Plant biomass-degrading loci play a role as determinants for lignocellulose degradation from the extremely thermophilic genus, *Caldicellulosiruptor*

Sara E. Blumer-Schuette[1, 4]* (seblumer@ncsu.edu), Richard J. Giannone,[2, 4] Jeffrey V. Zurawski,[1, 4] Inci Ozdemir,[1, 4] Dhaval B. Mistry,[1] Qin Ma,[3, 4] Yanbin Yin,[3, 4] Ying Xu,[3, 4] Farris L. Poole II,[3, 4] Irina Kataeva,[3, 4] Mike W. W. Adams,[3, 4] Scott D. Hamilton-Brehm,[2, 4] James G. Elkins, [2, 4] Frank W. Larimer,[2] Miriam L. Land,[2] Loren Hauser,[2] Robert W. Cottingham,[2, 4] Robert L. Hettich,[2. 4] **Robert M. Kelly**,[1, 4] and **Paul Gilna (BESC PI)**[4]

[1]North Carolina State University, Raleigh; [2]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee; [3]University of Georgia, Athens; [4]BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge, Tennessee

http://bioenergycenter.org

Project Goals: The BioEnergy Science Center (BESC) is focused on the fundamental understanding and elimination of biomass recalcitrance. BESC's approach to improve accessibility to the sugars within biomass involves 1) designing plant cell walls for rapid deconstruction and 2) developing multitalented microbes for converting plant biomass into biofuels in a single step (consolidated bioprocessing). BESC research in biomass deconstruction and conversion targets CBP by studying model organisms and thermophilic anaerobes to understand novel strategies and enzyme complexes for biomass deconstruction.

Consolidated bioprocessing (CBP) of second-generation biofuels from lignocellulose will require that the microbe(s) be able to deconstruct complex polysaccharides from plant biomass. The extremely thermophilic, non-cellulosomal genus *Caldicellulosiruptor* have come under renewed interest due to their plant biomass-degrading abilities at high temperatures (T_{opt}, 70~80°C). In order to fully identify the genetic diversity of the genus, eight *Caldicellulosiruptor* species, ranging from weakly to strongly cellulolytic were selected as a representative samples for whole genome sequencing (WGS). Comparative genomics using these eight genome sequences determined that the pan-genome for the genus *Caldicellulosiruptor* is not yet saturated. The *Caldicellulosiruptor* pan-genome (4,009 ORFs, based on eight species) encodes 92 glycoside hydrolases (GHs) representing 43 GH families, but only 25 GHs from 17 families are included in the core genome (1,543 ORFs). Three of these core CAZy-related genes are both multi-modular and extracellular and include a bi-functional cellobiohydrolase/ endo-xylanase, an □-amylase, and an endo-xylanase. Non-core ancestral glucan degrading and xylan degrading loci were identified based on protein homology and chromosomal location. Comparative analysis revealed that the key determinant for cellulolytic activity was not the number of GHs but rather the presence of a genomic locus encoding one or more novel multi-domain cellulases, coupled to a specific type of carbohydrate binding modules (CBM3). Weakly cellulolytic species have completely or partially lost this locus, specifically genes encoding for GH48 catalytic domains coupled with a CBM3 module. Furthermore, the cellulolytic *Caldicellulosiruptor* species produced adhesins (substrate-binding portions of type IV pili), encoded upstream of CBM3 gene clusters in the same locus, that were found to bind to crystalline cellulose. Overall, the use of comparative genomics has identified novel multi-modular CAZy-related enzymes, highlighted the importance of previously known and characterized cellulase loci and identified mechanisms for substrate attachment. The genus *Caldicellulosiruptor*, by virtue of their plant biomass-degrading abilities make an attractive model system for CBP organisms, and with advances in genetic systems, can become engineered CBP biocatalysts as well.

This work was supported by a grant from the BioEnergy Science Center (BESC), a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.

[The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science.]