**Abstract**

Biogasification is a process that utilizes the microbial community native to coalbeds to naturally convert currently unused coal into readily usable methane. One methodology involves injection of nutrients into the coal seam to stimulate biogenic coal degradation and methanogenesis. Identification of major functional pathways of biogenic coal degradation and subsequent methane production will lead to a better understanding of the coal-to-methane conversion, the microorganisms responsible for this conversion, and the nutrients required to bolster this conversion in situ. This study examines the metagenome of four produced water samples from the Central Appalachian Basin (Pocahontas 3 coal seam) to determine the composition (who’s there) and the potential functional pathways (what can they do) of the resident microbial communities. Nucleic acid was recovered from produced water samples using DNA isolation techniques and the quality and quantity of DNA was assessed. Illumina MiSeq next generation sequencing was employed, and the resultant nucleic acid sequence data was processed using a suite of bioinformatic software. Four metagenomes named K34, K35, BB137, and L32A were obtained from produced water samples from a depth of 1704 ft., 1980 ft., and 2575 ft., respectively. Methanogens were present in all samples, suggesting methanogenesis may occur. Furthermore, hydrocarbon degradation pathways were found, suggesting a route for biodegradation of coal. Importantly, a draft genome most closely related to Pseudomonas stutzeri CCUG was extracted from the K35 metagenome. Initial analysis of the draft genome revealed a complete nitrogen fixation pathway, and a naphthalene degradation pathway.

**Goals:**

- Investigate the microbial community in potential biogasification sites
- Characterize relevant functional pathways found in coal systems required for coal-to-methane biocconversion
- Construct draft genomes of abundant microorganisms in coal systems to complete a detailed characterization of prevalent functional potential

**Methods: Metagenome Analysis Pipeline (MAP)**

The first step in the metagenome analysis pipeline (MAP) involves careful and calculated sample collection. Samples are collected by drillers, or when possible, data site by NETL researchers. Importantly, to preserve sample integrity and prevent nucleic acid degradation, samples are immediately aliquoted and frozen. After transport from the field, the samples are thawed and processed for DNA extraction. The quality and quantity of DNA is assessed before preparing samples for sequencing.

The second step in the MAP involves processing DNA samples to generate a sequence library to be loaded into the sequencer. Illumina MiSeq. Processing involves cleaning, barcoding, and pooling DNA samples. The third MAP step is the most time-consuming and computationally intensive. Here, data that is retrieved from the sequencer is processed and analyzed. Processing involves removing barcodes and trimming reads based upon the quality score (a measure of the confidence of each base call). Analysis involves metagenome assembly, binning, and annotation.

**Genome Analysis**

**Contig mapping to reference genome sequence**

Pan-genome tree comparing the draft genome of P. stutzeri K35 to 31 complete and draft Pseudomonas genomes. The tree represents the degree of similarity between the predicted proteins encoded by each genome.

**Overview of functional potential**

Pan-genome Analysis

Genome Results: The K35 metagenome was estimated to be ~50% Pseudomonas. After careful contig binning and genome mapping, the Pseudomonas genome bin was 99.2% complete, estimated by the presence/absence of 833 marker genes. This genome bin was determined and the core genome (genes common across all strains tested) was estimated. The draft genome encodes for a complete nitrogen fixation pathway as well as the upper and lower naphthalene degradation pathways. The work presented here represents an initial metagenome/genome approach to functional characterization of coalbed methane microbial communities.

**Genome Binning**

**Contigs from the K35 metagenome assembly were binned. Each bar represents a contig and each cluster represents a potential genome.**

**K34** 1704 ft.

**K35** 1912 ft.

**BB137** 1980 ft.

**L32A** 2578 ft.

**Acknowledgements**

This research was funded by the Department of Energy, National Energy Technology Laboratory, an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor AECOM, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof.