# Comparative Metagenomics of Microbial Communities from Pristine and Contaminated Groundwater

Christopher L. Hemme[1,2], Ye Deng[1], Terry J. Gentry[1,3], Liyou Wu[1,2], Matthew W. Fields[4], Susannah Green Tringe[5], Kerrie Barry[5], David B. Watson[1], Nikos Kyrpides[5], Chris Detter[5,6], Paul Richardson[5], Terry C. Hazen[7], James M. Tiedje[8], Eddy Rubin[5] and Jizhong Zhou[1,2]

Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN[1], Institute for Environmental Genomics, University of Oklahoma, Norman, OK[2], Department of Soil & Crop Sciences , Texas A&M University, College Station, TX[3], Department of Microbiology, Montana State University, Bozeman, MT[4], DOE Joint Genome Institute, Walnut Creek, CA[5], Los Alamos National Laboratory, Los Alamos, NM[6], Earth Science Division, Lawrence Berkeley National Laboratory, Berkeley, CA[7], Center for Microbial Ecology, Michigan State University, East Lansing, MI[8]

ESPP2 Environmental Stress Pathway Project

DOE GENOMICS:GTL ACCELERATING DISCOVERY FOR ENERGY AND ENVIRONMENT — OFFICE OF SCIENCE U.S. DEPARTMENT OF ENERGY

VIMSS Virtual Institute for Microbial Stress and Survival — Massachusetts Institute of Technology — Missouri — University of Washington — Montana State University — Southern Illinois University Carbondale — Oklahoma — Sandia National Laboratories — Oak Ridge National Laboratory — Berkeley — 75 Years of Science

http://vimss.lbl.gov/

## INTRODUCTION

Microbial community DNA from contaminated groundwater from the US Dept. of Energy Field Research Center (FRC) has recently been analyzed to determine the effects of multiple stressors on microbial community structure. The sample was obtained from a site (FW106) experiencing long-term (~50 yrs) exposure to high concentrations of uranium, nitric acid and organic solvents. Analysis indicates a very low diversity community dominated by denitrifying γ- and β-proteobacteria. Furthermore, metabolic reconstruction reveals adaptations for specific geochemical parameters including: denitrification pathways; pathways for degradation of organic molecules including 1,2-dichloroethene, acetone, butanol, methanol and formaldehyde; a large variety of heavy metal resistance systems (czcABC, czcD, cadA, merACPT, etc.). Furthermore, many of these adaptations appear to be the result of lateral gene transfer. Abundance profiles of FW106 compared to all sequenced bacterial isolates show that several geochemically-relevant transporter genes, including *narK* nitrate/nitrite antiporter and *czcABC* and *czcD* divalent cation efflux, have been accumulated in the FW106 community. Comparisons of FW106 to the acid mine drainage community metagenome reveals community-specific gene profiles (e.g. geochemical genes, cytochromes, etc.) but also common mechanisms of adaptation, including the accumulation of czcABC metal transporter genes. Finally, the FW106 metagenome was compared to the preliminary (100 Mb) metagenomic sequence from pristine FRC groundwater (FW301). In contrast to the low species diversity of FW106, FW301 is represented by multiple phyla (predicted 400+ species) including α-, β-, γ- and δ-proteobacteria, Planctomycetes, Chloriflexi, Actinobacteria, Acidobacteria and Firmicutes. In contrast to the FW106 sample which yields large contigs, the FW301 sample is composed largely of single reads that do not assemble into contigs (95%). Most of the geochemical resistance genes identified in FW106 also exist in the ancestral pristine sample, thus accounting for the original source of genetic material necessary for adaptation of the community to contamination.

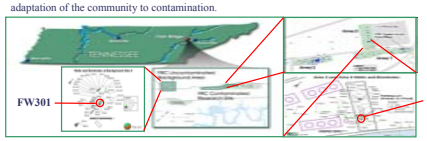


Fig. 1. Geographic location of sampling wells FW106 and FW301

## FW106 Groundwater Geochemistry

| | FW300 | FW106 |
|---|---|---|
| pH | ~7 | ~3.7 |
| Nitrate (mg/L) | 1.5 | 2331 |
| Sulfate (mg/L) | 6.3 | 1997 |
| Uranium (mg/L) | >0.0001 | 51 |
| Technetium-99 (pCi/L) | - | 3700 |
| 1,2-Dichloroethene (μg/mL) | 5 | 1153 |
| Tetrachloroethene (μg/mL) | 5 | 810 |
| 1-Butanol (μg/mL) | - | 475 |
| Acetone (μg/mL) | 10 | 823 |
| Benzoic Acid (μg/mL) | - | 1400 |
| Sodium (mg/L) | 2.96 | 826 |
| Chloride (mg/L) | 1.125 | 465 |
| Magnesium (mg/L) | 2.58 | 45.7 |

## FW106 Metagenome Statistics

| | FW106 (x2) | FW301 |
|---|---|---|
| Total Bases | 9554544 | 106573620 |
| DNA Coding # Bases | 8076611 | 84170988 |
| DNA G+C # Bases | 6011119 (63.20%) | 60905616 (57.16%) |
| Scaffolds | 6079 | 134883 |
| Total # Genes | 12493 | 170521 |
| 16S rRNA | 3 | 51 |
| Genes w/ Function Prediction | 8692 | 86183 |
| Genes Assigned to Enzymes | 1692 | 53901 |
| Genes Assigned to KEGG Pathways | 1423 | 44007 |
| Genes in COGs | 7961 | 85262 |
| Genes in Pfam | 5799 | 54364 |

## Phylogenetic Distribution of Genes

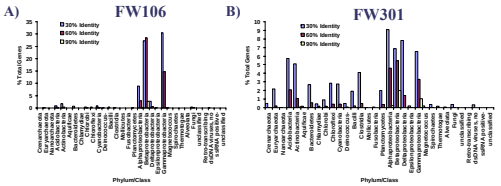


Fig. 2. Phylogenetic distribution of genes in the FW106 (A) and FW301 (B) metagenomes

**Introduction of contamination results in a massive loss of biodiversity, resulting in groundwater communities dominated by γ-, β- and α-proteobacterial species**

## Characteristics of FRC Communities

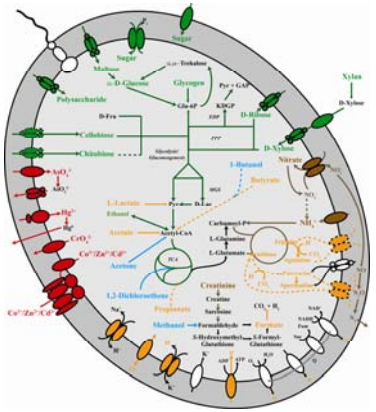

Fig. 3. Reconstructed metabolism of the dominant γ-proteobacterial species of the FW106 community color coded to show geochemically-relevant pathways (green, carbon metabolism; red, metal resistance; orange, acid resistance; blue, organic solvent resistance; brown, nitrogen metabolism). Dotted pathways/genes indicate missing or ambiguous pathways/genes.

## Lateral Gene Transfer in FW106 Community

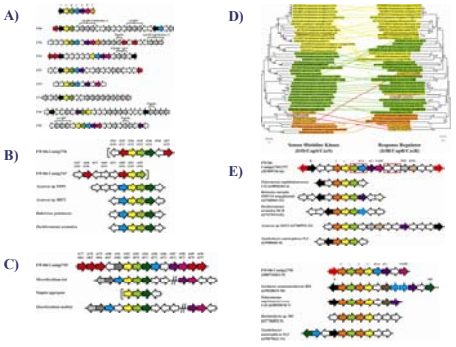


Fig. 4. Evidence of lateral transfer of geochemical resistance genes in the FW106 community (red genes indicate mobile elements, other colors represent orthologous sets). A) Distribution of Mer operon genes in FW106; B) Lateral transfer of β-proteobacterial genes into FW106 γ1; C) Lateral transfer of into FW106 γ1; D) Transposon insertion into L-Lys 6-aminotransferase gene of FW106 γ1; E) Phylogeny of heavy metal responsive two component systems in FW106, suggesting LGT from *Geobacter* species (lines connect syntenous pairs); E) Lateral transfer of two acetone carboxylase operons into FW106.

**Resistance to contamination is conferred by multiple specific genetic mechanisms, including heavy metal detoxification and export, denitrification, stabilization of chemiosmotic gradient under acidic conditions and degradation of organic contaminants. Evidence suggests that many of these resistance mechanisms were acquired via lateral gene transfer**

## Comparative Metagenomics of FW106

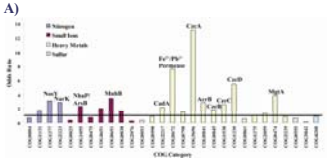


Fig. 5. Odds ratios of FW106 gene categories (based on geochemically-relevant COG categories) compared to all sequenced bacterial genomes (IMG database).


Fig. 6. Abundance profile of specific COG categories (geochemical resistance and cytochromes) between FW106 and acid mine drainage metagenomes (*z*-score normalized; red, most abundant; green, least abundant).

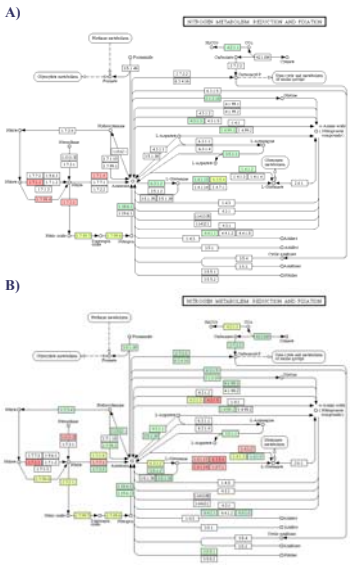


Fig. 7. Abundance profiles of nitrogen metabolism genes (based on KEGG database) for FW106 (A) and FW301 (B) (same color scheme as Fig. 6).

## Evolution of the FW106 Community

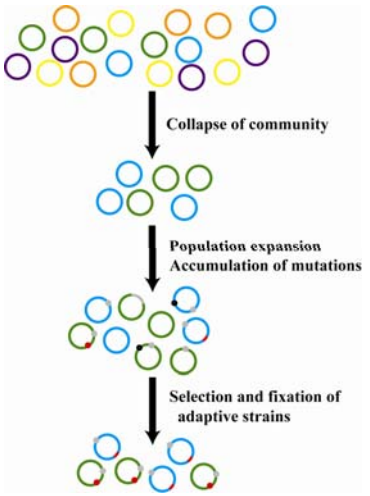


Fig. 8. Proposed model for the evolution of microbial communities under stressed conditions. Introduction of stressors to a diverse community results in a rapid loss of species diversity. In the case of severe contamination, this loss could result in the transient isolation of microcommunities and population bottlenecks which would in turn result in the loss of strain diversity within surviving species. As the surviving community begins to recover, nearly neutral mutations (via single-point mutation, LGT, gene duplication, etc., indicated by red dots/lines) will begin to accumulate. The cumulative effects of selective sweeps and background selection will serve to further purge local genomic diversity at linked loci. In cases where the populations remain isolated (e.g. geographical separation, strong geochemical gradients limiting gene flow, etc.), the cumulative effect of multiple selective sweeps may result in the emergence of new species specifically adapted to the prevalent conditions.

## Conclusions

- **Introduction of high levels of multiple stressors has resulted in the irreversible loss of species and strain biodiversity at FW106**
- **Selective pressures imposed by stressors have resulted in rapid adaptations to general and specific stresses (metals, acidity, organic solvents, nitrate/nitrite)**
- **The community employs a heterotrophic, respiratory lifestyle coupling denitrification to metabolism of simple and complex carbohydrates permeating from soil**
- **FW106 and AMD communities show distinct abundance profiles (cytochromes, *czcD*) but may share common adaptive strategies (accumulation of *czcABC* genes)**
- **Preliminary comparison of FW106 and FW301 metagenomes suggest most of the geochemically-relevant genes necessary for survival at the FW106 site were present in the ancestral population, but not all were maintained (i.e. aromatic degradation**

## ACKNOWLEDGEMENT

ESPP2 (MDCASE) is part of the Virtual Institute for Microbial Stress and Survival (VIMSS) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomics:GTL Program through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy.