S1-4

# Application of Microarrays for Microbial Community Studies

Sung-Keun Rhee, Jin-Woo Bae, Byung-Chun Kim, Jizhong Zhou[1], and Yong-Ha Park

*Biological Resources Center, Korea Research Institute of Bioscience and Biotechnology (KRIBB), Taejon, Korea; [1]Environmental Sciences Division, Oak Ridge National Laboratory (ORNL), Oak Ridge, TN, USA*

The microarray is a powerful genomic technology that is widely used to study various biological processes. Although microarray technology has been used successfully to analyze global gene expression in pure cultures, application of microarray to environmental studies presents great challenges in terms of specificity, sensitivity, and quantitation (Zhou and Thompson 2002). Although microarray-based genomic technology has attracted tremendous interests among microbial ecologists, it has only recently been extended to study microbial communities in the environment (Wu et al. 2001).

Based on the types of probes arrayed, microarrays used in environmental studies can be divided into three major classes: functional gene microarrays (FGA), Genome-Probing microarrays (GPM), and phylogenetic oligonucleotide microarrays (POAs). FGAs contain probes corresponding genes encoding key enzymes involved in various ecological and environmental processes, such as carbon fixation, nitrification, denitrification, sulfate reduction, and contaminant degradation. Both PCR-amplified DNA fragments (Wu et al. 2001) and oligonucleotides (oligos) derived from functional genes can be used to fabricate FGAs. To avoid confusion, the former is referred to as PCR product-based FGAs whereas the later is referred to as oligonucleotide-based FGAs. These types of arrays are useful in studying physiological status and functional activities of microbial communities in natural environments. GPMs are constructed using whole genomic DNA isolated from pure culture microorganisms and can be used to describe a microbial community in terms of its cultivable component.

In the presentation, for monitoring biodegradation potential and activity, we are going to show development of 1) a comprehensive 50-mer oligonucleotide microarrays containing probes (1,657) from all of the known genes (2,402) involved in biodegradation and metal resistance and 2) a genome-probing microarrays containing probes from lactic acid bacterial genomes. These microarrays were evaluated for bioremediation monitoring of contaminated soils and studying microbial community dynamics during mixed culture fermentation. Our results demonstrated the developed microarrays offer rapid, powerful, new high throughput tool for monitoring potential and functional activity of microbial communities.

**Probe specificity**. Since much shorter sequences were used as probes, the 50-mer FGAs should have higher resolution than the PCR products-based FGAs. To experimentally determine the resolution power of the hybridization with the 50-mer FGAs, the effects of probe sequence similarity on hybridization signal intensity of 6 target genes were evaluated (Fig. 1A). The probes were artificially designed to be different by 2% (i.e., 1 bp length). Little hybridization was observed for probes showing 76–88% identity to the target

DNA for all 6 genes, whereas the signal intensity increased substantially for probes showing more than 96% similarity to the target DNA (Fig. 1A). The SNRs varied from 0 to 0.5 for all of the genes when the probe sequence similarities were < 88%, which was much smaller than the generally accepted threshold value of SNR = 3.0. Specificity evaluation with pure cultures was also carried out using genomic DNA of four reference strains. Probes corresponding to all known genes reported in the reference strains had strong hybridization signals to their corresponding genomic DNA.

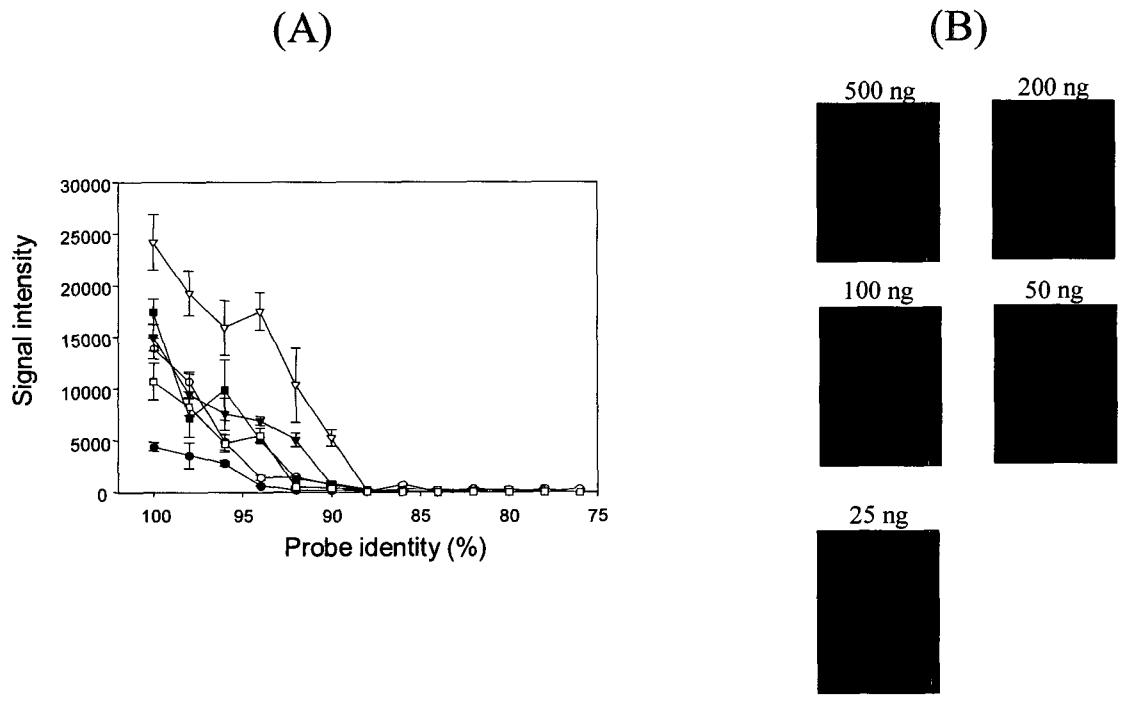


**Fig. 1. Effect of probe-target sequence identities on hybridization signal intensity.**

(A) The symbols correspond to the genes from *P. putida* PpG7 (●, gi151388 naphthalene dioxygenase, large subunit; ○, gi151385naphthalene dioxygenase reductase, ▼, gi5070351 dihydroxynaphthalene dehydrogenase) and Gpo1 ( ▽, gi5824143 alkane 1-monoxygenase; ■, gi5824146 aldehyde dehydrogenase, □, gi5824147 alcohol dehydrogenase. Error bars indicated the standard deviation from 4 replicates. (A) Artificial probes of 76–100% sequence identity to the target genes were hybridized with PCR-amplified target genes in triplicate.

(B) Fluorescence images showing DNA detection sensitivity with pure genomic DNA. Genomic DNA of *T. aromatica* K-172 was serially diluted in 1xTE buffer. The diluted genomic DNAs (5-1000 ng) were mixed with 1 μg of the negative control genomic DNA from *S. oneidensis* MR-1, respectively, labeled with Cy5 in a total hybridization solution volume of 20 μl, and hybridized with the microarrays. The arrays were scanned with 100% laser power and 100% PMT gain. A portion of the image is presented here, which contained probes from the following genes: succuinyl-CoA:benzylsuccinate CoA-transferase (1, gi9622538; 2, gi9622535; 3, gi9622533 4. gi9622531), benzoyl-CoA reductase (5, gi3724172; 6, gi3724140; 7, gi3724168) and 6-oxocyclohex-1-ene-1-carbonyl-CoA hydratase (8. gi3724166).

**Detection sensitivity.** The detection sensitivity of the 50-mer FGA-based hybridization was determined using 1–1000 ng of genomic DNA extracted from a pure culture of *T. aromatica* K172, which was randomly labeled with Cy5. In environmental samples, the microorganisms of interest are present together with other diverse microorganisms. The existence of other non-target DNAs may affect the hybridization with target DNA and hence decrease detection sensitivity. To evaluate the detection sensitivity in the presence of heterogeneous non-target DNA, genomic DNA (10–1000 ng) from *T. aromatica* K172 were mixed with 1 μg of *S. oneidensis* MR-1 DNA, respectively, and randomly labeled with Cy5. At a hybridization temperature of 50°C in the presence of 50% formamide, the strongest hybridization signals

were observed with 50 ng of *T. aromatica* K172 DNA for the target gene, 2-oxoglutarate ferredoxin-oxidoreductase beta subunit (gi19571178) (Fig. 1B). Hybridization signals using 25 ng of genomic DNA, however, were barely detectable.

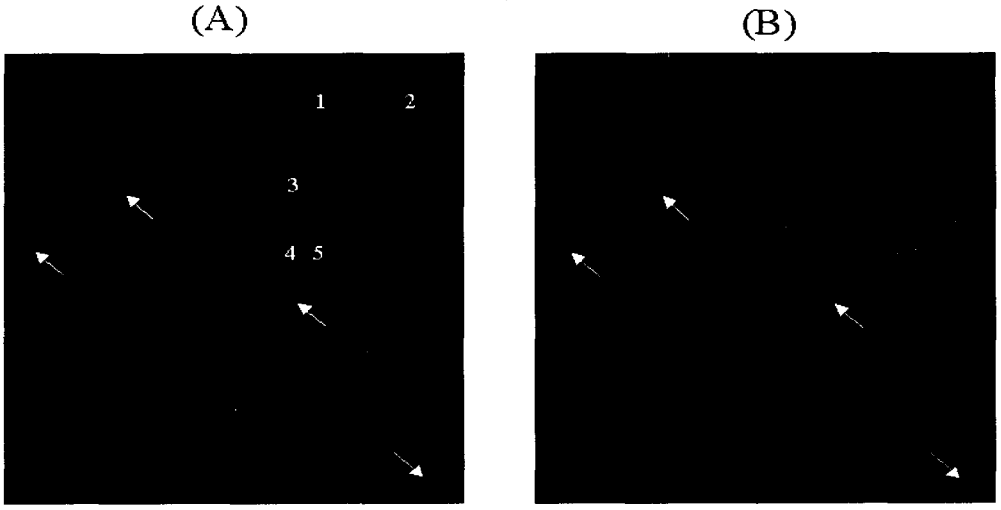


**Fig 2. Fluorescence images showing the specific detection of naphthalene degradation genes in the microcosm experiments.**

Only a part of the whole microarray 1 of 8 meta-grid was presented to show the difference of community between the soils incubated with naphthalene TFD-N and without any addition TFD. A community genomic DNA 5 ug from the TFD-N was labeled with Cy5 using the random primer labeling method. B Community genomic DNA 5ug from the TFD was labeled with Cy5 using the random primer labeling method. The scanned image display of the hybridization reaction is presented, with the arrow indicating the hybridization signals detected both for both TFD1 and TFD1-N and number indicating genes detected only in TFD-N 1. naphthalene dioxygenase Fe-S subunit gi18307552, 2. naphthalene dioxygenase gi17863951, 3. 2,4-DNT dioxygenase gi1477920, 4. fumarylpyruvate aldolase gi4220434, 5. glutathione-S-transferase protein gi4220432.

**Detection of naphthalene-degrading genes in soil microcosms.** To determine whether the developed microarrays could be used to detect microorganisms within the context of environmental applications, two microcosms were established with soil samples from TFD: one with naphthalene vapor, and the other without naphthalene vapor as a control. The hybridization signals of 40 genes were significantly different ($P = 0.05$). Among them, 22 genes were highly different with more than 3 fold difference (Table 1). Among them, *Ralstonia* sp. strain U2-type microorganisms containing the plasmid might be one of major constituents for degrading naphthalene in the microcosm. However, in contrast to the enrichment experiment, the plasmid-encoded genes in *Rhodococcus* sp. NCIMB12038 were not detected with the microarray, thus indicating that the enrichment process could cause severe bias of the population compositions.

Various contaminated soils were analyzed using this microarray. Clustering analysis of hybridization revealed 5 main groups (Fig. 25). Many genes related to naphthalene degradation were clustered together and they were abundant in naphthalene-amended soil.

**GPM for monitoring community dynamics during fermentation.** A genome-probing microarray was

constructed using 150 genomes of lactic acid bacteria. Evaluation and application of this format of microarray will be comparatively presented with the functional gene microarray.

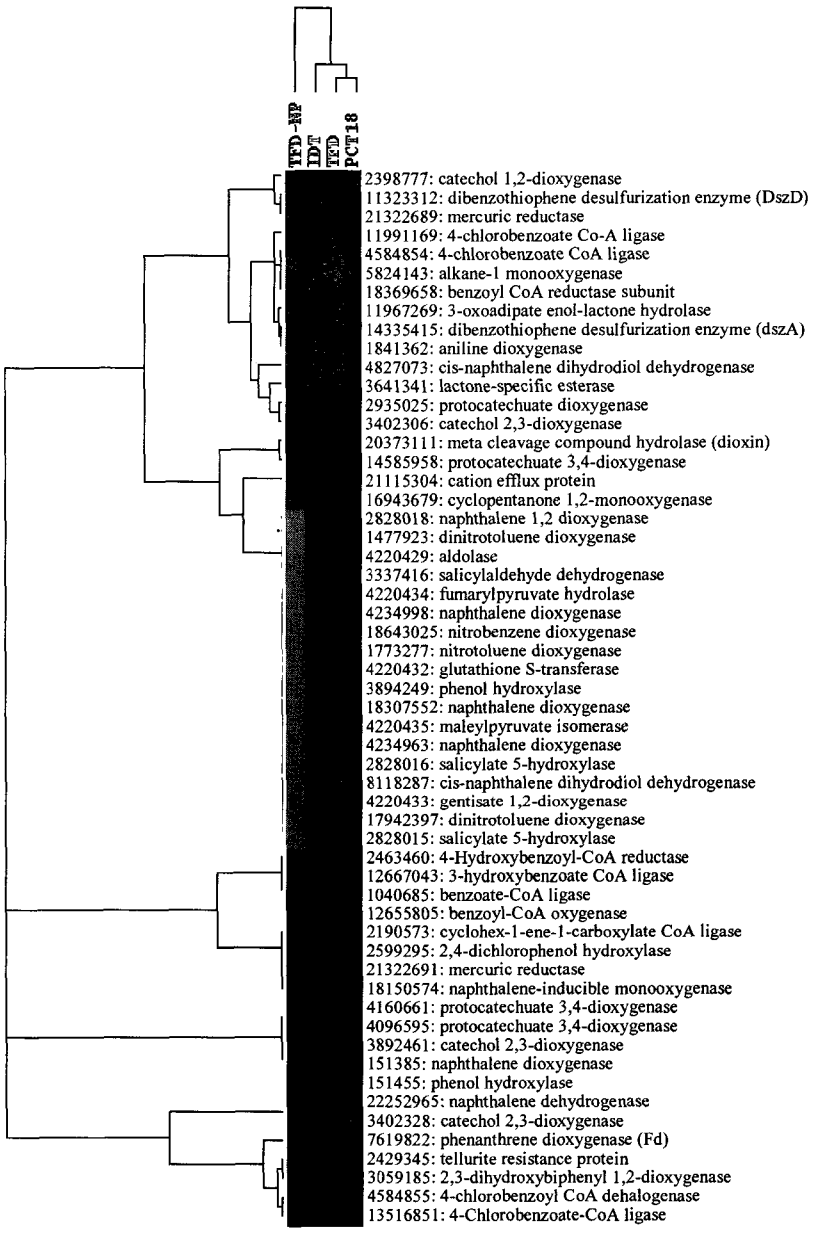


**Fig. 3.** Hierarchical cluster analysis of community relationships based on hybridization signal intensity ratios for the genes showing SNR > 3. This figure was generated using hierarchical cluster analysis CLUSTER and visualized with TREEVIEW (Eisen et al. 1998). The hybridization signals of genomic DNA from each of the four contaminated soils was divided by the hybridization signals from genomic DNA of OR. Microarray hybridization patterns with the labeled genomic DNAs from the soils are showed in each column. Each row represents the hybridization signal observed for each gene when the genomic DNA from the contaminated soil indicated in the column was used for hybridization. Black represents no detectable difference in the hybridization signal, while red represents the significant hybridization signal. The columns correspond to the hybridization patterns obtained with Cy5-labeled genomic DNA from the following contaminated soils: P-MC, PCT18, and TFD1, TFD1-N TFD soil incubated under naphthalene vapor.

**Table 1. Genes detected in soil microcosm incubated with naphthalene.**

| Gene (Gi number) | Gene function | Microorganism | Log of mean signal ratio[a] | SNR[b] |
|---|---|---|---|---|
| 1477920 | 2,4-DNT dioxygenase ferredoxin oxidoreductase | *Burkholderia* sp. RASC | 1.64 ± 0.18 | 3.23 ± 0.45 |
| 1477923 | 2,4-DNT dioxygenase large subunit | *Burkholderia* sp. RASC | 2.02 ± 0.23 | 3.42 ± 0.62 |
| 1773277 | 2-NT oxygenase large subunit | *Pseudomonas* sp. | 1.98 ± 0.23 | 6.32 ± 0.43 |
| 2828015 | salicylate-5-hydroxylase large subunit | *Ralstonia* sp. U2 | 1.26 ± 0.13 | 13.94 ± 1.03 |
| 2828016 | salicylate-5-hydroxylase small subunit | *Ralstonia* sp. U2 | 1.94 ± 0.22 | 15.02 ± 0.75 |
| 2828018 | naphthalene 1,2 dioxygenase large subunit | *Ralstonia* sp. U2 | 1.34 ± 0.14 | 3.77 ± 0.27 |
| 3337416 | salicylaldehyde dehydrogenase | *Ralstonia* sp. U2 | 1.49 ± 0.16 | 4.33 ± 0.24 |
| 3894249 | phenol hydroxylase | *unidentified bacterium* rN5 | 1.32 ± 0.14 | 6.85 ± 0.45 |
| 4220429 | putative aldolase | *Ralstonia* sp. U2 | 1.74 ± 0.19 | 3.99 ± 0.12 |
| 4220432 | glutathione-S-transferase protein | *Ralstonia* sp. U2 | 1.65 ± 0.18 | 6.62 ± 0.67 |
| 4220433 | gentisate 1,2-dioxygenase | *Ralstonia* sp. U2 | 2.12 ± 0.26 | 10.63 ± 0.45 |
| 4220434 | fumarylpyruvate hydrolase | *Ralstonia* sp. U2 | 2.34 ± 0.28 | 8.14 ± 0.27 |
| 4220435 | maleylpyruvate isomerase | *Ralstonia* sp. U2 | 2.27 ± 0.27 | 7.74 ± 0.56 |
| 4234963 | naphthalene dioxygenase | *uncultured bacterium* D1b | 1.95 ± 0.22 | 9.58 ± 0.88 |
| 4234998 | naphthalene dioxygenase | *uncultured bacterium* U3b | 1.64 ± 0.18 | 5.62 ± 0.59 |
| 8118285 | polyaromatic hydrocarbon dioxygenase large subunit | *Comamonas testosteroni* | 1.22 ± 0.13 | 3.94 ± 0.45 |
| 8118287 | cis-naphthalene dihydrodiol dehydrogenase | *Comamonas testosteroni* | 1.63 ± 0.18 | 13.96 ± 1.22 |
| 16943679 | cyclopentanone 1,2-monooxygenase | *Comamonas testosteroni* | 0.77 ± 0.08 | 13.92 ± 1.73 |
| 17863951 | naphthalene dioxygenase | *Burkholderia* sp. S1-17 | 1.72 ± 0.19 | 5.13 ± 0.76 |
| 17942397 | 2,4-DNT oxygenase iron sulfur protein | *Burkholderia cepacia* | 2.04 ± 0.24 | 11.93 ± 0.35 |
| 18307552 | naphthalene dioxygenase Fe-S protein | *Ralstonia* sp. NI1 | 2.32 ± 0.28 | 7.88 ± 0.67 |
| 18643025 | nitrobenzene dioyxgenase alpha subunit | *Comamonas* sp. JS765 | 1.12 ± 0.12 | 8.53 ± 0.34 |

[a]Hybridization signal ratios (log mean ± standard deviation from 6 replicates) of the naphthalene-spiked TFD soil to the control TFD soil. The hybridization signals of 40 genes were significantly different (P = 0.05), but only the genes with more than 3 fold difference were listed.
[b]Mean ± standard deviation from 6 replicates

## References


1. **Rhee et al.** 2004. Detection of biodegradation and biotransformation genes in microbial communities using 50-mer oligonucleotide microarrays. Appl. Environ. Microbiol. In press.
2. **Wu, L. Y., D. K. Thompson, G. Li, R. A. Hurt, J. M. Tiedjie, and J. Zhou.** 2001. Development and evaluation of functional gene arrays for detection of selected genes in the environment. Appl. Environ. Microbiol. **67**:5780–5790.
3. **Zhou, J.** 2003. Microarrays for bacterial detection and microbial community analysis. Curr. Opinion. Microbiol. **6**:288-294.
4. **Zhou, J.-Z., and D. K. Thompson.** 2002. Challenges in applying microarrays to environmental studies. Curr. Opion. Biotech. **13**:202–204.