Title

Discovery of substrate-targeted enzymes for the degradation of biomass by Metatranscriptomics

Authors

Hess, Matthias
Zhang, Tao
Tringe, Susanna
et al.

Publication Date

2009-03-26

Discovery of substrate-targeted enzymes for the degradation of biomass by Metatranscriptomics

M.Hess*, T.Zhang*, S.G.Tringe*, R.Mackie+ and E.M.Rubin*

* DOE Joint Genome Institute, Walnut Creek, CA, 94598
+ University of Illinois, Urbana-Champaign, ILL, 61801

Introduction

Current bioethanol production from lignocellulosic biomass is a relatively inefficient process requiring energy-intensive chemical and physical pretreatments (Fig.1). The identification of robust cellulolytic enzymes (e.g. cellulases and hemicellulases) from lignocellulolytic microbial communities is a necessity to achieve the ambitious goal to replace 30% of the national petroleum based gasoline with bioethanol by 2030.

Switchgrass (*Panicum virgatum*), a perennial, warm-season grass is native to most states of the U.S. and has drawn a lot of attention as a promising biofuel crop. Foregut fermenters such as cattle and sheep possess an area within the gastrointestinal tract that is well separated from the acid-secreting portion of the stomach.

This fermentation chamber harbors a microbial community that is able to degrade the plant celluloses, hemicelluloses, pectins, fructosans, starches and other polysaccharides to monomeric and dimeric sugars. The bovine rumen exceeds a volume of 100 L and represents an enormous, easily accessible and manipulable system (Fig. 2) to study the microbial community and its biocatalysts adapted to the degradation of selected biofuel crops such as *P. virgatum.*

Metagenomics and Metatranscriptomics of the microbial community associated with biofuel crops will allow us to identify the microbes responsible for lignocellulose degradation and the genes that are overexpressed during fiber hydrolysis.

Sequence-independent enzyme activity assays will complement our quest for genes and proteins that are required for efficient lignocellulose degradation. An outline of our project is shown in Figure 3.

We hope that the findings of our project will provide significant insight into the process of lignocellulose degradation and facilitate the construction of genetically modified organisms (i.e. *Escherichia coli*, *Saccharomyces cerevisiae*, and *Sulfolobus solfataricus*) and ultimately the development of an efficient process to convert lignocellulosic biomass into bioethanol at industrial scale.

Industrial biofuel production from biomass

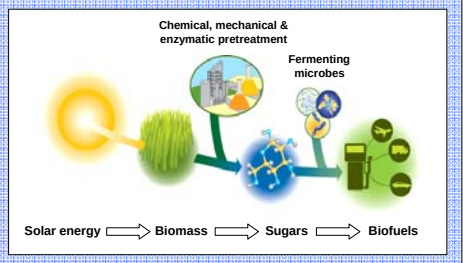

Fig. 1: Current industrial processes to convert biomass into biofuels are based on the conversion of substrates rich in starch, rather than of lignocellulosic material. The hydrolysis of complex sugars into monosaccharides requires an inefficient and expensive pretreatment step. More efficient enzymes from lignocellulolytic microbial communities are a promising avenue to render the complete process more efficient.

The fistulated cow: an excellent system to study lignocellulose degradation

Fig. 2: The cow is able to convert grasses and low growing plants rich in hemicelluloses, celluloses, pectins, fructosans, straches and other polysaccharides to monomeric and dimeric sugars. Conversion of the feedstock is mediated by the microbial community of the rumen which is part if the bovine digestive tract (Fig. 2A). A surgically inserted fistula and the rubber cannula guarantee permanent access to the rumen (~100 L) and its anaerobic microbial community (Fig. 2B). Large amounts of different substances can be inserted and removed from the rumen (Fig. 2C) at any time, permitting strictly controlled experiments.

Project outline

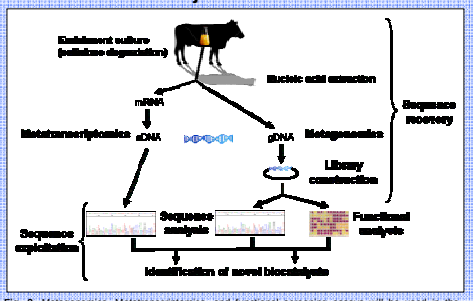

Fig. 3: Metagenomics, Metatranscriptomics, and functional screening assays will be used in the course of our project to identify the genes and proteins required for the efficient degradation of lignocellulose within the bovine rumen.

Switchgrass degradation in the rumen

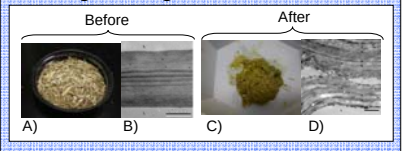

Fig. 4: Dried switchgrass (Fig. 4A & B) was incubated for 72 hrs in the rumen. Biochemical analysis showed that 26% of the initial fiber was degraded during this period. Figure 4C & D show the retrieved sample.

Microbes adherent to switchgrass

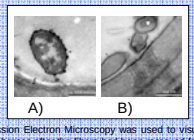

Fig. 5: Transmission Electron Microscopy was used to visualize microbes adherent to switchgrass after the fibers had been removed from the rumen

Results

Characterization of microbial rumen communities associated with different crops based on 16S ribosomal RNA clone libraries

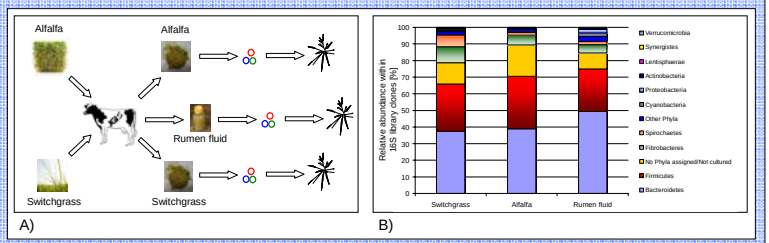

Fig. 6: 16S rRNA clone libraries were made for the communities associated with alfalfa, switchgrass and present in rumen fluid. The obtained almost-full length gene sequences were used to build phylogenetic trees (Fig. 6A) and to determine the community composition for the different substrates (Fig. 6B). Bacteroidetes and Firmicutes were the most abundant phyla. Relative abundance of Fibrobacteres and Synergistes increased in the switchgrass adherent community

Fingerprint of microbial communities associated with different crops

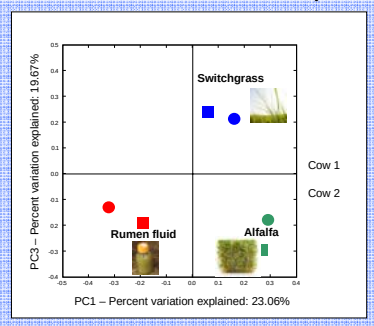

Fig. 7: The Phylogenetic trees determined from the 16S rRNA sequences were used for Principal Component Analysis. The results suggest that the composition of the microbial communities is substrate and not host dependent.

454-Pyrosequencing

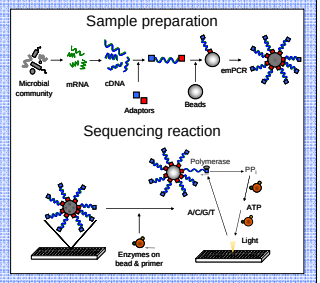

Fig. 8: Sample preparation: Total RNA was extracted from the microbial community associated with switchgrass. rRNA was partially and the remaining RNA was converted to cDNA. Sequencing specific adapters were linked to the double stranded cDNA. The construct attaches to a bead and was amplified by emulsion PCR. **Sequencing reaction:** One of the four nucleotides is flown over the sequencing plate and a inorganic pyrophosphate (PP$_i$) is released in case the polymerase incorporates the nucleotide. The PP$_i$ is converted to ATP and a luciferase uses the ATP to generate light, which can be detected by the sequencer. The cycle is iteratively repeated for each of the four bases.

Metatranscriptome of microbial community adherent to switchgrass

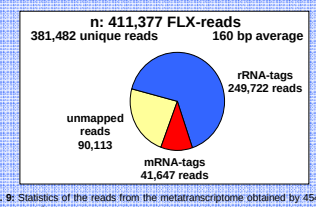

Fig. 9: Statistics of the reads from the metatranscriptome obtained by 454-FLX pyrosequencing

Gene expression distribution of the microbial community adherent to switchgrass

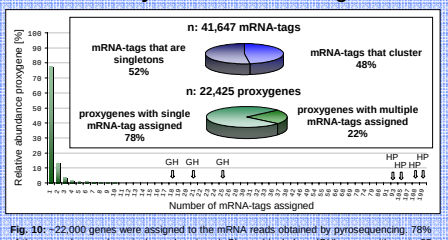

Fig. 10: ~22,000 genes were assigned to the mRNA reads obtained by pyrosequencing. 78% of the genes have only a single read assigned. Glycosyl-hydrolase (GH) genes with up to 25 assigned reads were identified in our data-set. The genes that had the most reads assigned were identified as hypothetical proteins (HP).

Functional COG classes detected in the metatranscriptome

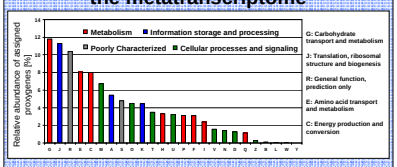

Fig. 11: Proxygenes and their assigned COGs were used to assigned function to the metatranscriptome. One letter code according to the NCBI COG database.

Identified glycosyl-hydrolases

Enzyme	# of distinct genes	# of distinct mRNA-reads
Glycosyl hydrolase	264	488
Cellulase	11	22
Xylanase	18	24
α-amylase	34	54

Summary

- Switchgrass is degraded within the bovine rumen
- Rumen microbes adherent to switchgrass were detected
- Composition of microbial populations within the rumen is substrate dependant
- Bacteroidetes, Firmicutes, and Fibrobacteres appear to be the dominant feedstock-associated phyla
- Relative abundance of Fibrobacteres and Spirochaetes is increased in the switchgrass associated population
- Transcripts of different expression levels were detected using 454-FLX pyrosequencing
- Most mRNA-tags are derived from unique genes
- Almost 500 transcripts of ~300 distinct glycosyl hydrolases were identified from the meta-transcriptome of switchgrass adherent rumen microbes

This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231, Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344, and Los Alamos National Laboratory under contract No. DE-AC02-06NA25396. TEM slides were prepared by Dr. Ding and Dr. Haas at NREL.