Development of Integrated Genomic Technology for Microbial Community Analysis

Jizhong (Joe) Zhou

Zhouj@ornl.gov, 865-576-7544

Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
Project Objectives

• Developing microarrays-based genomic technologies for microbial detection.

• Use of the developed genomic technologies for assessing microbial community structure and dynamics related to bioremediation.
Pioneering advances in microarray-based technologies to address challenges in microbial community genomics

• Challenges:
  – **Specificity**: Environmental sequence divergences.
  – **Sensitivity**: Low biomass.
  – **Quantification**:
    – **Existence of contaminants**: Humic materials, organic contaminants, metals and radionuclides.

• Solutions
  – Developing different types of microarrays and novel chemistry to address different levels of specificity.
  – Developing novel signal amplification strategy to increase sensitivity
  – Optimizing microarray protocols for reliable quantification.
Summary of 50mer-based FGAs for environmental studies

Oligonucleotide probe size: 50 bp

Rhee et al. 2004, AEM 70:4303-4317

- Nitrogen cycling: 302
- Sulfate reduction: 204
- Carbon cycling: 566
- Phosphorus utilization: 79
- Organic contaminant degradation: 770
- Metal resistance and oxidation: 85

- Total: 2,006 probes
- All probes are < 88% similarity
Specificity of 50 mer microarrays

Very specific hybridization was obtained

- 5 nirS genes were mixed together
- Only corresponding genes were hybridized

- 6 types of genes were mixed together
- Only corresponding genes were hybridized
Sensitivity

Detection limit
- 50 ng pure DNA in the presence of non-target templates
- $10^7$ cells
Quantification and validation

**Microarray hybridization**

- Good linear relationship
- Quantitative

- Microarray result is consistent with real-time PCR
Novel amplification approach for increasing hybridization sensitivity

As low as 10fg (2 cells) can be detected

Amplification is quantitative for majority of the genes

Submitted to PNAS
6 samples were taken to assess the effects of contaminants on microbial community structure.
Groundwater samples with very low biomass

- 2L groundwater from six different sites.
- Cell counts: $1-5 \times 10^5$/ml
- DNA was isolated, $1/20$ of the DNA was manipulated and used for hybridization.
- Nice hybridization was obtained with the DNA manipulated with the new method.
- No hybridization were obtained if the DNA is not manipulated.
Difference of functional genes in samples from NABIR Field Research Center

- Clear difference was observed among contaminated and noncontaminated sites.
- E.g., some genes are present in noncontaminated site but not in contaminated sites.

Reference site

Highly contaminated site
## Overall diversity among different samples

<table>
<thead>
<tr>
<th></th>
<th>FW300</th>
<th>FW003</th>
<th>FW021</th>
<th>FW010</th>
<th>FW024</th>
</tr>
</thead>
<tbody>
<tr>
<td>FW300</td>
<td>61(20%)</td>
<td>189(36%)</td>
<td>174(35%)</td>
<td>80(21%)</td>
<td>111(23%)</td>
</tr>
<tr>
<td>FW003</td>
<td></td>
<td>25(11%)</td>
<td>144(35%)</td>
<td>61(17%)</td>
<td>84(20%)</td>
</tr>
<tr>
<td>FW021</td>
<td></td>
<td></td>
<td>10(5%)</td>
<td>64(20%)</td>
<td>90(24%)</td>
</tr>
<tr>
<td>FW010</td>
<td></td>
<td></td>
<td></td>
<td>6(5%)</td>
<td>118(37%)</td>
</tr>
<tr>
<td>FW024</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30(16%)</td>
</tr>
<tr>
<td>Total Genes Detected</td>
<td>302</td>
<td>219</td>
<td>192</td>
<td>130</td>
<td>190</td>
</tr>
<tr>
<td>Genetic diversity, Simpson’s (1/D)$^a$</td>
<td>125.5</td>
<td>67.1</td>
<td>26.6</td>
<td>17.4</td>
<td>35.7</td>
</tr>
</tbody>
</table>

- Overall diversity correlates with contaminant level.
- The proportion of overlapping genes between samples was consistent with the contaminant level and geochemistry.
- A significant portion (5-20%) of all detected genes were unique to each sample, even though they are very close. Thus, important microbial populations appear to be highly heterogeneous in this groundwater system.
# CommOligo --- New oligo probe design program for community analysis

## Number and specificity of designed probes (50-mer) by different programs

<table>
<thead>
<tr>
<th>Programs used</th>
<th>Whole-genome sequences of <em>M. maripaludis</em> (1766 ORFs)</th>
<th>Group sequences of <em>nirS</em> and <em>nirK</em> (842 gene sequences)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total ORFs</td>
<td>ORFs rejected</td>
</tr>
<tr>
<td>ArrayOligosector</td>
<td>1766</td>
<td>7</td>
</tr>
<tr>
<td>OligoArray</td>
<td>1766</td>
<td>68</td>
</tr>
<tr>
<td>OligoArray 2.0</td>
<td>1766</td>
<td>68</td>
</tr>
<tr>
<td>OligoPicker</td>
<td>1766</td>
<td>18</td>
</tr>
<tr>
<td>CommOligo</td>
<td>1766</td>
<td>9</td>
</tr>
</tbody>
</table>

- Useful for both whole genome microarrays and community arrays
- Able to design group-specific probes
- Better performance than other programs
Probes Designed for a Second Generation FGA

- Nitrogen cycling: 5089
- Carbon cycling: 9198
- Sulfate reduction: 1006
- Phosphorus utilization: 438
- Organic contaminant degradation: 5359
- Metal resistance and oxidation: 2303

Total: 23,408 genes

- 23,000 probes designed

- Will be very useful for community and ecological studies
Biostimulation of microbial populations for Ur removal

1. Vacuum strip volatiles; chemically neutralize acid and precipitate metals

2. Denitrify water in FBR

3. Inject treated water into outer well

Above ground treatment

Source water

Ethanol addition

Multi-level wells

- Above ground denitrification and neutralization of groundwater
- *in situ* biostimulation with ethanol and reduction of U(VI)
Treatments and sampling

- Ethanol was injected nearly weekly from about Day 137.
- Groundwater was sampled (1-5L) nearly weekly for microarray analysis.
- Geochemical data were measured almost daily.
Overall community similarity

- Initially, 026 & 102-3 were similar but distinct from 101-2 although they are only a few meters away (Black box), indicating heterogeneity in the microbial populations.
- Over time, the populations in the different wells became more similar to each other (Red box), possibly due to continual influx of injected groundwater.
- These results suggest bioremediation treatment significantly altered community compositions.
Nitrite Reduction Genes

- Total N reduction gene signals correlated with nitrate levels.
- Additional samples after 354 d are being analyzed to see if the trend continued.
Nitrite Reductase Genes

- Total nirS and nirK signals correlated with nitrate levels.
- nirK genes were dominant but nirS increased in 354 d sample.
- Total numbers of nirS and nirK genes detected had same trends.
- Bacteria containing nirK genes respond to biostimulation more rapidly.
• Total S reduction gene signals correlated with sulfate levels.
• Total numbers of S reduction genes detected had same trends.
• Several FRC dsrA/B clones were detected.
N & S Reduction and Cytochrome Genes

- N and S reduction and cytochrome C genes followed trends in U(VI) levels.
Cytochrome Genes

- Both Geobacter- and Desulfovibrio-like species were detected.
- *Geobacter* sp.-like cytochrome C genes followed trends in U(VI) levels.
- Most prominent during initial denitrification phase.
Cytochrome Genes

- Most detected cytochrome genes were similar to *Geobacter*-like bacteria
- Desulfovibrio-like bacteria were also detected.
- But the result could be biased because many more genes from *Geobacter* were used as probes
Whole community sequencing

- Sample from NABIR Field Research Center at ORNL
- Sequenced by DOE Joint Genome Institute
- 20 species based on 16S rRNA
Current status of sequencing

- Collected 2,000 L groundwater
- Took about 6 months to optimize the protocols.
- 300 ug DNA was isolated.
- Sent DNA twice to JGI for library construction.
- Libraries
  - 40 kb fosmid library
  - 8 kb library
  - 3 kb library
- Very good 40kb library was obtained at the first but not 8 and 3 kb library.
- Sequencing is in process
Conclusions

• Development
  – Very comprehensive oligonucleotide arrays for environmental studies were developed. This is the most comprehensive arrays available today.
  – The arrays are specific and quantitative.
  – Novel approach for increasing sensitivity is developed. This made it possible to use microarrays for analyzing environmental samples.
  – New computer program was developed for probe designing.

• Applications
  – Microbial populations are highly heterogeneous in NABIR FRC.
  – Contaminants have significant effects on microbial community structure and dynamics.
  – Microbial populations at the FRC sites can be stimulated for removing uranium.
  – Geobacter and Desulfovibrio-like species could be responsible for uranium reduction after stimulation.
Acknowledgement

• ORNL
  – Liyou Wu
  – Xuedan Liu
  – Sung-Keun Rhee
  – Song Chong
  – Jost Liebich
  – Christopher Schadt
  – Terry Gentry
  – Zhili He
  – Sonia Tiquia
  – Zhijian Huang
  – Weimin Gao
  – David Watson
  – Phil Jardine
  – Baohua Gu
  – Tommy Phelps

• Stanford University
  – Craig Criddle
  – Weimin Wu
  – Peter Kitanidis
  – Many people from Craig’s group

• Michigan State University
  – James M. Tiedje
  – Terry Marshall

● LBL
  – Terry Hazen

● JGI
  – Eddy Rubin
  – Chris Detter

● Miami University
  – Matthew Fields