Development and Use of Integrated Microarray-Based Genomic Technologies for Assessing Microbial Community Composition and Dynamics

Jizhong Zhou (PI)[1], Liyou Wu[1], Terry Gentry[1], Christopher Schadt[1], Zhili He[1], and Xingyuan Li[2]

[1]Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN;
[2]Perkin Elmer Life and Analytical Sciences, Torrance, CA

The recent development of microarray technology provides a great opportunity for the simultaneous identification of thousands of microbial genes/populations. However, low microbial biomass often prevents application of this technology to microbial communities in natural settings. We have developed a whole community genome amplification (WCGA)-assisted microarray-based detection approach for analysis of microbial communities whose members can not be studied using conventional technology. With optimized buffer systems, as few as two bacterial cells could be detected. Whole genome microarray hybridization showed that representative detection of individual genes or genomes was obtained within the DNA concentrations of 1 to 100 ng from individual or mixed genomes. Significant linear relationships were observed between signal intensity and initial DNA concentration ranging from: (1) 40 pg to 125 ng for the majority of *Shewanella* genes (r^2 = 0.65–0.99) and other organisms as detected by whole genome arrays; (2) genomes in constructed communities from 0.1 to 1,000 ng (r^2 = 0.91) using community genome arrays; and (3) community DNAs diluted from a stimulated natural groundwater ranging from 0.01 ng to 250 ng (r^2 = 0.96–0.98) using functional gene arrays. We applied this technology to investigate microbial communities in six groundwater samples contaminated with uranium and other metals using functional gene arrays (~2,000 probes). The results indicated that microbial populations containing important genes involved in contaminant degradation and immobilization have locally heterogeneous distributions and that microbial diversity is greatly decreased in contaminated environments. This is the first time that microarrays have been successfully used to analyze low biomass communities, such as those commonly found in settings important to human health, industry, and environmental management.

We have also developed a software program, CommOligo, for designing probes from whole-genomes, meta-genomes, or a group of sequences. The program uses a new global alignment algorithm to design single or multiple unique probes for each gene with default settings for a maximal similarity of 85%, a maximal number of continuous match of 15 bases, and free energy of –30 kcal/mol. The program is also able to design single- or multiple-group-specific probes for a group of genes with a minimal similarity of 96% within a group and the same parameters as unique probes outside a group. The program was evaluated using both whole-genome and highly homologous sequence data and compared with other probe design software. The results clearly demonstrate that CommOligo performed better and can be used for oligonucleotide probe design from various types of sequence data. With this program, more comprehensive functional gene arrays containing ~23,000 probes for important biogeochemical cycling (C, N, S, and P), metal resistance, and contaminant degradation genes have been designed and constructed. This is the most comprehensive array currently available for environmental studies. We are applying this microarray to the study of several microbial communities and processes at the NABIR-FRC, including a denitrifying fluidized bed reactor and groundwater stimulated with ethanol for uranium reduction.