



## DESTINATIONS ON THE METAGENOMIC HIGHWAY: ANAEROBIC HYDROCARBON BIODEGRADATION

Julia Foght; Biological Sciences, University of Alberta  
Edmonton AB Canada T6G 2E9, julia.foght@ualberta.ca

*Key words: environment, biodegradation, metagenome*

**Introduction.** The Hydrocarbon Metagenomics Project (HMP; [www.hydrocarbonmetagenomics.com](http://www.hydrocarbonmetagenomics.com)) has been applying metagenomic approaches to fossil fuel exploitation in Canada, including oil sands recovery (through both surface mining and *in situ* extraction), coal bed methane resources and conventional oil reservoirs, as well as examining downstream consequences like tailings ponds and pipeline corrosion. Diverse microbial communities reside within these resources, commonly supported by methanogenic hydrocarbon biodegradation. By understanding the fundamental biological processes operating in these environments, the HMP aspires to enhance hydrocarbon recovery while reducing environmental impacts. To achieve these goals we have sequenced ~2,000 16S rRNA gene amplicons, ~20 metagenomes and a metatranscriptome from environmental samples and anaerobic enrichment cultures; generated and begun screening fosmid libraries constructed from metagenomic DNA; performed single cell sequencing of environmental samples; and sequenced whole genomes of cultivated pure isolates. The overarching objective of the HMP is mitigating the environmental impacts of fossil fuel production in Canada by harnessing the natural genetic potential of microbes, genes and biological processes residing in the natural microbial communities associated with these energy resources.

**Methods.** The project has advanced in several stages: (1) amplification of partial 16S rRNA genes using 'universal' barcoded primers for subsequent pyrosequencing (454-FLX); (2) extraction of total genomic DNA for construction of large-insert fosmid libraries, which have been end-sequenced and screened *in silico*, as well as by PCR amplification of functional genes and phenotype screening; (3) metagenomic sequencing of environmental samples and enrichment cultures using 454 and illumina platforms; and recently (4) using microfluidics for single cell isolation and sequencing, as well as Ion Torrent whole genome sequencing of pure cultivated isolates.

**Results.** Sequencing 16S rRNA gene amplicons revealed considerable taxonomic diversity in the hydrocarbon resource environments. Despite having similar anaerobic metabolic pathways (i.e., sulfidogenic and methanogenic biodegradation of hydrocarbons to signature metabolites, methane and CO<sub>2</sub>), the bacterial 'core community' is relatively small, with diverse ancillary taxa even within transects from a single environment, suggesting functional redundancy. Phylogenetic analysis suggests that many bacterial 16S sequences represent new species or genera. In contrast, the Archaea are almost exclusively mesophilic methanogens similar to known species, with far less diversity and variability. Assembly of metagenomic sequences into large contigs has proven difficult, due in part to lack of sequenced isolates as scaffolds. Thus, we have been sequencing isolated single cells and pure cultivated isolates to generate genomic scaffolds. Functional fosmid screening, e.g., for biosurfactant production, heavy metal resistance and key enzyme activity, is underway with some successes.

**Conclusions.** Combining metagenomic sequencing with biochemical analyses of enrichment culture activities (e.g., substrate specificity under various terminal electron acceptors) and chemical properties of the corresponding resource environments (e.g., residual hydrocarbons and signature metabolites) is providing insight into *in situ* processes as well as suggesting ways to harness these microbial communities for 'greener' resource extraction.

**Acknowledgements.** The Hydrocarbon Metagenomics Project is supported by Genome Canada, Genome Alberta, the Government of Alberta, Genome BC, and numerous energy industry partners.