Theodorou, M.K., Valentine, D.L., O'Malley, M.A., 2019. Top-Down Enrichment Guides in Formation of Synthetic Microbial Consortia for Biomass Degradation. *ACS Synth. Biol.* 8, 2174–2185. <https://doi.org/10.1021/acssynbio.9b00271>
- Gómez-Godínez, L.J., Fernandez-Valverde, S.L., Martinez Romero, J.C., Martínez-Romero, E., 2019. Metatranscriptomics and nitrogen fixation from the rhizoplane of maize plantlets inoculated with a group of PGPRs. *Syst. Appl. Microbiol.* 42, 517–525. <https://doi.org/10.1016/j.syapm.2019.05.003>
- Gorochowski, T.E., Matyjaszkiewicz, A., Todd, T., Oak, N., Kowalska, K., Reid, S., Tsaneva-Atanasova, K.T., Savery, N.J., Grierson, C.S., di Bernardo, M., 2012. BSim: An Agent-Based Tool for Modeling Bacterial Populations in Systems and Synthetic Biology. *PLoS One* 7, e42790. <https://doi.org/10.1371/journal.pone.0042790>
- Gottstein, W., Olivier, B.G., Bruggeman, F.J., Teusink, B., 2016. Constraint-based stoichiometric modelling from single organisms to microbial communities. *J. R. Soc. Interface* 13. <https://doi.org/10.1098/rsif.2016.0627>
- Granzier, H.L., Labeit, S., 2004. The Giant Protein Titin: A Major Player in Myocardial Mechanics, Signaling, and Disease. *Circ. Res.* 94, 284–295. <https://doi.org/10.1161/01.RES.0000117769.88862.F8>
- Greening, C., Ahmed, F.H., Mohamed, A.E., Lee, B.M., Pandey, G., Warden, A.C., Scott, C., Oakeshott, J.G., Taylor, M.C., Jackson, C.J., 2016. Physiology, Biochemistry, and Applications of F<sub>420</sub>- and F<sub>o</sub>-Dependent Redox Reactions. *Microbiol. Mol. Biol. Rev.* 80, 451–493. <https://doi.org/10.1128/mnbr.00070-15>
- Grigoriev, I. V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otilar, R., Riley, R., Salamov,

- A., Zhao, X., Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I., Shabalov, I., 2014. MycoCosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Res.* 42. <https://doi.org/10.1093/nar/gkt1183>
- Gruninger, R.J., Puniya, A.K., Callaghan, T.M., Edwards, J.E., Youssef, N., Dagar, S.S., Fliegerova, K., Griffith, G.W., Forster, R., Tsang, A., Mcallister, T., Elshahed, M.S., 2014. Anaerobic fungi (phylum Neocallimastigomycota): Advances in understanding their taxonomy, life cycle, ecology, role and biotechnological potential. *FEMS Microbiol. Ecol.* 90, 1–17. <https://doi.org/10.1111/1574-6941.12383>
- Gu, C., Kim, G.B., Kim, W.J., Kim, H.U., Lee, S.Y., 2019. Current status and applications of genome-scale metabolic models. *Genome Biol.* 20, 1–18. <https://doi.org/10.1186/s13059-019-1730-3>
- Gudmundsson, S., Thiele, I., 2010. Computationally efficient flux variability analysis. *BMC Bioinformatics* 11, 489–489. <https://doi.org/10.1186/1471-2105-11-489>
- Gutiérrez-Correa, M., Ludeña, Y., Ramage, G., Villena, G.K., 2012. Recent advances on filamentous fungal biofilms for industrial uses. *Appl. Biochem. Biotechnol.* 167, 1235–1253. <https://doi.org/10.1007/s12010-012-9555-5>
- Gutiérrez, N., Garrido, D., 2019. Species Deletions from Microbiome Consortia Reveal Key Metabolic Interactions between Gut Microbes. *mSystems* 4, e00185-19. <https://doi.org/10.1128/msystems.00185-19>
- Hackstein, J.H.P., Baker, S.E., van Hellemond, J.J., Tielens, A.G.M., 2019. Hydrogenosomes of Anaerobic Fungi: An Alternative Way to Adapt to Anaerobic Environments. Springer, Cham, pp. 159–175. [https://doi.org/10.1007/978-3-030-17941-0\\_7](https://doi.org/10.1007/978-3-030-17941-0_7)
- Hadadi, N., Pandey, V., Chiappino-Pepe, A., Morales, M., Gallart-Ayala, H., Mehl, F.,

- Ivanisevic, J., Sentchilo, V., Meer, J.R. va. der, 2020. Mechanistic insights into bacterial metabolic reprogramming from omics-integrated genome-scale models. *npj Syst. Biol. Appl.* 6, 1–11. <https://doi.org/10.1038/s41540-019-0121-4>
- Hanafy, R.A., Lanjekar, V.B., Dhakephalkar, P.K., Callaghan, T.M., Dagar, S.S., Griffith, G.W., Elshahed, M.S., Youssef, N.H., 2020. Seven new Neocallimastigomycota genera from wild, zoo-housed, and domesticated herbivores greatly expand the taxonomic diversity of the phylum. *Mycologia* 112, 1212–1239. <https://doi.org/10.1080/00275514.2019.1696619>
- Harcombe, W.R., Riehl, W.J., Dukovski, I., Granger, B.R., Betts, A., Lang, A.H., Bonilla, G., Kar, A., Leiby, N., Mehta, P., Marx, C.J., Segrè, D., 2014. Metabolic Resource Allocation in Individual Microbes Determines Ecosystem Interactions and Spatial Dynamics. *Cell Rep.* 7, 1104–1115. <https://doi.org/10.1016/j.celrep.2014.03.070>
- Hartmann, A., Vila-Santa, A., Kallscheuer, N., Vogt, M., Julien-Laferrrière, A., Sagot, M.F., Marienhagen, J., Vinga, S., 2017. OptPipe - a pipeline for optimizing metabolic engineering targets. *BMC Syst. Biol.* 11, 143. <https://doi.org/10.1186/s12918-017-0515-0>
- He, B., Jin, S., Cao, J., Mi, L., Wang, J., 2019. Metatranscriptomics of the Hu sheep rumen microbiome reveals novel cellulases. *Biotechnol. Biofuels* 12, 153. <https://doi.org/10.1186/s13068-019-1498-4>
- Heine-Dobbernack, E., Schoberth, S.M., Sahn, H., 1988. Relationship of Intracellular Coenzyme F420 Content to Growth and Metabolic Activity of *Methanobacterium bryantii* and *Methanosarcina barkeri*. *Appl. Environ. Microbiol.* 54, 454–459. <https://doi.org/10.1128/aem.54.2.454-459.1988>

- Henske, J.K., Gilmore, S.P., Haitjema, C.H., Solomon, K. V., O'Malley, M.A., 2018a. Biomass-degrading enzymes are catabolite repressed in anaerobic gut fungi. *AIChE J.* 64, 4263–4270. <https://doi.org/10.1002/aic.16395>
- Henske, J.K., Gilmore, S.P., Knop, D., Cunningham, F.J., Sexton, J.A., Smallwood, C.R., Shutthanandan, V., Evans, J.E., Theodorou, M.K., O'Malley, M.A., 2017. Transcriptomic characterization of *Caecomyces churrovis*: A novel, non-rhizoid-forming lignocellulolytic anaerobic fungus. *Biotechnol. Biofuels* 10, 305. <https://doi.org/10.1186/s13068-017-0997-4>
- Henske, J.K., Wilken, S.E., Solomon, K. V., Smallwood, C.R., Shutthanandan, V., Evans, J.E., Theodorou, M.K., O'Malley, M.A., 2018b. Metabolic characterization of anaerobic fungi provides a path forward for bioprocessing of crude lignocellulose. *Biotechnol. Bioeng.* 115, 874–884. <https://doi.org/10.1002/bit.26515>
- Henson, M., Phalak, P., 2017. Byproduct Cross Feeding and Community Stability in an In Silico Biofilm Model of the Gut Microbiome. *Processes* 5, 13. <https://doi.org/10.3390/pr5010013>
- Hoek, M.J.A. va., Merks, R.M.H., 2017. Emergence of microbial diversity due to cross-feeding interactions in a spatial model of gut microbial metabolism. *BMC Syst. Biol.* 11, 56. <https://doi.org/10.1186/s12918-017-0430-4>
- Hooker, C.A., Hanafy, R., Hillman, E.T., Muñ Oz Briones, J., Solomon, K. V, 2023. A Genetic Engineering Toolbox for the Lignocellulolytic Anaerobic Gut Fungus *Neocallimastix frontalis*. *ACS Synth. Biol.* <https://doi.org/10.1021/acssynbio.2c00502>
- Hooker, C.A., Lee, K.Z., Solomon, K. V, 2019. Leveraging anaerobic fungi for biotechnology. *Curr. Opin. Biotechnol.* 59, 103–110.

<https://doi.org/10.1016/J.COPBIO.2019.03.013>

Hungate, R.E., 1969. *The Rumen and Its Microbes*, 1st ed, Academic Press. Academic Press, New York-London. <https://doi.org/10.1016/C2013-0-12555-X>

Ibberson, C.B., Stacy, A., Fleming, D., Dees, J.L., Rumbaugh, K., Gilmore, M.S., Whiteley, M., 2017. Co-infecting microorganisms dramatically alter pathogen gene essentiality during polymicrobial infection. *Nat. Microbiol.* 2, 17079.

<https://doi.org/10.1038/nmicrobiol.2017.79>

Jawed, K., Yazdani, S.S., Koffas, M.A., 2019. Advances in the development and application of microbial consortia for metabolic engineering. *Metab. Eng. Commun.*

<https://doi.org/10.1016/j.mec.2019.e00095>

Jayathilake, P.G., Gupta, P., Li, B., Madsen, C., Oyebamiji, O., González-Cabaleiro, R., Rushton, S., Bridgens, B., Swailes, D., Allen, B., McGough, A.S., Zuliani, P., Ofiteru, I.D., Wilkinson, D., Chen, J., Curtis, T., 2017. A mechanistic Individual-based Model of microbial communities. *PLoS One* 12. <https://doi.org/10.1371/journal.pone.0181965>

Jeffryes, J.G., Colastani, R.L., Elbadawi-Sidhu, M., Kind, T., Niehaus, T.D., Broadbelt, L.J., Hanson, A.D., Fiehn, O., Tyo, K.E.J., Henry, C.S., 2015. MINEs: Open access databases of computationally predicted enzyme promiscuity products for untargeted metabolomics. *J. Cheminform.* 7, 44. <https://doi.org/10.1186/s13321-015-0087-1>

Jin, W., Cheng, Y.-F., Mao, S.-Y., Zhu, W.-Y., 2011. Isolation of natural cultures of anaerobic fungi and indigenously associated methanogens from herbivores and their bioconversion of lignocellulosic materials to methane.

<https://doi.org/10.1016/j.biortech.2011.06.026>

Joblin, K.N., Naylor, G.E., 1993. Inhibition of the rumen anaerobic fungus *Neocallimastix*

frontalis by fermentation products. *Lett. Appl. Microbiol.* 16, 254–256.

<https://doi.org/10.1111/j.1472-765X.1993.tb01412.x>

Joblin, K.N., Naylor, G.E., Williams, A.G., 1990. Effect of *Methanobrevibacter smithii* on Xylanolytic Activity of Anaerobic Ruminal Fungi. *Appl. Environ. Microbiol.* 56.

Joblin, K.N., Williams, A., 1991. Effect of cocultivation of ruminal chytrid fungi with *Methanobrevibacter smithii* on lucerne stem degradation and extracellular fungal enzyme activities. *Lett. Appl. Microbiol.* 12, 121–124. <https://doi.org/10.1111/j.1472-765X.1991.tb00520.x>

Junicke, H., Abbas, B., Oentoro, J., van Loosdrecht, M., Kleerebezem, R., 2014. Absolute Quantification of Individual Biomass Concentrations in a Methanogenic Coculture. *AMB Express* 4, 1–8. <https://doi.org/10.1186/s13568-014-0035-x>

Khandelwal, R.A., Olivier, B.G., Ling, R., Teusink, W., Bruggeman, B.J., 2013. Community Flux Balance Analysis for Microbial Consortia at Balanced Growth. *PLoS One* 8, 64567. <https://doi.org/10.1371/journal.pone.0064567>

Kim, B., Kim, W.J., Kim, D.I., Lee, S.Y., 2015. Applications of genome-scale metabolic network model in metabolic engineering. *J. Ind. Microbiol. Biotechnol.* 42, 339–348. <https://doi.org/10.1007/s10295-014-1554-9>

Kim, D., Paggi, J.M., Park, C., Bennett, C., Salzberg, S.L., 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* 37, 907–915. <https://doi.org/10.1038/s41587-019-0201-4>

Kleiner, M., 2019. Metaproteomics: Much More than Measuring Gene Expression in Microbial Communities. *mSystems* 4, e00115-19. <https://doi.org/10.1128/mSystems.00115-19>

- Kleiner, M., Thorson, E., Sharp, C.E., Dong, X., Liu, D., Li, C., Strous, M., 2017. Assessing species biomass contributions in microbial communities via metaproteomics. *Nat. Commun.* 8, 1558. <https://doi.org/10.1038/s41467-017-01544-x>
- Kovaka, S., Zimin, A. V., Pertea, G.M., Razaghi, R., Salzberg, S.L., Pertea, M., 2019. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* 2019 20, 1–13. <https://doi.org/10.1186/S13059-019-1910-1>
- Kuang, E., Marney, M., Cuevas, D., Edwards, R.A., Forsberg, E.M., 2020. Towards Predicting Gut Microbial Metabolism: Integration of Flux Balance Analysis and Untargeted Metabolomics. *Metabolites* 10, 156. <https://doi.org/10.3390/metabo10040156>
- Kumar, M., Ji, B., Babaei, P., Das, P., Lappa, D., Ramakrishnan, G., Fox, T.E., Haque, R., Petri, W.A., Bäckhed, F., Nielsen, J., 2018. Gut microbiota dysbiosis is associated with malnutrition and reduced plasma amino acid levels: Lessons from genome-scale metabolic modeling. *Metab. Eng.* 49, 128–142. <https://doi.org/10.1016/j.ymben.2018.07.018>
- Kumar, M., Ji, B., Zengler, K., Nielsen, J., 2019. Modelling approaches for studying the microbiome. *Nat. Microbiol.* 4, 1253–1267. <https://doi.org/10.1038/s41564-019-0491-9>
- Lankiewicz, T.S., Choudhary, H., Gao, Y., Amer, B., Lillington, S.P., Leggieri, P.A., Brown, J.L., Swift, C.L., Lipzen, A., Na, H., Amirebrahimi, M., Theodorou, M.K., Baidoo, E.E.K., Barry, K., Grigoriev, I. V., Timokhin, V.I., Gladden, J., Singh, S., Mortimer, J.C., Ralph, J., Simmons, B.A., Singer, S.W., O'Malley, M.A., 2023. Lignin deconstruction by anaerobic fungi. *Nat. Microbiol.* <https://doi.org/10.1038/S41564-023-01336-8>

- Lawson, C.E., Harcombe, W.R., Hatzenpichler, R., Lindemann, S.R., Löffler, F.E., O'Malley, M.A., García Martín, H., Pflieger, B.F., Raskin, L., Venturelli, O.S., Weissbrodt, D.G., Noguera, D.R., McMahon, K.D., 2019. Common principles and best practices for engineering microbiomes. *Nat. Rev. Microbiol.* 17, 725–741. <https://doi.org/10.1038/s41579-019-0255-9>
- Lee, S.S., Ha, J.K., Cheng, K., Beever, D.E., Prins, R.A., 2000. Relative contributions of bacteria, protozoa, and fungi to in vitro degradation of orchard grass cell walls and their interactions. *Appl. Environ. Microbiol.* 66, 3807–13. <https://doi.org/10.1128/aem.66.9.3807-3813.2000>
- Leggieri, P.A., Kerdman-Andrade, C., Lankiewicz, T.S., Valentine, M.T., O'Malley, M.A., 2021a. Non-destructive quantification of anaerobic gut fungi and methanogens in co-culture reveals increased fungal growth rate and changes in metabolic flux relative to mono-culture. *Microb. Cell Factories* 2021 201 20, 1–16. <https://doi.org/10.1186/S12934-021-01684-2>
- Leggieri, P.A., Liu, Y., Hayes, M., Connors, B., Seppälä, S., O'Malley, M.A., Venturelli, O.S., 2021b. Integrating Systems and Synthetic Biology to Understand and Engineer Microbiomes. *Annu. Rev. Biomed. Eng.* 23. <https://doi.org/10.1146/annurev-bioeng-082120-022836>
- Leggieri, P.A., Valentine, M.T., O'Malley, M.A., 2022. Biofilm disruption enhances growth rate and carbohydrate-active enzyme production in anaerobic fungi. *Bioresour. Technol.* 358, 127361. <https://doi.org/10.1016/J.BIORTECH.2022.127361>
- Leng, R.A., 2017. Biofilm compartmentalisation of the rumen microbiome: modification of fermentation and degradation of dietary toxins. *Anim. Prod. Sci.* 57, 2188–2203.

<https://doi.org/10.1071/an17382>

Leroy, C., Delbarre, C., Ghillebaert, F., Compere, C., Combes, D., 2008. Influence of subtilisin on the adhesion of a marine bacterium which produces mainly proteins as extracellular polymers. *J. Appl. Microbiol.* 105, 791–799.

<https://doi.org/10.1111/J.1365-2672.2008.03837.X>

Li, Y., Jin, W., Mu, C., Cheng, Y., Zhu, W., 2017. Indigenously associated methanogens intensified the metabolism in hydrogenosomes of anaerobic fungi with xylose as substrate. *J. Basic Microbiol.* 57, 933–940. <https://doi.org/10.1002/jobm.201700132>

Li, Yuanfei, Li, Yuqi, Jin, W., Sharpton, T.J., Mackie, R.I., Cann, I., Cheng, Y., Zhu, W., 2019. 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.* 10.

<https://doi.org/10.3389/fmicb.2019.00435>

Lieven, C., Beber, M.E., Olivier, B.G., Bergmann, F.T., Ataman, M., Babaei, P., Bartell, J.A., Blank, L.M., Chauhan, S., Correia, K., Diener, C., Dräger, A., Ebert, B.E., Edirisinghe, J.N., Faria, J.P., Feist, A.M., Fengos, G., Fleming, R.M.T., García-Jiménez, B., Hatzimanikatis, V., van Helvoirt, W., Henry, C.S., Hermjakob, H., Herrgård, M.J., Kaafarani, A., Kim, H.U., King, Z., Klamt, S., Klipp, E., Koehorst, J.J., König, M., Lakshmanan, M., Lee, D.Y., Lee, S.Y., Lee, S., Lewis, N.E., Liu, F., Ma, H., Machado, D., Mahadevan, R., Maia, P., Mardinoglu, A., Medlock, G.L., Monk, J.M., Nielsen, J., Nielsen, L.K., Nogales, J., Nookaew, I., Palsson, B.O., Papin, J.A., Patil, K.R., Poolman, M., Price, N.D., Resendis-Antonio, O., Richelle, A., Rocha, I., Sánchez, B.J., Schaap, P.J., Malik Sheriff, R.S., Shoaie, S., Sonnenschein, N., Teusink, B., Vilaça, P.,

- Vik, J.O., Wodke, J.A.H., Xavier, J.C., Yuan, Q., Zakhartsev, M., Zhang, C., 2020. MEMOTE for standardized genome-scale metabolic model testing. *Nat. Biotechnol.* 38, 272–276. <https://doi.org/10.1038/s41587-020-0446-y>
- Lillington, S.P., Leggieri, P.A., Heom, K.A., O'Malley, M.A., 2020. Nature's recyclers: anaerobic microbial communities drive crude biomass deconstruction. *Curr. Opin. Biotechnol.* 62, 38–47. <https://doi.org/10.1016/j.copbio.2019.08.015>
- Liu, J.J., Madec, J.Y., Bousquet-Mélou, A., Haenni, M., Ferran, A.A., 2021. Destruction of *Staphylococcus aureus* biofilms by combining an antibiotic with subtilisin A or calcium gluconate. *Sci. Reports* 2021 11, 1–12. <https://doi.org/10.1038/s41598-021-85722-4>
- Long, C.P., Antoniewicz, M.R., 2019. High-resolution <sup>13</sup>C metabolic flux analysis. *Nat. Protoc.* 14, 2856–2877. <https://doi.org/10.1038/s41596-019-0204-0>
- Lou, J., Yang, L., Wang, H., Wu, L., Xu, J., 2018. Assessing soil bacterial community and dynamics by integrated high-throughput absolute abundance quantification. *PeerJ* 2018. <https://doi.org/10.7717/peerj.4514>
- Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014 15, 1–21. <https://doi.org/10.1186/S13059-014-0550-8>
- Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Marchler, G.H., Song, J.S., Thanki, N., Yamashita, R.A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C.J., Marchler-Bauer, A., 2020. CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* 48. <https://doi.org/10.1093/nar/gkz991>
- Magnúsdóttir, S., Heinken, A., Fleming, R.M.T., Thiele, I., 2018. Magnúsdóttir et al. Reply:

- Babaei et al. *Nat. Biotechnol.* 36, 686–691. <https://doi.org/10.1038/nbt.4212>
- Magnúsdóttir, S., Heinken, A., Kutt, L., Ravcheev, D.A., Bauer, E., Noronha, A., Greenhalgh, K., Jäger, C., Baginska, J., Wilmes, P., T Fleming, R.M., Thiele, I., 2017. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat. Biotechnol.* Vol. 35. <https://doi.org/10.1038/nbt.3703>
- Mao, W., Yuan, Q., Qi, H., Wang, Z., Ma, H., Chen, T., 2020. Recent progress in metabolic engineering of microbial formate assimilation. *Appl. Microbiol. Biotechnol.* 104, 6905–6917. <https://doi.org/10.1007/s00253-020-10725-6>
- Marchler-Bauer, A., Bryant, S.H., 2004. CD-Search: Protein domain annotations on the fly. *Nucleic Acids Res.* 32. <https://doi.org/10.1093/nar/gkh454>
- Marvin-Sikkema, F.D., Driessen, A.J.M., Gottschal, J.C., Prins, R.A., 1994. Metabolic energy generation in hydrogenosomes of the anaerobic fungus *Neocallimastix*: evidence for a functional relationship with mitochondria. *Mycol. Res.* 98, 205–212. [https://doi.org/10.1016/S0953-7562\(09\)80187-1](https://doi.org/10.1016/S0953-7562(09)80187-1)
- Marvin-Sikkema, F.D., Pedro Gomes, T.M., Grivet, J.P., Gottschal, J.C., Prins, R.A., 1993. Characterization of hydrogenosomes and their role in glucose metabolism of *Neocallimastix* sp. L2. *Arch. Microbiol.* 160, 388–396. <https://doi.org/10.1007/BF00252226>
- Marvin-Sikkema, F.D., Richardson, A.J., Stewart, C.S., Gottschal, J.C., Prins, R.A., 1990. Influence of hydrogen-consuming bacteria on cellulose degradation by anaerobic fungi. *Appl. Environ. Microbiol.* 56, 3793–3797. <https://doi.org/10.1128/aem.56.12.3793-3797.1990>
- Matsunaga, A., Tsugawa, M., Fortes, J., 2008. CloudBLAST: combining MapReduce and

- virtualization on distributed resources for bioinformatics applications, in: Proceedings - 4th IEEE International Conference on EScience, EScience 2008. pp. 222–229.  
<https://doi.org/10.1109/eScience.2008.62>
- Mendoza, S.N., Olivier, B.G., Molenaar, D., Teusink, B., 2019. A systematic assessment of current genome-scale metabolic reconstruction tools. *Genome Biol.* 20, 158.  
<https://doi.org/10.1186/s13059-019-1769-1>
- Miller, T.L., Lin, C., 2002. Description of *Methanobrevibacter gottschalkii* sp. nov., *Methanobrevibacter thaueri* sp. nov., *Methanobrevibacter woesei* sp. nov. and *Methanobrevibacter wolinii* sp. nov.. *Int. J. Syst. Evol. Microbiol.* 52, 819–822.  
<https://doi.org/10.1099/00207713-52-3-819>
- Mitrofanova, O., Mardanova, A., Evtugyn, V., Bogomolnaya, L., Sharipova, M., 2017. Effects of *Bacillus* Serine Proteases on the Bacterial Biofilms. *Biomed Res. Int.* 2017.  
<https://doi.org/10.1155/2017/8525912>
- Mosbæk, F., Kjeldal, H., Mulat, D.G., Albertsen, M., Ward, A.J., Feilberg, A., Nielsen, J.L., 2016. Identification of syntrophic acetate-oxidizing bacteria in anaerobic digesters by combined protein-based stable isotope probing and metagenomics. *ISME J.* 10, 2405–2418. <https://doi.org/10.1038/ismej.2016.39>
- Mountfort, D.O., Asher, R.A., Bauchop, T., 1982. 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.* 44, 128–134.  
<https://doi.org/10.1128/aem.44.1.128-134.1982>
- Mueller, R.S., Pan, C., 2013. Sample Handling and Mass Spectrometry for Microbial Metaproteomic Analyses. *Methods Enzymol.* 531, 289–303.

<https://doi.org/10.1016/B978-0-12-407863-5.00015-0>

Müller, M., Mentel, M., van Hellemond, J.J., Henze, K., Woehle, C., Gould, S.B., Yu, R.-Y., van der Giezen, M., Tielens, A.G.M., Martin, W.F., 2012. Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol. Mol. Biol. Rev.* 76, 444–95.

<https://doi.org/10.1128/MMBR.05024-11>

Namkung, J., 2020. Machine learning methods for microbiome studies. *J. Microbiol.* 58, 206–216. <https://doi.org/10.1007/s12275-020-0066-8>

Nayfach, S., Rodriguez-Mueller, B., Garud, N., Pollard, K.S., 2016. An integrated metagenomics pipeline for strain profiling reveals novel patterns of bacterial transmission and biogeography. *Genome Res.* 26, 1612–1625.

<https://doi.org/10.1101/gr.201863.115>

Nayfach, S., Shi, Z.J., Seshadri, R., Pollard, K.S., Kyrpides, N.C., 2019. New insights from uncultivated genomes of the global human gut microbiome. *Nature* 568, 505–510.

<https://doi.org/10.1038/s41586-019-1058-x>

Nettmann, E., Fröhling, A., Heeg, K., Klocke, M., Schlüter, O., Mumme, J., 2013.

Development of a flow-fluorescence in situ hybridization protocol for the analysis of microbial communities in anaerobic fermentation liquor. *BMC Microbiol.* 13, 278.

<https://doi.org/10.1186/1471-2180-13-278>

Nilsson, A., Nielsen, J., Palsson, B.O., 2017. Metabolic Models of Protein Allocation Call for the Kinetome. *Cell Syst.* 5, 538–541. <https://doi.org/10.1016/j.cels.2017.11.013>

Noronha, A., Modamio, J., Jarosz, Y., Guerard, E., Sompairac, N., Preciat, G., Dröfn, A., Dröfn, D., Daníelsdóttir, D., Krecke, M., Merten, D., Haraldsdóttir, H.S., Haraldsdóttir, H., Heinken, A., Heirendt, L., Magnúsdóttir, S., Magnúsd, M.,

Magnúsdóttir, M., Ravcheev, D.A., Sahoo, S., Gawron, P., Friscioni, L., Garcia, B., Prendergast, M., Puente, A., Rodrigues, M., Roy, A., Rouquaya, M., Wiltgen, L., Zagare, A., John, E., Krueger, M., Kuperstein, I., Zinovyev, A., Schneider, R., Fleming, R.M.T., Thiele, I., 2019. The Virtual Metabolic Human database: integrating human and gut microbiome metabolism with nutrition and disease. *Nucleic Acids Res.* 47.  
<https://doi.org/10.1093/nar/gky992>

Norsigian, C.J., Pusarla, N., Mcconn, J.L., Yurkovich, J.T., Dräger, A., Dräger, D., Palsson, B.O., King, Z., 2020. BiGG Models 2020: multi-strain genome-scale models and expansion across the phylogenetic tree. *Nucleic Acids Res.* 48.  
<https://doi.org/10.1093/nar/gkz1054>

Nowak, M.A., 2006. *Evolutionary Dynamics: Exploring the Equations of Life*. Harvard University Press.

Orth, J.D., Thiele, I., Palsson, B.O., 2010. What is flux balance analysis? *Nat. Biotechnol.* 28, 245–248. <https://doi.org/10.1038/nbt.1614>

Øyås, O., Stelling, J., 2018. Genome-scale metabolic networks in time and space. *Curr. Opin. Syst. Biol.* 8, 51–58. <https://doi.org/10.1016/J.COISB.2017.12.003>

Ozbakir, H.F., Anderson, N.T., Fan, K.C., Mukherjee, A., 2020. Beyond the Green Fluorescent Protein: Biomolecular Reporters for Anaerobic and Deep-Tissue Imaging. *Bioconjug. Chem.* 31, 293–302. <https://doi.org/10.1021/acs.bioconjchem.9b00688>

Pandey, V., Hadadi, N., Hatzimanikatis, V., 2019. Enhanced flux prediction by integrating relative expression and relative metabolite abundance into thermodynamically consistent metabolic models. *PLOS Comput. Biol.* 15, e1007036.  
<https://doi.org/10.1371/journal.pcbi.1007036>

- Patel, A., Carlson, R.P., Henson, M.A., 2019. In Silico Metabolic Design of Two-Strain Biofilm Systems Predicts Enhanced Biomass Production and Biochemical Synthesis. *Biotechnol. J.* 14. <https://doi.org/10.1002/biot.201800511>
- Patel, A., Shah, A.R., 2021. Integrated lignocellulosic biorefinery: Gateway for production of second generation ethanol and value added products. *J. Bioresour. Bioprod.* 6, 108–128. <https://doi.org/https://doi.org/10.1016/j.jobab.2021.02.001>
- Peck, M.W., 1989. Changes in concentrations of coenzyme F420 analogs during batch growth of *Methanosarcina barkeri* and *Methanosarcina mazei*. *Appl. Environ. Microbiol.* 55, 940–945. <https://doi.org/10.1128/aem.55.4.940-945.1989>
- Peng, X. “Nick,” Gilmore, S.P., O’Malley, M.A., 2016. Microbial communities for bioprocessing: lessons learned from nature. *Curr. Opin. Chem. Eng.* 14, 103–109. <https://doi.org/10.1016/J.COCHE.2016.09.003>
- Peng, X., Swift, C.L., Theodorou, M.K., O’Malley, M.A., 2018. Methods for genomic characterization and maintenance of anaerobic fungi, in: *Methods in Molecular Biology*. Humana Press Inc., pp. 53–67. [https://doi.org/10.1007/978-1-4939-7804-5\\_5](https://doi.org/10.1007/978-1-4939-7804-5_5)
- Peng, X., Wilken, S.E., Lankiewicz, T.S., Gilmore, S.P., Brown, J.L., Henske, J.K., Swift, C.L., Salamov, A., Barry, K., Grigoriev, I. V., Theodorou, M.K., Valentine, D.L., O’Malley, M.A., 2021. Genomic and functional analyses of fungal and bacterial consortia that enable lignocellulose breakdown in goat gut microbiomes. *Nat. Microbiol.* 6, 499–511. <https://doi.org/10.1038/s41564-020-00861-0>
- Petrova, O.E., Garcia-Alcalde, F., Zampaloni, C., Sauer, K., 2017. Comparative evaluation of rRNA depletion procedures for the improved analysis of bacterial biofilm and mixed pathogen culture transcriptomes. *Sci. Rep.* 7, 41114. <https://doi.org/10.1038/srep41114>

- Phalak, P., Chen, J., Carlson, R.P., Henson, M.A., 2016. Metabolic modeling of a chronic wound biofilm consortium predicts spatial partitioning of bacterial species. *BMC Syst. Biol.* 10, 90. <https://doi.org/10.1186/s12918-016-0334-8>
- Price, M.N., Wetmore, K.M., Waters, R.J., Callaghan, M., Ray, J., Liu, H., Kuehl, J. V., Melnyk, R.A., Lamson, J.S., Suh, Y., Carlson, H.K., Esquivel, Z., Sadeeshkumar, H., Chakraborty, R., Zane, G.M., Rubin, B.E., Wall, J.D., Visel, A., Bristow, J., Blow, M.J., Arkin, A.P., Deutschbauer, A.M., 2018. Mutant phenotypes for thousands of bacterial genes of unknown function. *Nature* 557, 503–509. <https://doi.org/10.1038/s41586-018-0124-0>
- Pusa, T., Wannagat, M., Sagot, M.-F., 2019. Metabolic Games. *Front. Appl. Math. Stat.* 5, 18. <https://doi.org/10.3389/fams.2019.00018>
- Radajewski, S., McDonald, I.R., Murrell, J.C., 2003. Stable-isotope probing of nucleic acids: a window to the function of uncultured microorganisms. *Curr. Opin. Biotechnol.* 14, 296–302. [https://doi.org/10.1016/S0958-1669\(03\)00064-8](https://doi.org/10.1016/S0958-1669(03)00064-8)
- Raja, H.A., Miller, A.N., Pearce, C.J., Oberlies, N.H., 2017. Fungal Identification Using Molecular Tools: A Primer for the Natural Products Research Community. *J. Nat. Prod.* 80, 756–770. <https://doi.org/10.1021/acs.jnatprod.6b01085>
- Reimers, A.-M., Lindhorst, H., Waldherr, S., 2017. A Protocol for Generating and Exchanging (Genome-Scale) Metabolic Resource Allocation Models. *Metabolites* 7, 47. <https://doi.org/10.3390/metabo7030047>
- Robinson, C.D., Auchtung, J.M., Collins, J., Britton, R.A., 2014. Epidemic *Clostridium difficile* Strains Demonstrate Increased Competitive Fitness Compared to Nonepidemic Isolates. *Infect. Immun.* 82, 2815–2825. <https://doi.org/10.1128/IAI.01524-14>

- Robinson, J.L., Kocabaş, P., Wang, H., Cholley, P.E., Cook, D., Nilsson, A., Anton, M., Ferreira, R., Domenzain, I., Billa, V., Limeta, A., Hedin, A., Gustafsson, J., Kerkhoven, E.J., Svensson, L.T., Palsson, B.O., Mardinoglu, A., Hansson, L., Uhlén, M., Nielsen, J., 2020. An atlas of human metabolism. *Sci. Signal.* 13.  
<https://doi.org/10.1126/scisignal.aaz1482>
- Saye, L.M.G., Navaratna, T.A., Chong, J.P.J., O'Malley, M.A., Theodorou, M.K., Reilly, M., 2021. The anaerobic fungi: Challenges and opportunities for industrial lignocellulosic biofuel production. *Microorganisms* 9, 694.  
<https://doi.org/10.3390/microorganisms9040694>
- Scarborough, M.J., Lawson, C.E., Hamilton, J.J., Donohue, T.J., Noguera, D.R., 2018. Metatranscriptomic and Thermodynamic Insights into Medium-Chain Fatty Acid Production Using an Anaerobic Microbiome. *mSystems* 3, e00221-18.  
<https://doi.org/10.1128/mSystems.00221-18>
- Scholz, M., Ward, D. V., Pasolli, E., Tolio, T., Zolfo, M., Asnicar, F., Truong, D.T., Tett, A., Morrow, A.L., Segata, N., 2016. Strain-level microbial epidemiology and population genomics from shotgun metagenomics. *Nat. Methods* 13, 435–438.  
<https://doi.org/10.1038/nmeth.3802>
- Schut, G.J., Adams, M.W.W., 2009. The iron-hydrogenase of *Thermotoga maritima* utilizes ferredoxin and NADH synergistically: A new perspective on anaerobic hydrogen production. *J. Bacteriol.* 191. <https://doi.org/10.1128/JB.01582-08>
- Seaver, S.M.D., Liu, F., Zhang, Q., Jeffryes, J., Faria, J.P., Edirisinghe, J.N., Mundy, M., Chia, N., Noor, E., Beber, M.E., Best, A.A., DeJongh, M., Kimbrel, J.A., D'haeseleer, P., Pearson, E., Canon, S., Wood-Charlson, E.M., Cottingham, R.W., Arkin, A.P.,

- Henry, C.S., 2020. The ModelSEED Database for the integration of metabolic annotations and the reconstruction, comparison, and analysis of metabolic models for plants, fungi, and microbes. *bioRxiv* 2020.03.31.018663.  
<https://doi.org/10.1101/2020.03.31.018663>
- Seifert, J., Taubert, M., Jehmlich, N., Schmidt, F., Völker, U., Vogt, C., Richnow, H.-H., von Bergen, M., 2012. Protein-based stable isotope probing (protein-SIP) in functional metaproteomics. *Mass Spectrom. Rev.* 31, 683–697. <https://doi.org/10.1002/mas.21346>
- Seppälä, S., Wilken, S.E., Knop, D., Solomon, K. V., O'Malley, M.A., 2017. The importance of sourcing enzymes from non-conventional fungi for metabolic engineering and biomass breakdown. *Metab. Eng.* 44, 45–59.  
<https://doi.org/10.1016/j.ymben.2017.09.008>
- Sgobba, E., Wendisch, V.F., 2020. Synthetic microbial consortia for small molecule production. *Curr. Opin. Biotechnol.* 62, 72–79.  
<https://doi.org/10.1016/j.copbio.2019.09.011>
- Shaffer, M., Borton, M.A., McGivern, B.B., Zayed, A.A., Rosa, S.L. La, Solden, L.M., Liu, P., Narrowe, A.B., Rodriguez-Ramos, J., Bolduc, B., Gazitua, M.C., Daly, R.A., Smith, G.J., Vik, D.R., Pope, P.B., Sullivan, M.B., Roux, S., Wrighton, K.C., 2020. DRAM for distilling microbial metabolism to automate the curation of microbiome function. *bioRxiv* 2020.06.29.177501. <https://doi.org/10.1101/2020.06.29.177501>
- Shamir, M., Bar-On, Y., Phillips, R., Milo, R., 2016. SnapShot: Timescales in Cell Biology. *Cell* 164, 1302-1302.e1. <https://doi.org/10.1016/j.cell.2016.02.058>
- Sheth, R.U., Li, M., Jiang, W., Sims, P.A., Leong, K.W., Wang, H.H., 2019. Spatial metagenomic characterization of microbial biogeography in the gut. *Nat. Biotechnol.*

37, 877–883. <https://doi.org/10.1038/s41587-019-0183-2>

Shi, X., Shao, C., Luo, C., Chu, Y., Wang, J., Meng, Q., Yu, J., Gao, Z., Kang, Y., 2019.

Microfluidics-Based Enrichment and Whole-Genome Amplification Enable Strain-Level Resolution for Airway Metagenomics. *mSystems* 4.

<https://doi.org/10.1128/msystems.00198-19>

Shoaie, S., Ghaffari, P., Kovatcheva-Datchary, P., Mardinoglu, A., Sen, P., Pujos-Guillot, E.,

de Wouters, T., Juste, C., Rizkalla, S., Chilloux, J., Hoyles, L., Nicholson, J.K., Dore, J.,

Dumas, M.E., Clement, K., Bäckhed, F., Nielsen, J., 2015. Quantifying Diet-Induced Metabolic Changes of the Human Gut Microbiome. *Cell Metab.* 22, 320–331.

<https://doi.org/10.1016/j.cmet.2015.07.001>

Solden, L.M., Naas, A.E., Roux, S., Daly, R.A., Collins, W.B., Nicora, C.D., Purvine, S.O.,

Hoyt, D.W., Schückel, J., Jørgensen, B., Willats, W., Spalinger, D.E., Firkins, J.L.,

Lipton, M.S., Sullivan, M.B., Pope, P.B., Wrighton, K.C., 2018. Interspecies cross-feeding orchestrates carbon degradation in the rumen ecosystem. *Nat. Microbiol.* 3,

1274–1284. <https://doi.org/10.1038/s41564-018-0225-4>

Solomon, K. V, Haitjema, C.H., Henske, J.K., Gilmore, S.P., Borges-Rivera, D., Lipzen, A.,

Brewer, H.M., Purvine, S.O., Wright, A.T., Theodorou, M.K., Grigoriev, I. V, Regev,

A., Thompson, D.A., O'Malley, M.A., 2016. Early-branching gut fungi possess large, comprehensive array of biomass-degrading enzymes. *Science* 351, 1192–1195.

<https://doi.org/10.1126/science.aad1431>

Song, H.S., Cannon, W.R., Beliaev, A.S., Konopka, A., 2014. Mathematical modeling of

microbial community dynamics: A methodological review. *Processes* 2, 711–752.

<https://doi.org/10.3390/pr2040711>

- Speda, J., Jonsson, B.-H., Carlsson, U., Karlsson, M., 2017. Metaproteomics-guided selection of targeted enzymes for bioprospecting of mixed microbial communities. *Biotechnol. Biofuels* 10, 128. <https://doi.org/10.1186/s13068-017-0815-z>
- Srinivasan, K., Murakami, M., Nakashimada, Y., Nishio, N., 2001. Efficient production of cellulolytic and xylanolytic enzymes by the rumen anaerobic fungus, *neocallimastix frontalis*, in a repeated batch culture. *J. Biosci. Bioeng.* 91, 153–158. [https://doi.org/10.1016/S1389-1723\(01\)80058-X](https://doi.org/10.1016/S1389-1723(01)80058-X)
- Stamatopoulou, P., Malkowski, J., Conrado, L., Brown, K., Scarborough, M., 2020. Fermentation of Organic Residues to Beneficial Chemicals: A Review of Medium-Chain Fatty Acid Production. *Processes* 8. <https://doi.org/10.3390/pr8121571>
- Starr, E.P., Shi, S., Blazewicz, S.J., Probst, A.J., Herman, D.J., Firestone, M.K., Banfield, J.F., 2018. Stable isotope informed genome-resolved metagenomics reveals that *Saccharibacteria* utilize microbially-processed plant-derived carbon. *Microbiome* 6, 122. <https://doi.org/10.1186/s40168-018-0499-z>
- Stein, R.R., Bucci, V., Toussaint, N.C., Buffie, C.G., Räscht, G., Pamer, E.G., Sander, C., Xavier, J.B., 2013. Ecological Modeling from Time-Series Inference: Insight into Dynamics and Stability of Intestinal Microbiota. *PLoS Comput. Biol.* 9, e1003388. <https://doi.org/10.1371/journal.pcbi.1003388>
- Stolyar, S., Van Dien, S., Hillesland, K.L., Pineda, N., Lie, T.J., Leigh, J.A., Stahl, D.A., 2007. Metabolic modeling of a mutualistic microbial community. *Mol. Syst. Biol.* 3, 92. <https://doi.org/10.1038/msb4100131>
- Swift, C.L., Brown, J.L., Seppälä, S., O'Malley, M.A., 2019. Co-cultivation of the anaerobic fungus *Anaeromyces robustus* with *Methanobacterium bryantii* enhances transcription

of carbohydrate active enzymes. *J. Ind. Microbiol. Biotechnol.* 46, 1427–1433.

<https://doi.org/10.1007/s10295-019-02188-0>

Taya, M., Aoki, N., Kobayashi, T., 1986. Kinetic evaluation and monitoring of methanogen culture based upon fluorometry. *Appl. Microbiol. Biotechnol.* 23, 342–347.

<https://doi.org/10.1007/BF00257030>

Teunissen, M.J., Kets, E.P.W., Op den Camp, H.J.M., Huis in't Veld, J.H.J., Vogels, G.D., 1992. Effect of coculture of anaerobic fungi isolated from ruminants and non-ruminants with methanogenic bacteria on cellulolytic and xylanolytic enzyme activities. *Arch. Microbiol.* 157, 176–182. <https://doi.org/10.1007/BF00245287>

The Gene Ontology Consortium, 2021. The Gene Ontology resource: Enriching a Gold mine. *Nucleic Acids Res.* 49, D325–D334. <https://doi.org/10.1093/nar/gkaa1113>

The Gene Ontology Consortium, Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene Ontology: tool for the unification of biology. *Nat. Genet.* 25, 25. <https://doi.org/10.1038/75556>

Theodorou, M.K., Davies, D.R., Nielsen, B.B., Lawrence, M.I.G., Trinci, A.P.J., 1995.

Determination of growth of anaerobic fungi on soluble and cellulosic substrates using a pressure transducer. *Microbiology* 141, 671–678. <https://doi.org/10.1099/13500872-141-3-671>

Thiele, I., Palsson, B., 2010. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* 5, 93–121.

<https://doi.org/10.1038/nprot.2009.203>

- Thiele, I., Sahoo, S., Heinken, A., Hertel, J., Heirendt, L., Aurich, M.K., Fleming, R.M., 2020. Personalized whole-body models integrate metabolism, physiology, and the gut microbiome. *Mol. Syst. Biol.* 16. <https://doi.org/10.15252/msb.20198982>
- Thornbury, M., Sicheri, J., Slaine, P., Getz, L.J., Finlayson-Trick, E., Cook, J., Guinard, C., Boudreau, N., Jakeman, D., Rohde, J., McCormick, C., 2019. Characterization of novel lignocellulose-degrading enzymes from the porcupine microbiome using synthetic metagenomics. *PLoS One* 14, e0209221. <https://doi.org/10.1371/journal.pone.0209221>
- Tian, M., Reed, J.L., 2018. Integrating Proteomic or Transcriptomic Data into Metabolic Models Using Linear Bound Flux Balance Analysis. *Bioinformatics* 34, 3882–3888. <https://doi.org/10.1093/bioinformatics/bty445>
- Tkacz, A., Hortala, M., Poole, P.S., 2018. Absolute quantitation of microbiota abundance in environmental samples. *Microbiome* 6. <https://doi.org/10.1186/s40168-018-0491-7>
- Toor, M., Kumar, S.S., Malyan, S.K., Bishnoi, N.R., Mathimani, T., Rajendran, K., Pugazhendhi, A., 2020. An overview on bioethanol production from lignocellulosic feedstocks. *Chemosphere* 242, 125080. <https://doi.org/https://doi.org/10.1016/j.chemosphere.2019.125080>
- Traversi, D., Villa, S., Lorenzi, E., Degan, R., Gilli, G., 2012. Application of a real-time qPCR method to measure the methanogen concentration during anaerobic digestion as an indicator of biogas production capacity. *J. Environ. Manage.* 111, 173–177. <https://doi.org/10.1016/j.jenvman.2012.07.021>
- Truong, D.T., Tett, A., Pasolli, E., Huttenhower, C., Segata, N., 2017. Microbial strain-level population structure and genetic diversity from metagenomes. *Genome Res.* 27, 626–638. <https://doi.org/10.1101/gr.216242.116>

van Dijk, E.L., Jaszczyszyn, Y., Naquin, D., Thermes, C., 2018. The Third Revolution in Sequencing Technology. *Trends Genet.* 34, 666–681.

<https://doi.org/10.1016/J.TIG.2018.05.008>

Vandeputte, D., Kathagen, G., D’Hoe, K., Vieira-Silva, S., Valles-Colomer, M., Sabino, J., Wang, J., Tito, R.Y., De Commer, L., Darzi, Y., Vermeire, S., Falony, G., Raes, J., 2017. Quantitative microbiome profiling links gut community variation to microbial load. *Nature* 551, 507–511. <https://doi.org/10.1038/nature24460>

Venturelli, O.S., Carr, A. V, Fisher, G., Hsu, R.H., Lau, R., Bowen, B.P., Hromada, S., Northen, T., Arkin, A.P., 2018. Deciphering microbial interactions in synthetic human gut microbiome communities. *Mol. Syst. Biol.* 14, e8157.

<https://doi.org/10.15252/msb.20178157>

Vinzelj, J., Joshi, A., Insam, H., Podmirseg, S.M., 2020. Employing anaerobic fungi in biogas production: challenges & opportunities. *Bioresour. Technol.* 300.

<https://doi.org/10.1016/j.biortech.2019.122687>

Vit, O., Petrak, J., 2017. Integral membrane proteins in proteomics. How to break open the black box? *J. Proteomics* 153, 8–20. <https://doi.org/10.1016/j.jprot.2016.08.006>

Wang, R., Zhao, S., Wang, Z., Koffas, M.A.,. Recent advances in modular co-culture engineering for synthesis of natural products, *Current Opinion in Biotechnology.*

Elsevier Ltd. <https://doi.org/10.1016/j.copbio.2019.09.004>

Wangsanuwat, C., Heom, K.A., Liu, E., O’malley, M.A., Dey, S.S., 2020. Efficient and cost-effective bacterial mRNA sequencing from low input samples through ribosomal RNA

depletion. *bioRxiv* 2020.06.19.162412. <https://doi.org/10.1101/2020.06.19.162412>

Wilken, S., Saxena, M., Petzold, L., O’Malley, M., 2018. In Silico Identification of Microbial

Partners to Form Consortia with Anaerobic Fungi. *Processes* 6, 7.

<https://doi.org/10.3390/pr6010007>

Wilken, S.E., Leggieri, P.A., Kerdman-Andrade, C., Reilly, M., Theodorou, M.K., O'Malley, M.A., 2020. An Arduino based Automatic Pressure Evaluation System (A-APES) to quantify growth of non-model anaerobes in culture. *AIChE J.* 66.

<https://doi.org/10.1002/aic.16540>

Wilken, S.E., Monk, J.M., Leggieri, P.A., Lawson, C.E., Lankiewicz, T.S., Seppälä, S., Daum, C.G., Jenkins, J., Lipzen, A.M., Mondo, S.J., Barry, K.W., Grigoriev, I. V., Henske, J.K., Theodorou, M.K., Palsson, B.O., Petzold, L.R., O'Malley, M.A., 2021. Experimentally Validated Reconstruction and Analysis of a Genome-Scale Metabolic Model of an Anaerobic Neocallimastigomycota Fungus. *mSystems* 6.

<https://doi.org/10.1128/msystems.00002-21>

Wilson, J.-J., Brandon-Mong, G.-J., Gan, H.-M., Sing, K.-W., 2019. High-throughput terrestrial biodiversity assessments: mitochondrial metabarcoding, metagenomics or metatranscriptomics? *Mitochondrial DNA Part A* 30, 60–67.

<https://doi.org/10.1080/24701394.2018.1455189>

Xiros, C., Studer, M.H., 2017. A Multispecies Fungal Biofilm Approach to Enhance the Cellulolytic Efficiency of Membrane Reactors for Consolidated Bioprocessing of Plant Biomass. *Front. Microbiol.* 8, 1930. <https://doi.org/10.3389/fmicb.2017.01930>

Yarlett, N., Orpin, C.G., Munn, E.A., Greenwood, C.A., 1986. Hydrogenosomes in the rumen fungus *Neocallimastix patriciarum*. *Biochem. J.* 236, 729–739.

<https://doi.org/10.1042/bj2360729>

Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., Xu, Y., 2012. dbCAN: a web resource for

automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 40, W445–W451. <https://doi.org/10.1093/NAR/GKS479>

Youssef, N.H., Couger, M.B., Struchtemeyer, C.G., Liggenstoffer, A.S., Prade, R.A., Najjar, F.Z., Atiyeh, H.K., Wilkins, M.R., Elshahed, M.S., 2013. The Genome of the Anaerobic Fungus *Orpinomyces* sp. Strain C1A Reveals the Unique Evolutionary History of a Remarkable Plant Biomass Degrader. *Appl. Environ. Microbiol.* 79, 4620–4634. <https://doi.org/10.1128/AEM.00821-13>

Zamboni, N., Saghatelian, A., Patti, G.J., 2015. Defining the Metabolome: Size, Flux, and Regulation. *Mol. Cell* 58, 699–706. <https://doi.org/10.1016/j.molcel.2015.04.021>

Zampieri, G., Vijayakumar, S., Yaneske, E., Angione, C., 2019. Machine and deep learning meet genome-scale metabolic modeling. *PLOS Comput. Biol.* 15, e1007084. <https://doi.org/10.1371/journal.pcbi.1007084>

Zeng, H., Yang, A., 2020. Bridging substrate intake kinetics and bacterial growth phenotypes with flux balance analysis incorporating proteome allocation. *Sci. Rep.* 10, 1–10. <https://doi.org/10.1038/s41598-020-61174-0>

Zhang, C., Hua, Q., 2016. Applications of genome-scale metabolic models in biotechnology and systems medicine. *Front. Physiol.* 6, 413. <https://doi.org/10.3389/fphys.2015.00413>

Zhao, S., Liu, Q., Wang, J.X., Liao, X.Z., Guo, H., Li, C.X., Zhang, F.F., Liao, L.S., Luo, X.M., Feng, J.X., 2019. Differential transcriptomic profiling of filamentous fungus during solid-state and submerged fermentation and identification of an essential regulatory gene *PoxMBF1* that directly regulated cellulase and xylanase gene expression. *Biotechnol. Biofuels* 12, 103. <https://doi.org/10.1186/s13068-019-1445-4>

Zhou, Y.-H., Gallins, P., 2019. A Review and Tutorial of Machine Learning Methods for

Microbiome Host Trait Prediction. *Front. Genet.* 10, 579.

<https://doi.org/10.3389/fgene.2019.00579>

Zhu, W.Y., Theodorou, M.K., Longland, A.C., Nielsen, B.B., Dijkstra, J., Trinci, A.P.J., 1996. Growth and survival of anaerobic fungi in batch and continuous-flow cultures.

*Anaerobe* 2, 29–37. <https://doi.org/10.1006/ANAE.1996.0004>

Zhu, W.Y., Theodorou, M.K., Nielsen, B.B., Trinci, A.P.J., 1997. Dilution rate increases production of plant cell-wall degrading enzymes by anaerobic fungi in continuous-flow culture. *Anaerobe* 3, 49–59. <https://doi.org/10.1006/ANAE.1997.0070>

Ziels, R.M., Sousa, D.Z., Stensel, H.D., Beck, D.A.C., 2018. DNA-SIP based genome-centric metagenomics identifies key long-chain fatty acid-degrading populations in anaerobic digesters with different feeding frequencies. *ISME J.* 12, 112–123.

<https://doi.org/10.1038/ismej.2017.143>

Zomorodi, A.R., Maranas, C.D., 2012. OptCom: A Multi-Level Optimization Framework for the Metabolic Modeling and Analysis of Microbial Communities. *PLoS Comput. Biol.* 8, e1002363. <https://doi.org/10.1371/journal.pcbi.1002363>

<https://doi.org/10.1371/journal.pcbi.1002363>

Zomorodi, A.R., Segrè, D., 2017. Genome-driven evolutionary game theory helps understand the rise of metabolic interdependencies in microbial communities. *Nat. Commun.* 8, 1–12. <https://doi.org/10.1038/s41467-017-01407-5>

<https://doi.org/10.1038/s41467-017-01407-5>

Zuroff, T.R., Curtis, W.R., 2012. Developing symbiotic consortia for lignocellulosic biofuel production. *Appl. Microbiol. Biotechnol.* 93, 1423–1435.

<https://doi.org/10.1007/s00253-011-3762-9>

## VII. Appendices

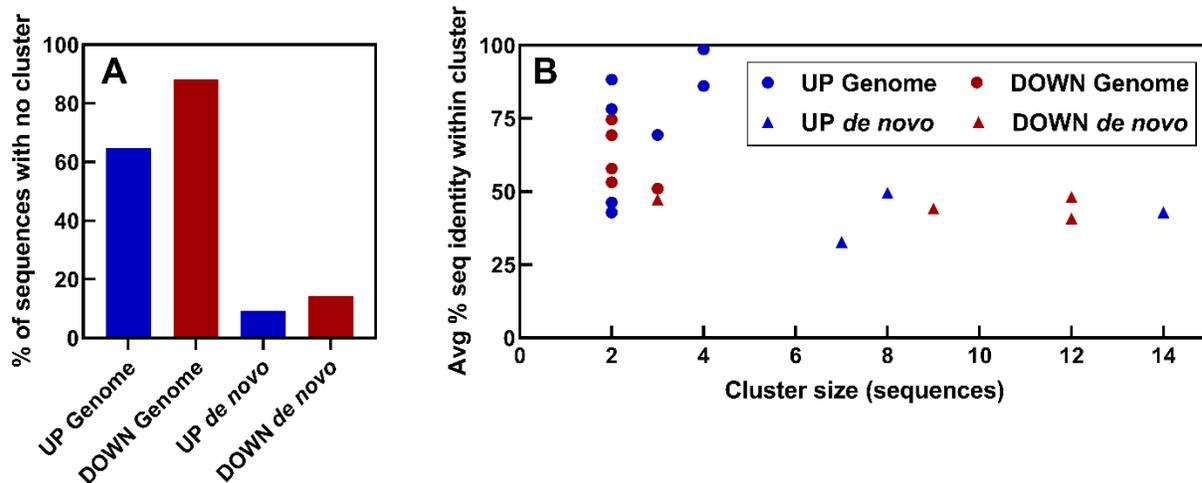
### A. Appendix: Supplementary Materials for Chapter II

**Supplementary Table S1A.** Significantly differentially expressed, secreted, unannotated proteins in *C. churrovis* with homology to potential biofilm-implicated proteins or homology to no AGF proteins at all.

Upregulated, unannotated, secreted proteins of interest	
Protein ID	Note
465686	No BLAST hits
450845	No BLAST hits
626073	No AGF BLAST hits
625875	Titin-like homology
527394	Subtilisin-like homology
527396	Subtilisin-like homology
527393	Subtilisin-like homology
527391	Subtilisin-like homology
Downregulated, unannotated, secreted proteins of interest	
Protein ID	Note
604814	No BLAST hits
442778	No BLAST hits
465611	No BLAST hits
266897	No BLAST hits
472781	No BLAST hits
447947	No BLAST hits
457444	No BLAST hits
535584	No BLAST hits
273016	No BLAST hits
444564	No BLAST hits
74656	No BLAST hits
496745	No BLAST hits
457460	No BLAST hits
457154	No BLAST hits
69190	No BLAST hits
551766	No BLAST hits
628690	No AGF BLAST hits, homology to mammalian sperm-egg adhesion protein (zonadhesin)
531507	No AGF BLAST hits, homology to extracellular peptidase in <i>Streptococcus</i>
454075	No AGF BLAST hits, homology to 1,4-alpha-glucan branching protein in <i>Blastococcus</i>
548363	Many AGF hits, also homology to zonadhesin in eel
468469	Subtilisin-like homology
522134	Subtilisin-like homology

**Supplementary Table S1B.** Significantly differentially expressed, unannotated *de novo* transcripts in *C. churrovis* with homology to potential biofilm-implicated proteins.

Upregulated, unannotated <i>de novo</i> transcripts of interest	
<i>de novo</i> Transcript ID	Note
Novel_ST_.10883	Chitin-binding GO annotation
Novel_ST_.11534	Subtilisin-like homology, chitin-binding GO annotation
Novel_ST_.11829	Subtilisin-like homology
Novel_ST_.12094	Chitin-binding GO annotation
Novel_ST_.3363	Chitin-binding GO annotation
Novel_ST_.3974	Subtilisin-like homology
Novel_ST_.774	Chitin-binding GO annotation
Novel_ST_.802	Subtilisin-like homology, chitin-binding GO annotation
Novel_ST_.8811	Chitin-binding GO annotation
Novel_ST_.9077	Chitin-binding GO annotation
Novel_ST_.9371	Chitin-binding GO annotation
Novel_ST_.9904	Chitin-binding GO annotation
Downregulated, unannotated <i>de novo</i> transcripts of interest	
<i>de novo</i> Transcript ID	Note
Novel_ST_.12702	Chitin-binding GO annotation
Novel_ST_.1576	Chitin-binding GO annotation
Novel_ST_.3350	Chitin-binding GO annotation
Novel_ST_.3931	Chitin-binding GO annotation
Novel_ST_.4336	Chitin-binding GO annotation
Novel_ST_.5771	Subtilisin-like homology, chitin-binding GO annotation
Novel_ST_.7672	Chitin-binding GO annotation
Novel_ST_.8864	Subtilisin-like homology, chitin-binding GO annotation
Novel_ST_.9690	Chitin-binding GO annotation

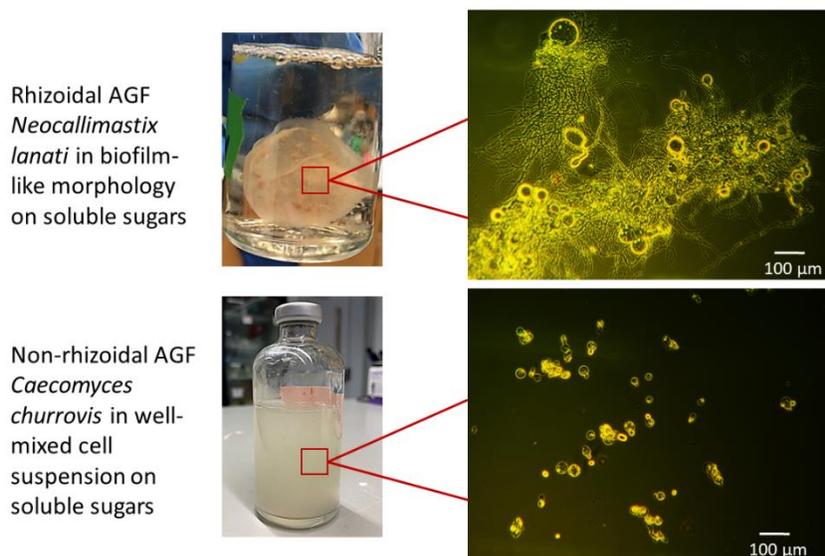


**Supplementary Figure S1.** A) More unannotated *de novo* transcripts show sequence similarity clustering than unannotated proteins in the *C. churrovis* genome. B) However, the percent sequence similarity is lower within *de novo* transcript clusters. More unannotated upregulated proteins cluster than downregulated proteins, however the function of these proteins remains uncertain.

**Supplementary Spreadsheet S1.** Spreadsheet of all *C. churrovis de novo* transcripts with CAZyme annotation and sequence. Available online (Leggieri et al., 2022).

**Supplementary Spreadsheet S2.** Master spreadsheet of all *C. churrovis* genes and *de novo* transcripts with expression metrics, differential expression statistics, functional annotation, CAZyme annotation, secretion, and sequence. Available online (Leggieri et al., 2022).

## B. Appendix: Supplementary Materials for Chapter III



**Supplementary Figure S2.** Images (left) and micrographs (right) of rhizoidal AGF *N. lanati* (top) in biofilm-like morphology and non-rhizoidal AGF *C. churrovis* (bottom) in well-mixed cell suspension. Both cultures shown here grown in Medium B on soluble sugars. The *C. churrovis* culture is amenable to growth tracking via optical density of small culture samples, while the *N. lanati* culture is not.

**Supplementary Table S2.** Comparison of *C. churrovis* metabolic EC numbers with the rest of the AGF phylum (Neocallimastigomycota) and with *N. lanati* in particular shows significant similarity in metabolic potential between *C. churrovis* and the rest of the AGF. Bottom row represents the total EC numbers present only in the indicated organism, relative to all of that organism's EC numbers.

	<i>C. churrovis</i> only (no other AGF)	<i>C. churrovis</i> but not <i>N. lanati</i>	<i>N. lanati</i> but not <i>C. churrovis</i>
Amino acids	1	5	6
Carbon	2	3	12
Glycans	2	1	0
Lipids	0	4	6
Nucleotides	1	2	4
Secondary met.	0	0	1
Terpenoids	2	2	0
Vitamins	3	3	3
Uncertain	11	33	41
Total	22	53	73
<b>Percent of total organism ECs</b>	<b>3.1%</b>	<b>7.6%</b>	<b>10.1%</b>

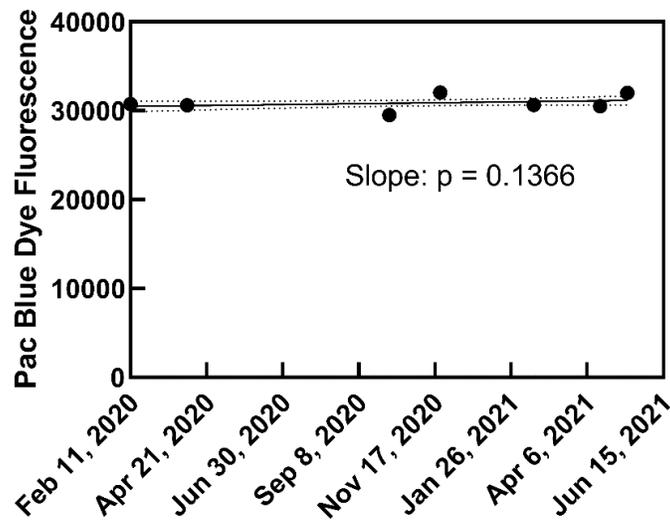
**Supplementary Table S3.** Mathematical workflow for calculating absorbance and associated uncertainty of each species (A: AGF; B: methanogen) from total co-culture fluorescence (F) and absorbance (Abs) signals.  $\varepsilon$  is the pure species absorbance per cell and  $\mathcal{F}$  is the pure species normalized fluorescence intensity per cell.

### Fluorescence

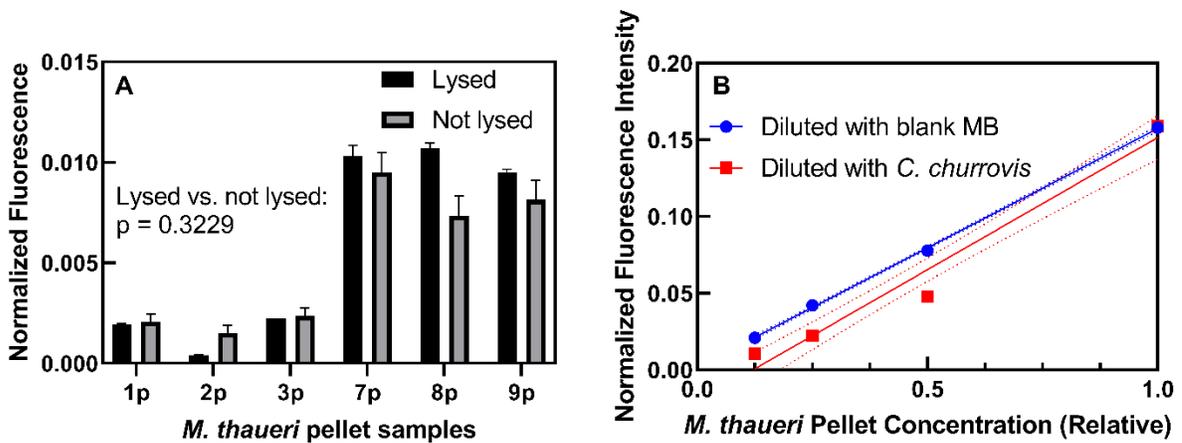
Values	Uncertainty
$F_{Sample} = \text{mean}(3 \text{ replicates} \times 2 \text{ wells each})$	$\sigma_{Sample (fluor)} = \text{stdev}(3 \text{ replicates} \times 2 \text{ wells each})$
$F_{Blank} = \text{mean}(3 \text{ replicates} \times 2 \text{ wells each})$	$\sigma_{Blank (fluor)} = \text{stdev}(3 \text{ replicates} \times 2 \text{ wells each})$
$F_{Dye} = \text{mean}(2 \text{ replicates} \times 2 \text{ wells each})$	$\sigma_{Dye (fluor)} = \text{stdev}(2 \text{ replicates} \times 2 \text{ wells each})$
$F_{Minus blank} = F_{Sample} - F_{Blank}$	$\sigma_{Minus blank (fluor)} = \sqrt{\sigma_{Sample (fluor)}^2 + \sigma_{Blank (fluor)}^2}$
$F_{Normalized} = \frac{F_{Minus blank}}{F_{Dye}}$	$\sigma_{Normalized (fluor)}$ $\approx \left  \frac{F_{Minus blank}}{F_{Dye}} \right  \sqrt{\left( \frac{\sigma_{Minus blank (fluor)}}{F_{Minus blank}} \right)^2 + \left( \frac{\sigma_{Dye}}{F_{Dye}} \right)^2}$

### Absorbance

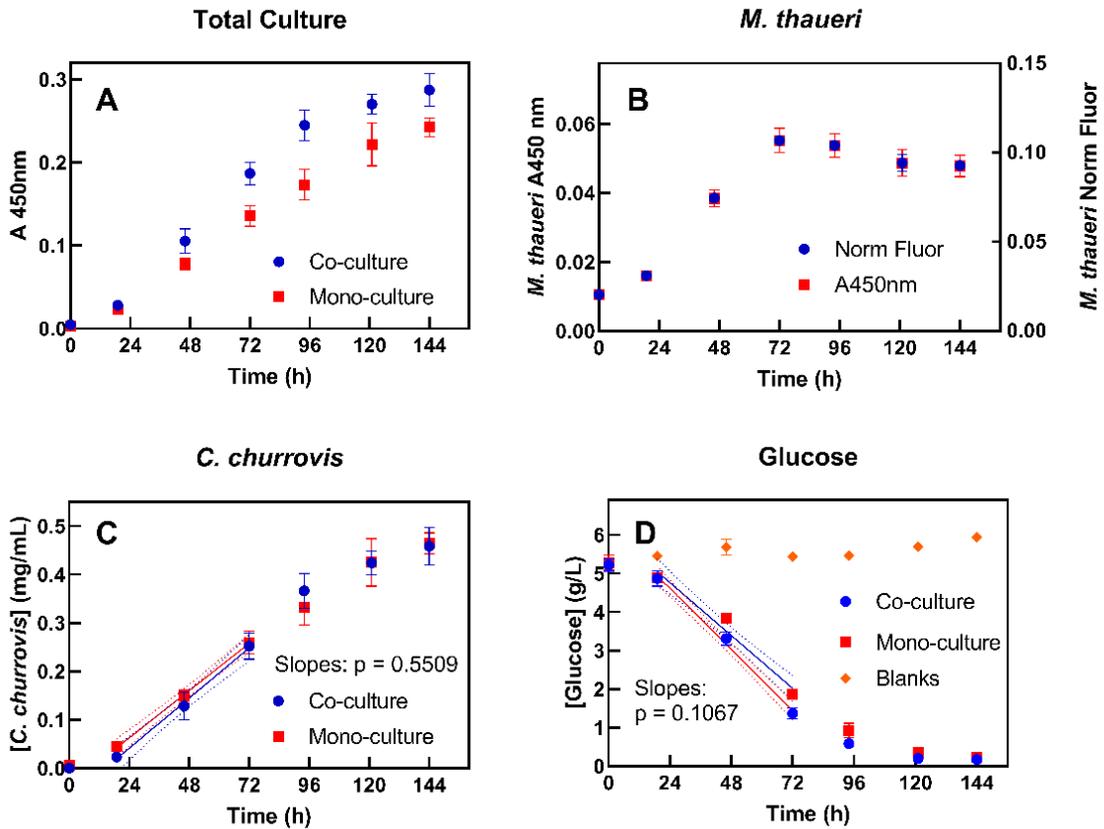
Values	Uncertainty
$Abs_{Sample} = \text{mean}(3 \text{ replicates} \times 2 \text{ wells each})$	$\sigma_{Sample (abs)} = \text{stdev}(3 \text{ replicates} \times 2 \text{ wells each})$
$Abs_{Blank} = \text{mean}(3 \text{ replicates} \times 2 \text{ wells each})$	$\sigma_{Blank (abs)} = \text{stdev}(3 \text{ replicates} \times 2 \text{ wells each})$
$Abs_{Minus blank} = Abs_{Sample} - Abs_{Blank}$	$\sigma_{Minus blank (abs)} = \sqrt{\sigma_{Sample (abs)}^2 + \sigma_{Blank (abs)}^2}$
$Abs_B = \frac{F_{Normalized}}{\frac{\mathcal{F}_B}{\varepsilon_B}}$	$\sigma_B (abs)$ $\approx \left  \frac{F_{Normalized}}{\frac{\mathcal{F}_B}{\varepsilon_B}} \right  \sqrt{\left( \frac{\sigma_{Normalized (fluor)}}{F_{Normalized}} \right)^2 + \left( \frac{\sigma_{\frac{\mathcal{F}_B}{\varepsilon_B}}}{\frac{\mathcal{F}_B}{\varepsilon_B}} \right)^2}$
$Abs_A = Abs_{Minus blank} - Abs_B$	$\sigma_A (abs) = \sqrt{\sigma_{Minus blank (abs)}^2 + \sigma_B^2 (abs)}$



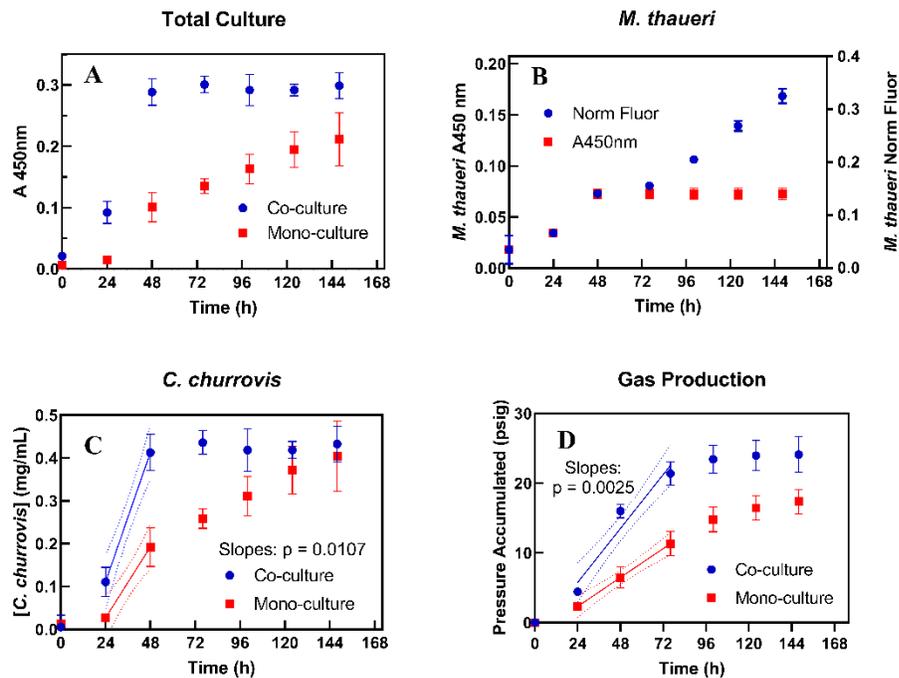
**Supplementary Figure S3.** The fluorescence intensity of aliquots of Pacific Blue dye in dimethyl sulfoxide (100  $\mu\text{g/L}$ ) stored at  $-20\text{ }^\circ\text{C}$  did not significantly change over 15 months of storage, indicating its utility as a standard for fluorescence normalization. The slope of the regression of fluorescence intensity vs. time (in weeks) is not significantly different from zero ( $p = 0.1366$ ), suggesting that fluorescence remains constant over the time period shown. Dotted lines represent the 95% confidence interval of the regression.



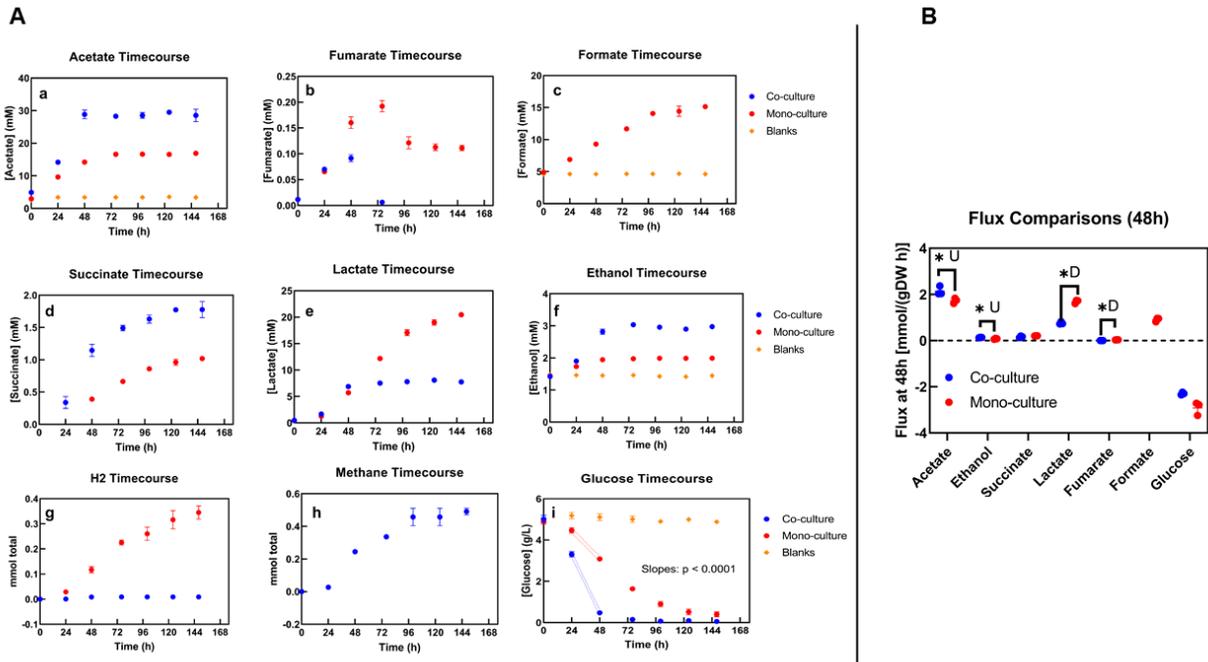
**Supplementary Figure S4.** The fluorescence intensity of *M. thaueri* cell pellets did not significantly increase when lysed according to the protocol outlined in (Peck, 1989) relative to unlysed (A) (paired t-test  $p = 0.3229$ ). The fluorescence intensity of *M. thaueri* pellets did not scale linearly with concentration when diluted with concentrated *C. churrovis* (B), suggesting that *C. churrovis* may interfere with the fluorescence of *M. thaueri* pellets, and the combined pellet and supernatant samples of co-cultures should be used to quantify methanogens in co-culture with AGF.



**Supplementary Figure S5.** Total culture absorbance (A), *M. thaueri* fluorescence + absorbance (B), *C. churrovis* concentration (C), and glucose consumption (D) curves from co-cultures inoculated with seven day-old *M. thaueri* culture. While methanogen growth occurred (B), neither the growth rate (C) nor glucose consumption rate (D) of *C. churrovis* increased in co-culture. The fluorescence of the methanogen does not diverge relative to the absorbance (B) because these slow-growing co-cultures did not fully reach stationary phase. All of these results differ from the glucose co-cultures presented in Additional File 6, in which a 48 hour-old *M. thaueri* culture was used for inoculation. Dotted lines represent the 95% confidence interval of each regression. The p-values in panels C and D represents a test for significant difference in the values of the slopes of the two regressions.

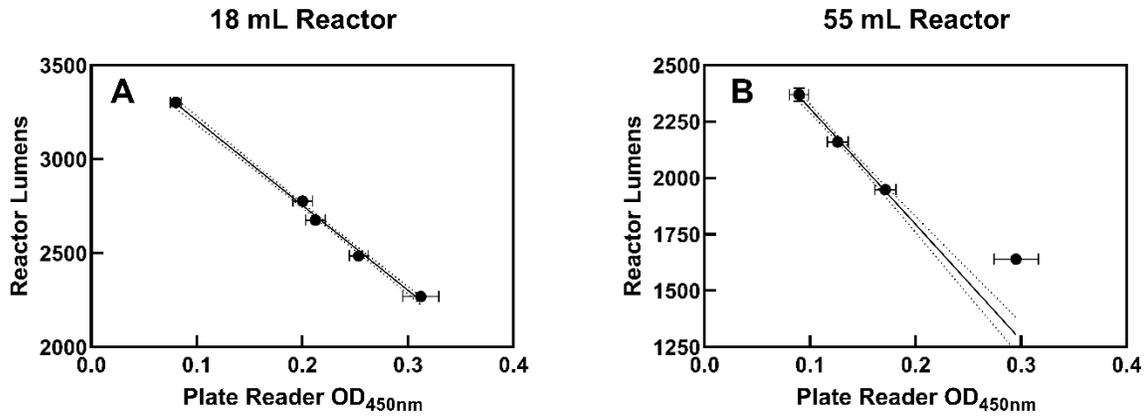


**Supplementary Figure S6.** Total culture absorbance (A), *M. thaueri* fluorescence + absorbance (B), *C. churrovis* concentration (C) and accumulated pressure (D) curves show that both the growth rate of *C. churrovis* and the rate of gas production are significantly increased in co-cultures with *M. thaueri* grown on glucose, relative to monocultures. Panel B shows the divergence of *M. thaueri* fluorescence relative to absorbance in stationary phase also observed in mono-culture; the absorbance of the methanogen was assumed to remain constant after the absorbance of the co-culture stops increasing (96h and after). Dotted lines represent the 95% confidence interval of each regression. The p-values in panels C and D represents a test for significant difference in the values of the slopes of the two regressions.

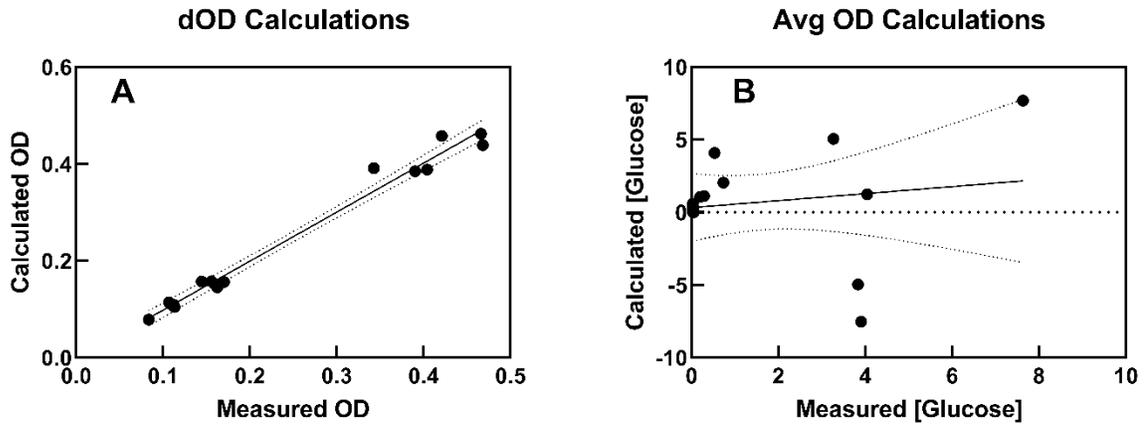


**Supplementary Figure S7.** Metabolite profiles (A) and cell mass-normalized fluxes (B) reveal significant upregulation (\*U) of acetate and ethanol fluxes, and significant downregulation (\*D) of lactate flux in co-cultures. Fumarate is an intermediate to succinate production, and it is consumed more quickly in co-cultures. Formate and hydrogen are consumed by *M. thaueri* and therefore do not accumulate in co-cultures. While glucose is consumed more quickly in co-culture, the flux of glucose into *C. churrovis* is equal in mono- and co-cultures. Dotted lines represent the 95% confidence interval of each regression. The p-value in panel (i) represents a test for significant difference in the values of the slopes of the two regressions.

C. Appendix: Supplementary Materials for Chapter IV



**Supplementary Figure S8)** Calibrations of the reactor's light sensor (lumens) against OD at 450 nm measured *via* plate reader. The 18 mL reactor configuration (A) showed linearity at all tested cell concentrations ( $R^2 = 0.9941$ ). The 55 mL reactor configuration (B) showed linearity at low to intermediate cell concentrations, with deviations at high cell concentrations (note the point near OD 0.30). Therefore, the 55 mL reactor was run with turbidostat control only up to OD setpoint 0.25. For higher titers in the 55 mL configuration, chemostat control was used. The calibration using only low to intermediate cell concentrations is sufficiently linear ( $R^2 = 0.9847$ ).



**Supplementary Figure S9)** The ratio of (change in *C. churrovius* OD)/(change in [glucose]) is constant, while the ratio of (average *C. churrovius* OD)/(glucose consumption rate) is not. To demonstrate, the former ratio (R1) was used to calculate the expected OD for batch cultures by dividing the measured change in [glucose] concentration between timepoints 1 and 2 by R1, and adding the result to the OD at timepoint 1. As shown in (A), the agreement is good ( $R^2 = 0.9830$ ), suggesting that there is a constant, reproducible ratio between glucose consumed and increase in OD, and that this ratio holds for all measured cell titers, even late in batch growth. In (B), the expected glucose concentration is calculated by dividing the average *C. churrovius* concentration between timepoints 1 and 2 by the latter ratio (R2), multiplying the time between measurements, and subtracting from the glucose concentration at timepoint 1. (B) shows no agreement between measured and calculated glucose concentrations, suggesting that there is no fixed per-cell rate of glucose consumption in *C. churrovius*; this rate varies over the course of batch growth. The linearity of (A) and lack thereof in (B) suggest that the change in AGF concentration between two timepoints, not the average, should be used to normalize flux measurements in batch culture.

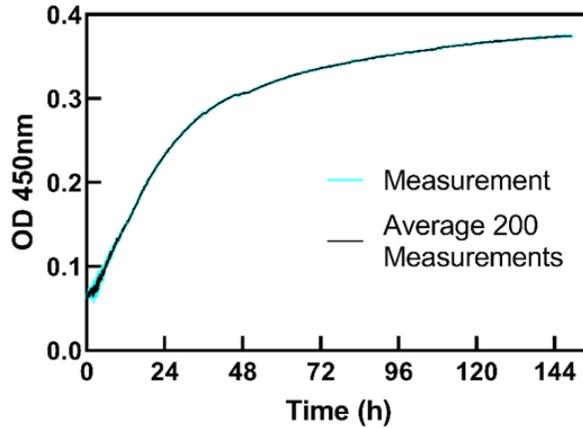
**Supplementary Table S4) Parts list with costs and hookup guide for Arduino-based bioreactor.**

Part	Connection to Arduino	Link	Cost
Blue LED (+)	2	<a href="https://a.co/d/gF5o5mZ">https://a.co/d/gF5o5mZ</a>	\$ 5.49
Blue LED (-)	GND		
TSL2591 Light Sensor (SDA)	20	<a href="https://www.adafruit.com/product/1980">https://www.adafruit.com/product/1980</a>	\$ 6.95
TSL2591 Light Sensor (SCL)	21		
TSL2591 Light Sensor (VCC)	5V		
TSL2591 Light Sensor (GND)	GND		
Pump Motor Driver (In) (AIN1)	13		
Pump Motor Driver (In) (AIN2)	11	<a href="https://www.sparkfun.com/products/14451">https://www.sparkfun.com/products/14451</a>	\$ 13.50
Pump Motor Driver (In) (PWMA)	10		
Pump Motor Driver (Out) (BIN1)	12		
Pump Motor Driver (Out) (BIN2)	8		
Pump Motor Driver (Out) (PWMB)	6		
Pump Motor Driver (VCC)	5V		
Pump Motor Driver (STBY)	7		
Pump Motor Driver (VM)	9V, 2A DC (not Arduino)		
Pump Motor Driver (GND)	GND		
Stirrer Motor Driver (AIN1)	43		
Stirrer Motor Driver (AIN2)	41		
Stirrer Motor Driver (PWMA)	5		
Stirrer Motor Driver (VCC)	5V		
Stirrer Motor Driver (STBY)	37		
Stirrer Motor Driver (VM)	9V, 2A DC (not Arduino)		
Stirrer Motor Driver (GND)	GND		
XBee S2C Wireless Communicator (DIN)	24	<a href="https://www.adafruit.com/product/968">https://www.adafruit.com/product/968</a>	\$ 22.95 (x2)
XBee S2C Wireless Communicator (DOOUT)	22		
XBee S2C Wireless Communicator (VCC)	5V		
XBee S2C Wireless Communicator (GND)	GND		
XBee S2C Wireless Communicator (Receiver)	Connected to PC via USB		
Peristaltic Pumps	Through motor driver	<a href="https://a.co/d/jj9tvF1">https://a.co/d/jj9tvF1</a>	\$ 11.98 (x2)
Magnetic Stirrer Motor	Through motor driver	Taken from <a href="https://www.vernier.com/product/stir-station/">https://www.vernier.com/product/stir-station/</a>	\$ 139.00
Arduino Mega 2560	-	<a href="https://store-usa.arduino.cc/products/arduino-mega-2560-rev3">https://store-usa.arduino.cc/products/arduino-mega-2560-rev3</a>	\$ 38.72
Heat Sink for LED	-	<a href="https://a.co/d/FP2YRqa">https://a.co/d/FP2YRqa</a>	\$ 7.99
Tubing	-	<a href="https://a.co/d/5FQI8RL">https://a.co/d/5FQI8RL</a>	\$ 6.99
Fubing Fittings to Connect Needles	-	<a href="https://www.fishersci.com/shop/products/luer-w-locknut-m-1-16-nyl-25pk/NC1108514">https://www.fishersci.com/shop/products/luer-w-locknut-m-1-16-nyl-25pk/NC1108514</a>	\$ 29.95
Total cost			\$ 331.95

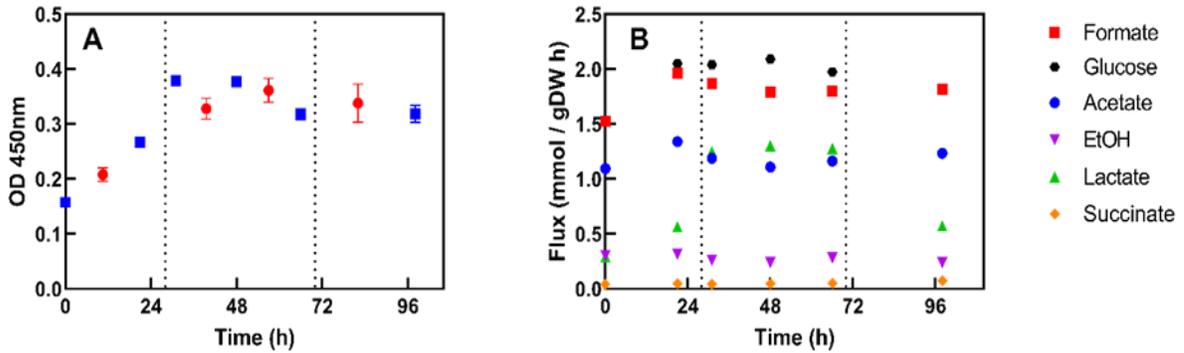
**Supplementary Script S1)** Representative Arduino IDE script for bioreactor turbidostat operation. This particular script is designed to operate the 18 mL configuration at high titer with turbidostat control. Available in online version of article.

**Supplementary Script S2)** Representative Arduino IDE script for bioreactor operation. This particular script is designed to operate the 55 mL configuration at high titer with chemostat control with a fixed dilution rate. Pumps were calibrated offline; the known flow rate is used to set the dilution rate.

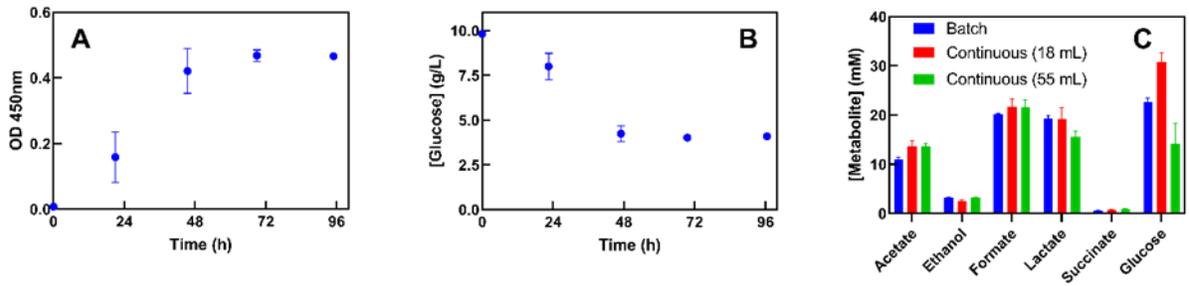
### *C. churrovis* Batch Growth



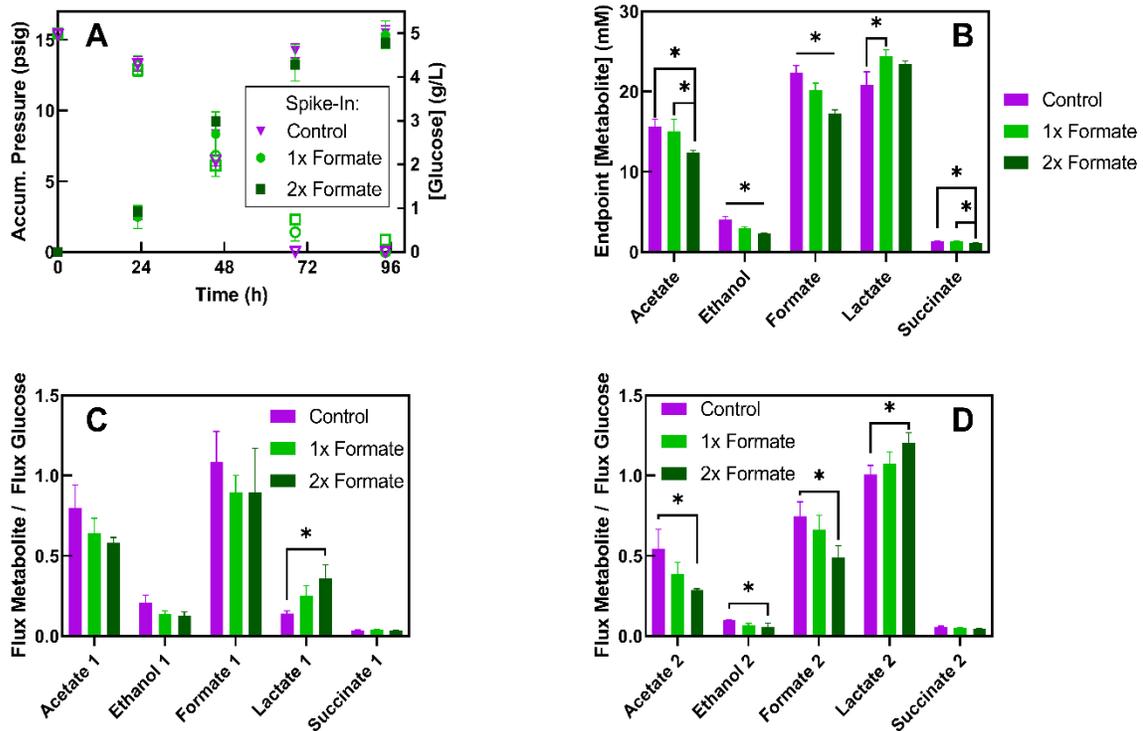
**Supplementary Figure S10)** Representative high-resolution growth curve for *C. churrovis* in 18 mL reactor configuration. Black line is moving average of 200 measurements taken every 300 ms, blue shadow shows every measurement with no averaging.



**Supplementary Figure S11)** High-titer chemostat operation showing constant OD (A) and constant metabolite fluxes (B) in the 55 mL reactor configuration. Blue squares in (A) represent OD measured with a single sample at that timepoint. Red circles in (A) represent measurements using the combined reactor effluent over several hours. The light path length in the 55 mL configuration is too long for linear detection of OD at this high titer setpoint, so chemostat control is required. The operating flow rate was 3.2 mL/h, corresponding to a dilution rate of 0.063 h<sup>-1</sup>. The high-titer setpoint is reached by choosing this slow dilution rate; *C. churrovis* concentration increases over the first 28 hours of operation at this dilution rate until a titer is reached that exhibits a growth rate matching the setpoint dilution rate. At this point, 28h (dotted line), OD remained constant for over 3 days. Metabolite fluxes were constant over this time with the exception lactate, which decreased after 72 h. It is not known why lactate flux alone decreased, however lactate flux appears to be the most variable in *C. churrovis* in continuous culture. Metabolic steady state is only assumed between 28 h and 70 h.



**Supplementary Figure S12)** Growth stops in *C. churrovis* batch cultures once they reach concentrations near OD 0.40 (A), even though substantial substrate remains (B). Continuous cultures are able to grow for several days near this OD (setpoint 0.35) in both turbidostat (18 mL) and chemostat (55 mL) configurations. The operating metabolite concentrations in continuous cultures are comparable to the endpoint concentrations in batch (C), so metabolite buildup alone is likely not responsible for growth stopping shy of complete substrate utilization in batch culture.

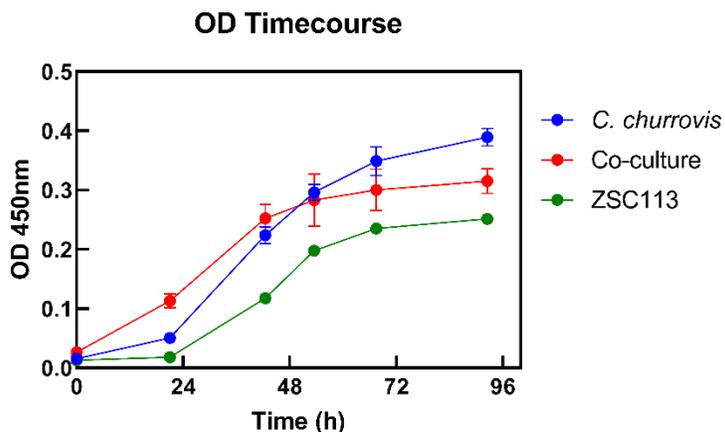


**Supplementary Figure S13)** Spike-in of formate during inoculation at 1x and 2x typical batch endpoint concentration (25 mM and 50 mM, respectively) had no effect on growth rate, yield, or glucose consumption (A) in *N. lanati*; solid shapes represent culture accumulated pressure (growth) and hollow shapes represent [glucose]. Formate spike-in significantly increased day 1 lactate flux relative to control (C). Day 2 lactate flux increased with formate spike-in; day 2 ethanol acetate, ethanol, and formate flux decreased with formate spike-in (D). Formate spike-in yielded significantly greater lactate concentrations, significantly lower acetate and succinate concentrations, and significantly lower ethanol, and formate concentrations with a dose-response relationship (B). Results align with those observed in *C. churrovis*, suggesting that formate may be used to control fluxes in AGF at large.

#### D. Appendix: Co-cultures of *C. churrovis* and *E. coli* ZSC113

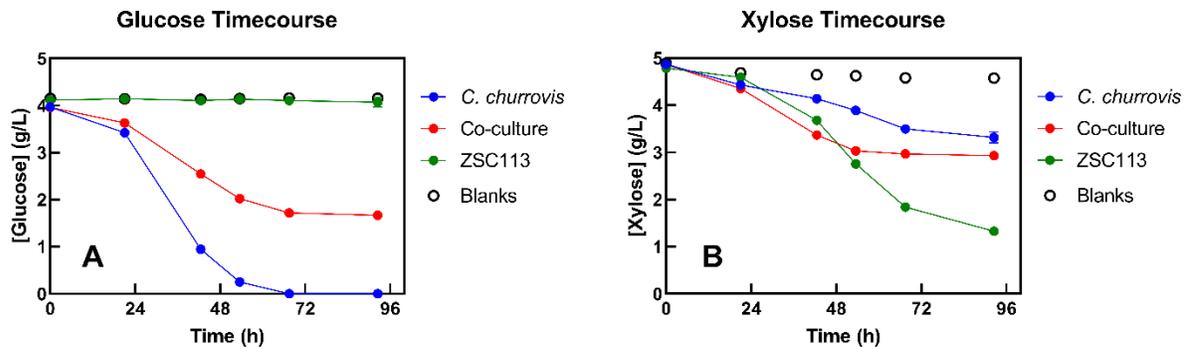
AGF liberate many different sugars from plant biomass (Henske et al., 2018b), but they are genetically intractable, precluding metabolic engineering efforts to produce designer biomolecules directly from lignocellulose using AGF alone. However, a workhorse strain of a genetically tractable model microbe, such as *E. coli*, could be co-cultured alongside AGF for bioproduction provided that the two species would not compete over sugars.

*E. coli* strain ZSC113 consumes xylose, but not glucose (Eiteman et al., 2008). Therefore, if AGF release both xylose and glucose from biomass, ZSC113 and AGF should not compete for sugars provided that AGF preferentially utilize glucose when both glucose and xylose are available. Further, the method presented in Chapter III of this dissertation is, in principle, extendable to quantify co-cultures of *C. churrovis* and non-fluorescent prokaryotes, enabling growth and flux tracking. To test the feasibility of co-culturing *C. churrovis* with ZSC113 on plant biomass for one-pot bioproduction, we assessed whether the two microbes could both grow in co-culture in defined Medium B with both glucose and xylose available.



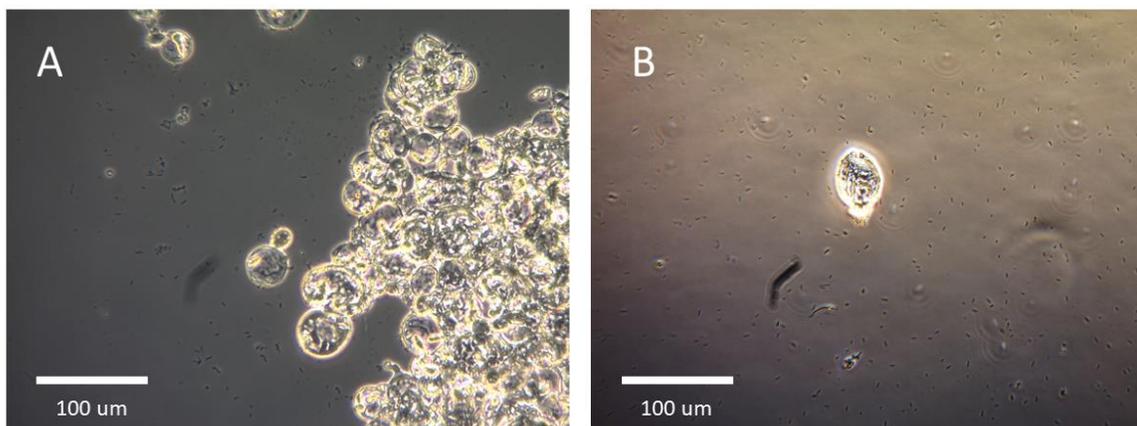
**Supplementary Figure S14.** OD curves of *C. churrovis* (blue), ZSC113 (green) and co-cultures of *C. churrovis* and ZSC113 (red) grown in Medium B with both glucose and xylose.

As shown in **Supplementary Figure S14**, *C. churrovis* and ZSC113 both grew in mono-culture in Medium B with glucose and xylose. Co-cultures showed growth of at least one species. **Supplementary Figure S15** shows that, in mono-culture, *C. churrovis* preferentially utilized glucose when both glucose and xylose were available, suggesting that *C. churrovis* and ZSC113 should not compete for substrates in co-culture. Mono-cultures of ZSC113 consumed xylose, but not glucose, as expected. In co-cultures, both glucose and xylose were partially consumed, indicating growth of at least *C. churrovis*.



**Supplementary Figure S15** Timecourses of glucose (A) and xylose (B) in cultures of *C. churrovis* (blue), ZSC113 (green) and co-cultures of *C. churrovis* and ZSC113 (red) grown in Medium B with both glucose and xylose.

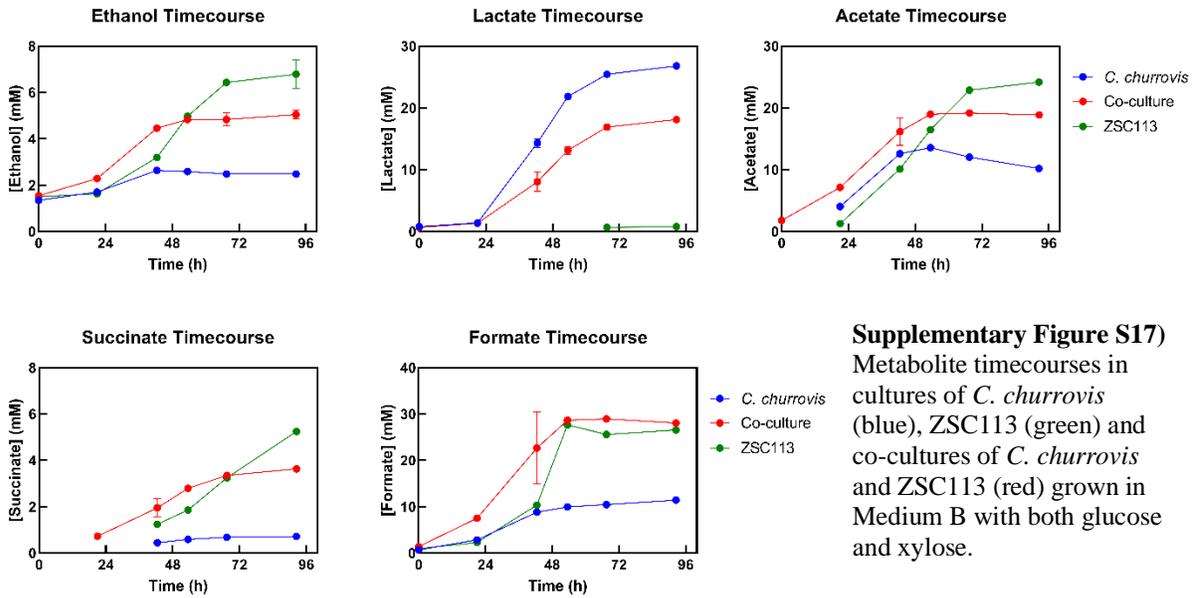
Micrographs of co-cultures show that, despite both species having substantial substrate available, only one species thrives in co-culture. Sometimes *C. churrovis* grew well (**Supplementary Figure S16A**), and sometimes ZSC113 grew well (**Supplementary Figure S16B**), but we did not observe any co-cultures with hearty growth of both species. We also attempted to co-culture *C. churrovis* and ZSC113 in Medium B with a plant biomass substrate and observed no growth of ZSC113.



**Supplementary Figure S16** Micrographs of co-cultures of *C. churrovis* and ZSC113 grown in Medium B with both glucose and xylose. (A) shows good growth of *C. churrovis*, but not ZSC113. (B) shows growth of ZSC113, but not *C. churrovis*.

It is possible that accumulation of fermentation products stunts growth in *C. churrovis* – ZSC113 co-cultures. The two species produce slightly different metabolite profiles (**Supplementary Figure S17**), so accumulation of products to higher titers than each species typically sees in mono-culture could result in reduced growth, however this hypothesis is unconfirmed.

Because co-cultures did not grow robustly, we could not definitively verify whether an augmented version of the method presented in Chapter III (two absorbance signals rather than one absorbance and one fluorescence signal) would work to quantify *C. churrovis* and ZSC113 in co-culture. However, preliminary results suggest that the average size and shape of *C. churrovis* changes too much over the course of growth to precisely quantify both species using two different absorbance wavelengths.



The negative results presented here underscore the complexity of microbial co-cultures and the challenges related to designing community systems for bioproduction. All mono-culture data suggests that these two species should grow together in co-culture; they grow in the same media, and they utilize different substrates. However, in practice, they could not coexist. Perhaps it is an interesting science question to determine exactly why. In an engineering context, this system is not capable of the application we designed it for, so we should look for other routes of AGF-based bioproduction from lignocellulose.

**E. Appendix: Proteomic analysis of *C. churrovis* cultures and hydrogenosome enrichments and *N. lanati* cultures and zoospore enrichments**

Many different pathways appear to be transcribed in the AGF hydrogenosome, however it is not clear which of these pathways carry significant flux, and whether pathway activity changes over the course of batch growth. To determine which pathways may be the most active in AGF hydrogenosomes, we attempted to quantify proteins in hydrogenosome-enriched lysate pellets from *C. churrovis* and zoospores (which may have greater hydrogenosome densities than sporangia (Yarlett et al., 1986)) enriched from *N. lanati* cultures.

Total culture proteomic samples of *C. churrovis* and *N. lanati*, with no enrichment of hydrogenosomes or zoospores, were analyzed for comparison, to determine whether putative hydrogenosome proteins were overrepresented in the samples that were expected to have greater hydrogenosome densities. As for hydrogenosome enrichments from *C. churrovis*, this was not the case. Further, we compared protein counts in samples from *C. churrovis* cultures at several different points of batch growth to assess whether enzyme counts aligned with observed changes in flux; they did not. These two shortcomings highlight the insufficient sensitivity of global proteomics to address these types of questions.

Targeted proteomics, with peptide standards, could provide better quantification of hydrogenosome proteins, and would eliminate any enrichment requirements. However, even targeted proteomics cannot identify proteins that are not annotated in the genome, as proteomic results are aligned to our reference genomes. As of today, no AGF genomes are of sufficient quality to fully understand the AGF hydrogenosome, motivating resequencing or at least better functional annotation.

Protocols for zoospore enrichment with quantification and hydrogenosome enrichment with validation are provided below, followed by representative validation data. Proteomic results collected by PNNL and processed by PA Leggieri are available in O'Malley Lab data archives.

Protocol: Zoospore Enrichment from *N. lanati*

- Grow *N. lanati* in any suitable medium with a soluble carbon source. Plant debris makes zoospore enrichment intractable.
- Harvest cultures by pouring them over a double layer of miracloth in a funnel, collect the filtrate. This step, and all others, can be done aerobically.
- Centrifuge the filtrate for 5 minutes at 2,500 g, 4 °C. Remove the supernatant, resuspend the pellet in 1 mL of media (no substrate necessary) or PBS.
- Centrifuge for 5 minutes at 2,500 g, 4 °C. Remove the supernatant.
- Resuspend the pellet in 1 mL of media (no substrate necessary) or PBS. Use 10 µL of this suspension to quantify zoospores *via* hemocytometry.
- Centrifuge for 5 minutes at 2,500 g, 4 °C. Remove the supernatant, store the pellet at -80 °C.
- To estimate the mass of each zoospore sample for proteomic sample preparation, use the zoospore counts obtained *via* hemocytometry, and assume each zoospore is a sphere of 10 µm diameter with roughly the same density as water.

Protocol: Hydrogenosome Enrichment from *C. churrovis*

- In chamber, transfer culture(s) to 15 mL centrifuge tube(s) and close lid.
- Remove centrifuge tubes from chamber, spin them 5 minutes, 2500 g, 4 °C to pellet.

- In chamber, remove supernatant, transfer cell pellet to 1.5 or 2mL microcentrifuge tube.
- Wash pellet with MB salts + water solution, PBS, or M2 twice.
- Apply chitinase solution to pellet. Vortex, incubate at 39 °C for 1 hour.
- After incubation, centrifuge in chamber for 5 minutes, 2500 g. Wash, repeat.
- Remove supernatant, transfer to Dounce homogenizer with 1 mL “hydrogenosome lysis buffer solution.”
- Homogenize, 25 strokes with rotation.
- Transfer homogenate to a new 1.5 mL microcentrifuge tube.
- Spin 5 minutes, 400 g in chamber.
- Remove supernatant and keep it, label it as **S1**. Set the pellet aside (**P1**).
- Spin **S1** 5000g, 20 minutes in chamber. There should be a visible pellet after this step. Remove the supernatant, set it aside (**S2**). Wash the pellet, then repeat the spin.
- Discard the supernatant from the second spin above. Keep the pellet (**P2**); this pellet should contain the most hydrogenosomes.
- **P2** is the fraction that can be used for further purification via ultracentrifugation.

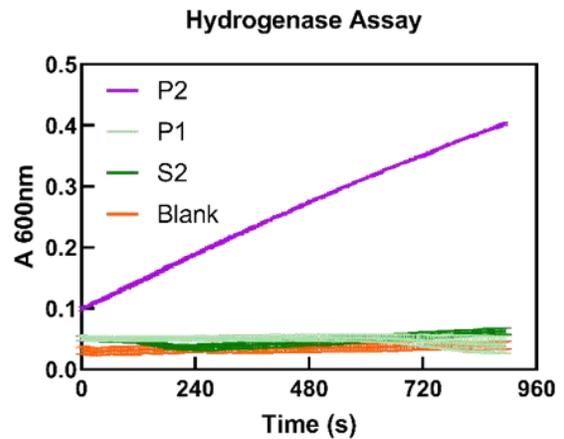
#### Protocol: Hydrogenase Assay

- For each fraction (S1, S2, P1, P2), use the same volume of methyl viologen working solution, same total sample volume, and same microwell volume. Apply the methyl viologen quickly to each well immediately before starting the absorbance run.
- Using the Absorbance96 software, plug the mini plate reader into the laptop, and set up your run to measure a 20 minute timecourse of absorbance at 600nm for all wells.

- Immediately after applying methyl viologen solution to each well, place the microplate in the plate reader in the incubator in the chamber. Start the run.
- Keep in mind, if you want to further purify **P2**, you only want to use a small fraction of your pellet for this assay.

### Reagents

- Hydrogenosome lysis buffer solution
- MB salts + water (wash buffer)
- Methyl viologen working solution
  - 200 mM phosphate buffer
  - Beta-mercaptoethanol
  - Methyl viologen



**Supplementary Figure S18)** Example hydrogenase assay on a successful enrichment (PL March 2022)

## **F. Appendix: Live-cell AGF Microscopy**

Given the temperature and oxygen sensitivities of AGF, a resourceful approach is required to capture microscopic videos of their growth. The approach described below applies to all strains of AGF. Media and substrates can be varied as desired.

### Protocol: Live-cell AGF Microscopy

- Place the light microscope (Zeiss Primovert) with camera (SPOT Idea 28.2) inside a 2418 incubator set to 39 °C, with the middle rack removed. The microscope must be inserted at an angle to fit. Run the USB cable from the camera to the bottom right corner of the glass screen and close it. Plug that cable into a laptop near the incubator.
- Use the SPOT imaging software to capture videos; the advanced version is preferable to the basic version because it displays the last captured image in a sequence, enabling real time monitoring. Capture images as TIFFs to avoid quality loss. Set the frequency and duration of image capture based on the goals of the experiment. Set exposure manually to achieve the best possible image quality. Remember that the focus through the eyepiece is different than the focus through the camera, so set fine focus based on the camera.
- Long image sequences result in immense file sizes, so set the software to store all images (even the temporary cache) on an external solid-state drive with at least 50 GB of free memory. Move or delete runs from the drive after analysis is complete.

- Grow AGF within sealable cuvettes from FireflySci. We have several different cuvettes with 1-2mm lightpaths, and PTFE plug or threaded cap stoppers.
- Cleaning the cuvettes is somewhat challenging, but soaking in 10% HCl then using a needle and syringe to carefully rinse the inside several times with DI water works well. Be careful not to scratch the cuvettes with the needles.
- Any typical growth medium or substrate should work. I used Medium B and glucose in my work. Plant substrates should work too and could provide interesting results.
- Inoculate the cuvettes in the anaerobic chamber. No zoospore enrichment is necessary; a healthy, actively growing culture contains more than enough zoospores to inoculate using only a small sample of culture supernatant.
- The cuvettes with stoppers do not lie flat, so it is prudent to create a small housing for the cuvette using tape, plastic, or whatever is available to provide a flat, fixed lightpath for the cuvette as it sits on the microscope stage.
- Since AGF produce pressure, bubbles are very common, and they ruin timelapse videos. This remains the main challenge in capturing live-cell AGF videos using this method.

- I had small, airtight, translucent boxes made to fit 96-well or other plates for live-cell imaging. The acrylic is too thick to image using the 10X objective, but 4X and 20X seem to work. You'll still want to seal the plate using a translucent seal to prevent evaporation within the box. Even with 384-well plates, our current microscope/camera setup likely only allows imaging of one well at a time. Bubbles may not be as much of an issue with this system, but condensation could be.