Identification of Novel Biomass Deconstructing Enzymes via Phylogenetic and Structural Analyses

Adam J. Book[1,2], Ragothaman M. Yennamalli[3], Taichi E. Takasuka[2,3], Gina R. Lewin[1,2] Brian G. Fox[2,3], George N. Phillips Jr. [2,3] and Cameron R. Currie[1,2]

1 Department of Bacteriology, University of Wisconsin-Madison
2 DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison
3 Department of Biochemistry, University of Wisconsin-Madison

Project Goals: Develop more accurate functional annotation models for cellulase enzyme families based on phylogenetic, structural, and enzymatic analyses.

The tremendous amount of -omic data (genomes, metagenomes, transcriptomes, and proteomes) has provided a great resource for the discovery of highly efficient or novel enzymes for biomass deconstruction. Work in the Currie Lab alone has provided >50 bacterial genomes and >10 metagenomes from lignocellulose-rich environments. The standard annotation processes for carbohydrate active enzymes (CAZymes) gives good general classification; however, high quality functional predictions are elusive because each CAZy superfamily contains multiple enzymatic functional groups (*e.g.* GH5 contains endoglucanases, mannanases and more). The overall goal of this project is to develop more accurate functional annotation models for cellulase based on phylogenetic, structural, and enzymatic analyses. To achieve this we have constructed phylogenetic trees for two major cellulase CAZy superfamilies, the highly populated GH5 superfamily, and the relatively uncharacterized CBM33 superfamily (oxygenase). When these trees are layered with protein structures (solved and modeled), gene expression data, and biochemically tested enzymatic information, predicted functional annotations can be assigned with much higher precision. These improved annotations will then be used to create prediction algorithms so that new sequences can be quickly and accurately classified into functional groups.

This work was funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-FC02-07ER64494).