Genetic tools to optimize lignocellulose conversion in anaerobic fungi and interrogate their genomes

Ethan Hillman[1,2*] (ehillman@purdue.edu), Casey Hooker[1,3*] (hookerc@purdue.edu), and **Kevin Solomon[1,2,3]**

[1]Department of Agricultural and Biological Engineering, Purdue University,West Lafayette, IN; [2]Purdue University Interdisciplinary Life Sciences (PULSe) Program, Purdue University, West Lafayette, IN; [3]Laboratory of Renewable Resources Engineering (LORRE), Purdue University, West Lafayette, IN

http://solomonlab.weebly.com

Project Goals: This project develops genetic and epigenetic tools for emerging model anaerobic fungi to identify the genomic determinants of their powerful biomass-degrading capabilities, facilitate their study, and enable direct fungal conversion of untreated lignocellulose to bioproducts.

Deconstruction of plant cell wall biomass is a significant bottleneck to the production of affordable biofuels and bioproducts. Anaerobic fungi (*Neocallimastigomycota*) from the digestive tracts of large herbivores, however, have evolved unique abilities to degrade untreated fiber-rich plant biomass by combining hydrolytic strategies from the bacterial and fungal kingdoms[1]. Anaerobic fungi secrete the largest known diversity of lignocellulolytic carbohydrate active enzymes (CAZymes) in the fungal kingdom (>300 CAZymes), which unaided can degrade up to 60% of the ingested plant material within the animal digestive tract[2,3]. Unlike many other fungal systems, these CAZymes are tightly regulated and assembled in fungal cellulosomes to synergistically degrade plant material, including untreated agricultural residues, bioenergy crops, and woody biomass, with comparable efficiency regardless of composition[1,4–6]. However, the specific role of individual enzymes in maintaining hydrolytic efficiency remains unknown due to a lack of genetic tools that facilitate testing of gene function in its natural context. Thus, there is a critical need to create methods that manipulate CAZyme expression and rapidly interrogate gene function in anaerobic fungi to identify targets that will advance biofuel and bioproducts production.

In this project we study three novel specimens of anaerobic fungi isolated from giraffes, wildebeest, and donkeys. Initial characterization of these species confirm that they form distinct species from 2 genera of anaerobic fungi that exhibit high enzymatic activity against a range of untreated lignocellulolytic substrates regardless of lignin composition (e.g. GM poplar lines, sorghum, alfalfa, corn stover). Anaerobic fungi extensively tailor the relative abundance of secreted CAZymes to adapt to differences in substrate composition and achieve consistently high-levels of synergistic activity. To better understand this response and identify tools for genetic engineering, in the first phase of the project we are sequencing the transcriptomes of these isolates across various substrates and their genomes in partnership with the DOE-JGI. Using these resources, we intend to identify sequences such as inducible promoters and environmental conditions that regulate them to develop novel tools for gene expression. Our ongoing work with DOE resources such as MycoCosm has already validated constitutive promoters that can express

distinct genes in specified cellular compartments and alter cellular phenotypes of anaerobic fungi for the first time.

As anaerobic fungal genomes have extremely low GC contents (~15%) that impede accurate genome assembly, in parallel we are leveraging Hi-C (chromosomal conformation capture) sequencing technologies to improve sequence assembly and determine genomic structure to better infer gene function. While efforts are currently underway, early results have already identified a handful of assembly artifacts that were erroneously generated by conventional assembly algorithms. Ongoing work aims to generate and annotate the most accurate and high-resolution genomes of anaerobic fungi to date and determine ploidy, enabling more sophisticated genetics studies.

Anaerobic fungi display many classical hallmarks of epigenetic regulation within their genomes. Their genomes contain dozens of annotated histone proteins, and histone- and DNA- modifying proteins that are regulated with growth substrate[1,3,4]. This regulation is also strongly correlated with CAZyme expression in a gene-specific manner[4]. Our preliminary studies with the histone deacetylase inhibitor suberoylanilide hydroxamic acid (SAHA) corroborate this as we can globally modulate H3K4 and H3K27 trimethylation levels with an effective doubling in xylanase activity. To date we have investigated the effect of a panel of epigenetic inhibitors on anaerobic fungi establishing for the first time that anaerobic fungal CAZymes are epigenetically regulated. We are beginning to identify the epigenetic marks that are affected by these inhibitors via Western Blots and are working in partnership with PNNL EMSL and the DOE-JGI to establish the effects of specific epigenetic marks at a given loci on CAZyme expression and activity.

In summary, the ongoing work creates new tools to manipulate anaerobic fungi facilitating new opportunities for study and engineering. These approaches will be used to generate a deeper systems-level understanding of anaerobic fungal physiology while establishing fundamental knowledge about epigenetic regulation of gut fungal CAZymes. Ultimately, we enable predictive biology in anaerobic fungi and derive insight into microbial plant deconstruction to advance the development of economical biofuels and bioproducts.

References

1 Haitjema, C. H. *et al. Nature Microbiology* 2, 17087 (2017)
2 Seppälä, S. *et al. Metabolic Engineering* 44, 45–59 (2017)
3 Youssef, N. H. *et al. Appl. Environ. Microbiol.* 79, 4620–4634 (2013)
4 Solomon, K. V. *et al. Science* 351, 1192–1195 (2016)
5 Solomon, K. V. *et al. Fungal Genetics and Biology* 121, 1–9 (2018)
6 Hooker, C. A. *et al. Biotechnology for Biofuels* 11, 293 (2018)

This project is supported by the Office of Biological and Environmental Research through the DOE Office of Science (Early Career Program) Contract No DE-SC0020117. A portion of this research was performed under the Facilities Integrating Collaborations for User Science (FICUS) initiative and used resources at the DOE Joint Genome Institute and the Environmental Molecular Sciences Laboratory, which are DOE Office of Science User Facilities. Both facilities are sponsored by the Office of Biological and Environmental Research and operated under Contract Nos. DE-AC02-05CH11231 (JGI) and DE-AC05-76RL01830 (EMSL).