Chapter 5
Analysis of Microbial Communities by Functional Gene Arrays

Joy D. Van Nostrand, Zhili He, and Jizhong Zhou

Introduction

A major hurdle to the study of microbial communities is that only about 1% of microorganisms are cultivated (Whitman et al. 1998). As such, culture-independent approaches are necessary in order to examine the vast majority of environmental microorganisms. Many molecular techniques are available for community analysis, and most of these techniques utilize phylogenetic markers such as the 16S rRNA or the DNA gyrase gene (*gyrB*) (Wilson et al. 1990; Yamamoto and Harayama 1995; Hugenholtz et al. 1998; Brodie et al. 2006). While the use of these genes provides information regarding phylogenic diversity and structure of a microbial community, they don't provide much, if any information relating to the functional potential and/or activity of the community. Functional genes have been used to examine both phylogenetic and functional diversities (e.g., McDonald et al. 1995; Braker et al. 1998). However, even if multiple functional genes are examined, conventional molecular techniques only provide information on a small fraction of the community. This is because conserved PCR primers cannot be designed for many functional genes of interest due to a lack of sequence homology or a lack of a sufficient number of sequences. Consequently, conventional PCR-based approaches cannot be used to detect and quantify many functional genes of interest. As such, a more comprehensive technique is required to provide a full picture of microbial community activity and dynamics in a rapid, parallel, and high-through-put manner.

Use of microarrays is one way to overcome these limitations. Microarrays provide a rapid and high-throughput method of examining thousands of functional

J.D. Van Nostrand, Z. He, and J. Zhou (✉)
Department of Botany and Microbiology, Institute for Environmental Genomics, University of Oklahoma, 101 David L. Boren Blvd, Norman, OK 73019, USA
e-mail: jzhou@ou.edu

L.L. Barton et al. (eds.), *Geomicrobiology: Molecular and Environmental Perspective*,
DOI 10.1007/978-90-481-9204-5_5,

genes at one time without the need for conserved primers. Microarray technology was first developed for analyzing gene expression in the plant *Arabidopsis thaliana* (Schena et al. 1995). Such gene expression arrays are now commonly used to study individual microorganisms. However, the potential usefulness of microarray technology was greatly expanded when Guschin and colleagues (1997) proposed and tested the use of microarrays to study microbial communities. Now, several different types of arrays are available for the study of microbial communities in the environment (Zhou and Thompson 2002; Zhou 2003; Gentry et al. 2006). Phylogenetic oligonucleotide arrays (POA) are designed to determine community composition or phylogenetic relatedness using 16S rRNA or other phylogenetically informative genes (Small et al. 2001; Loy et al. 2001; Wilson et al. 2002). The most comprehensive POA is the PhyloChip, which contains almost 300,000 perfect-match and mismatch probes for 842 subfamilies (Brodie et al. 2006); although the 23S rRNA gene has also been used for POA construction (Lee et al. 2006). Community genome arrays (CGA) are used to examine the relatedness of microbial strains or to detect specific microorganisms in the environment using whole genomic DNA of individual species or strains as probes (Wu et al. 2004; Zhang et al. 2004; Wu et al. 2008). Metagenomic arrays (MGA) are made using clone libraries created from environmental DNA as probes (Sebat et al. 2003). This array was used as a high-throughput screening method for clone libraries. Whole-genome open reading frame (ORF) arrays (WGA) are comprised of probes for all ORFs in one or more genomes (Wilson et al. 1999). This type of array has been used to examine the diversity and relatedness of several metal-reducing *Shewanella* strains (Murray et al. 2001). A WGA of 353 virulence factors was used to evaluate over 100 strains of *Pseudomonas syringae* to determine those genes associated with host specificity and several genes were identified that were statistically associated with specific hosts (Sarkar et al. 2006). Functional gene arrays (FGA) are composed of probes for key genes involved in microbial functional processes of interest (Wu et al. 2001; Gentry et al. 2006; He et al. 2007). FGAs allow for the simultaneous examination of many functional gene groups (Wu et al. 2001; Zhou and Thompson 2002; Gentry et al. 2006; Wu et al. 2006; He et al. 2007; Wagner et al. 2007; Zhou et al. 2008; Wang et al. 2009) unlike PCR-based techniques, which limit the number of genes that can be examined at once. The focus of this chapter will be the development and application of FGAs.

Functional Gene Array Development

The first reported FGA contained ~100 PCR-amplicon probes targeting N-cycling genes (*nirS*, *nirK*, *amoA*, and *pmoA*) (Wu et al. 2001). However, the use of PCR-based probes limits the comprehensiveness of an array since a very large number of diverse bacterial strains and environmental clones would be required. Another issue with the use of PCR probes is the need to develop conserved primers for each gene or gene group, which would pose a problem as no conserved PCR primers can be designed for many functional genes of interest. In addition, although fairly conserved

primers can be designed for some functional genes of interest, it is sometimes difficult to amplify these genes from environmental samples. To overcome these technical challenges, oligonucleotide probes have been used instead of PCR amplicons. Oligonucleotide probes have higher specificity but lower sensitivity than PCR-based probes (Zhou 2003), can be easily customized allowing more targeted probe design (Denef et al. 2003; Zhou 2003; Gentry et al. 2006), and are relatively inexpensive. As such, FGAs are often constructed using oligonucleotide probes.

In the decade since the first FGA was reported, several different FGAs have been developed (Cho and Tiedje 2002; Bodrossy et al. 2003; Rhee et al. 2004; Gentry et al. 2006; Zhang et al. 2006). Some have targeted specific functional groups or genes, such as antibiotic resistance (Call et al. 2003), organic contaminant degradation and metal resistance (Rhee et al. 2004), N-cycling (Taroncher-Oldenburg et al. 2003; Steward et al. 2004; Tiquia et al. 2004), methanotrophs (Stralis-Pavese et al. 2004), virulence factors and pathogen-specific markers (Miller et al. 2008; Palka-Santini et al. 2009), *nodC* variants (Bontemps et al. 2005), or specific locations like acid mine drainage sites (Yin et al. 2007). To date, the most comprehensive FGA reported is the GeoChip 2.0 (He et al. 2007), a high density FGA, with 24,243 50-mer oligonucleotide probes, targeting ~10,000 functional genes from 150 gene families involved in the geochemical cycling of C, N, and P cycling, sulfate reduction, metal reduction and resistance, and organic contaminant degradation. The GeoChip was designed to provide sufficient oligonucleotide probe specificity for genes that have high homologies and to provide a truly comprehensive FGA probe set, both of which were lacking in previous FGAs (He et al. 2007). A newer version, GeoChip 3.0, which covers ~47,000 sequences from 292 gene families, covering twice as many functional gene groups as GeoChip 2.0, including the phylogenetic marker *gyrB*, has been developed (He et al. 2010a).

Comparison of FGA to Other High-Throughput Genomic Technologies

Several high-throughput methods are available for microbial community studies in addition to GeoChip, including the PhyloChip and barcode-based high-throughput sequencing. In contrast to 16S rRNA gene-based microarrays (e.g., PhyloChip) and sequencing technologies (e.g., 454), functional gene arrays (e.g., GeoChips) have several advantages: (a) Detecting functions. While PhyloChip is a powerful tool for examining microbial communities, this array only has probes for the 16S rRNA gene to detect the presence of a strain but not its functional activity. Sequencing allows for obtaining hundreds of thousands of sequences, but requires an initial PCR step. As mentioned above, the use of PCR primer based methods are limited since primers are only available for a limited number of functional genes. In addition, the sheer number of reactions would be prohibitive. Therefore, an examination of a wide variety of functional genes of interest is nearly impossible with this method. FGAs, however, use functional gene markers and thus provide information

on the potential metabolic functions of a community (Wu et al. 2001; He et al. 2007). As such, GeoChip, which contains probes for thousands of functional genes, is particularly useful in linking microbial community structure to community function (He et al. 2007; 2010a). (b) Higher resolution. GeoChip can provide resolution at the species-strain level (Tiquia et al. 2004); whereas the resolution of PhyloChip is at the family-subfamily level (DeSantis et al. 2007) and 16S rRNA gene-based sequencing can generally provide resolution at the genus-species level (He Z et al., unpublished data). (c) Quantitation. Many ecological studies require quantitative information regarding microbial abundance. Since PhyloChip and sequencing require PCR amplification steps, these techniques may not provide quantitative results because amplification bias is a well-known phenomenon in PCR (Warnecke et al. 1997; Lueders and Friedrich 2003; Suzuki and Giovannoni 1996). In contrast, GeoChip does not rely on PCR amplification for detection. Previous studies with FGAs have shown that hybridization of template DNA with or without random amplification are quite quantitative (Wu et al. 2001; Tiquia et al. 2004; Rhee et al. 2004; Wu et al. 2006; Gao et al. 2007). In addition, whole community genome amplification (WCGA), which can be used to increase the amount of DNA available for hybridization has been shown to produce minimal bias (Wu et al. 2006).

In addition to the process differences discussed above, these technologies also have advantages and disadvantages specific to the study of microbial community structure and dynamics. These include (a) Random sampling errors. Although metagenomic technologies are able to provide a lot of data, only a small portion of the microbial community is actually sampled in most studies. Mckenna et al. (2008) obtained only ~1,400 rRNA gene sequences per sample from the gut microbiome of Rhesus macaques using 454 sequencing, an underestimate of the gut community based on Chao1 estimates. With an estimated density of 10^{11}–10^{12} bacterial cells per mL (Whitman et al. 1998) or 10^{9}–10^{10} g^{-1} of stool (Palmer et al. 2007) in the human gut, results of 454 pyrosequencing would be expected to greatly underestimate the microbial community. If the sampling process is completely random, theoretically, the probability of sampling the same portion of a community over multiple sampling events would be small (Zhou et al. 2008). Although dominant populations would, in all probability, be sampled multiple times, it is still not possible to ensure that the same populations of a microbial community are measured across different sampling events. As such, the estimated species richness would be subjected to random sampling errors. In contrast, microarray-based approaches compare all samples against the same set of probes (i.e., those contained on the array), ensuring that the same population are sampled for comparison across all samples in a study. As a result, the artifact due to the nature of random sampling can be minimized if not eliminated. (b) Relative abundance. Unlike pure cultures, the abundance of different species in a microbial community varies greatly. Sequencing-based approaches will be very sensitive to the distribution of species abundance. For instance, if the abundance of species A is 5% of the abundance of species B; theoretically, only one of the 20 molecules sequenced will be from species A while the remainder would be from species B. To have a 1x sequence coverage for species A, 20-fold more sequencing effort would be required. Thus, detecting all of the rare species within a community would not be cost effective or feasible even with next-generation sequencing technologies.

Microarray-based detection approaches are not affected by the relative abundance of species as long as their abundance is above the detection limit. (c) Detecting new sequences. Sequenced-based technologies can find new or novel sequences since there are no limitations on what sequences are sequenced. However, because microarrays are limited to detecting only sequences covered by the probe sequences on an array, detecting new sequences is impossible. (d) Community comparison. Randomly sequencing a small portion of the microbial communities from different environments or conditions may not be informative or meaningful for comparative purposes due to random sampling errors (Zhou et al. 2008), as discussed above, unless sufficient sequencing coverage is achieved. In contrast, since microarrays interrogate communities with the same set of probes across samples, comparisons between samples and environments can easily be made. (e) Cost. In addition, although sequencing technology has developed rapidly and the cost per base pair has decreased considerably, sequencing capacity and cost are still limiting factors when entire microbial communities and/or multiple communities are considered. In contrast, after the initial output for printing and imaging equipment, microarray analysis is much less expensive than 454 pyroequencing, even for multiple samples with a barcode approach.

Design and Development of Geochip

Probe Design

A flowchart of the basic design protocol is shown in Fig. 5.1. First, specific functional genes for key functional processes of interest are selected. Genes should be

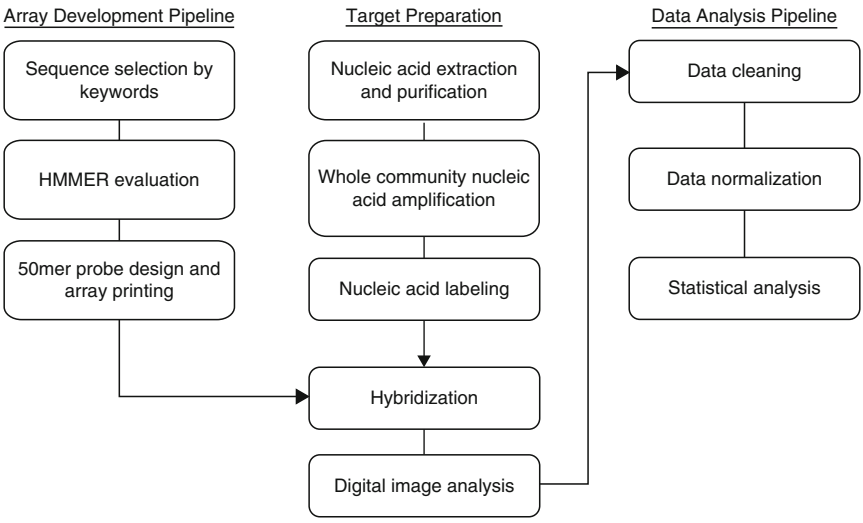

Fig. 5.1 Major steps for GeoChip design and use. See text for full explanation of all steps

chosen for key enzymes or proteins that are vital to pathways or functions of interest. Public databases (e.g., GenBank) are searched automatically using selected keywords and resulting sequences are downloaded. Care should be taken in selecting keywords as genes may be annotated differently in different microorganisms or may have a more general or specific description so using very broad key words is often best. Second, the downloaded sequences are evaluated using HMMER alignment (http://hmmer.wustl.edu/) with seed sequences, which have had protein identity and function and experimentally confirmed. The selection of seed sequences is a critical step in probe design and care should be taken in choosing appropriate sequences. The seed sequences are stored in a database for later array updates. Sequences passing HMMER alignment are deposited to a local sequence database. Third, gene-specific or group-specific 50-mer oligonucleotide probes are designed with CommOligo (Li et al. 2005) using experimentally determined criteria based on sequence homology (≤90% identity for gene-specific; ≥96% for group-specific), continuous stretch length (≤20 bases for gene-specific; ≥35 for group-specific), and free energy (≥ −35 kJ mol^{-1} for gene-specific; ≤ −60 kJ mol^{-1} for group-specific (He et al. 2005b; Leibich et al. 2006). In addition, to ensure specificity, all designed probes are screened against the GenBank database. Finally, the resultant probes are then commercially synthesized and used for array construction. Probes can be spotted onto glass slides (Taroncher-Oldenburg et al. 2003; Tiquia et al. 2004; Rhee et al. 2004) or nylon membranes (Steward et al. 2004). Glass slides are generally used since they produce less background fluorescence (Schena et al. 1995, 1996) and allow higher probe density (Ehrenreich 2006).

Target Preparation

An important factor in obtaining reliable microarray data is to use high quality DNA or RNA. The key steps in target preparation are shown in Fig. 5.1. Microbial community DNA from environmental samples is generally extracted and purified using a well-established freeze-grind method since it results in large fragments of genomic DNA (Zhou et al. 1996; Hurt et al. 2001) which are important if the DNA needs to be amplified. The purified DNA should have A_{260}:A_{280} > 1.8 and A_{260}:A_{230} > 1.7. We have had success with the use of agarose gel purification followed by a phenol-chloroform-butanol extraction (Liang et al. 2009b). Impurities in the DNA can inhibit subsequent amplification, labeling and hybridization processes. If < 2.0 μg of DNA is obtained, WCGA can be used to increase the amount of DNA available for hybridization with small quantities of DNA (1–100 ng) (Wu et al. 2006). WCGA provides a sensitive (10 fg detection limit) and representative amplification (<0.5% of amplified genes showed >twofold different from unamplified) (Wu et al. 2006).

While the use of DNA provides information on the community structure and the functional potential of the microbial community, it does not provide information on the activity of the community. To examine activity of microbial communities, mRNA can also be used with FGAs. However, two major challenges in using RNA

are the relatively low abundance of mRNA in environmental samples, and short turnover rates of mRNA. Several methods are available for extraction of community RNA from environmental samples, including simultaneous extraction of DNA and RNA (Hurt et al. 2001), RNA extraction via a bead-beating method (Burgmann et al. 2003), or use of gel electrophoresis to isolate mRNA from total RNA (McGrath et al. 2008). Purified RNA should have A_{260}:A_{280} >1.90 and A_{260}:A_{230} >1.70. Since only a small portion of the total RNA is mRNA, a large quantity of RNA (10–20 μg) is required for hybridization. However, environmental samples often do not provide a sufficient quantity of RNA, so whole community RNA amplification (WCRA) (Gao et al. 2007) may be required. WCRA employs a fusion primer comprised of a short (6–9) set of random nucleotides and a T7 promoter. Amplification of 50–100 ng of total RNA resulted in a representative amplification that maintained the original relationship of mRNA (Gao et al. 2007). Another option is the use of stable isotope probing of active community members (Leigh et al. 2007).

The DNA or RNA is then labeled with fluorescent dyes (e.g., Cy3, Cy5). For DNA, random priming with the Klenow fragment of DNA polymerase is used (Wu et al. 2006). RNA is labeled using Superscript™ II/III RNase H-reverse transcriptase (He et al. 2005b). The labeled nucleic acids are then purified and dried for hybridization.

Hybridization

Labeled DNA or RNA is suspended in hybridization buffer for hybridization. GeoChips can be hybridized at 42–50°C and 50% formamide (He et al. 2007; Mason et al. 2009; Liang et al. 2009a, b; Van Nostrand et al. 2009; Waldron et al. 2009). The hybridization temperature and formamide concentration can be adjusted to increase or decrease stringency in order to detect more or less diverse sequences. The effective hybridization temperature can be increased by the use of formamide (0.6°C for every 1%).

Hybridizations using glass arrays can be carried out manually or using automated or semi-automated hybridization stations. Manual hybridizations are performed using a water bath or hybridization oven and specially designed hybridization chambers that help maintain humidity levels within the chamber. Several hybridization stations provide incubation at controlled temperatures and mixing (e.g., Mai Tai® from SciGene, SlideBooster from Advalytix, Maui from BioMicro Systems). Washing after hybridization can be accomplished manually or using an automated wash station (e.g., Maui Wash Station, BioMicro Systems). Other systems are completely automated from pre-hybridization through post-hybridization washes (e.g., Tecan HS4800Pro, TECAN US).

Image Analysis

After hybridization, the array is imaged using a microarray scanner with a resolution of 10 μm or better. The image is then digitally analyzed by quantifying

the pixel density (intensity) of each spot using microarray-analysis software. The analysis software can also be used to evaluate spot quality using predetermined criteria and flag poor or low quality spots for later removal. Distinguishing a positive spot from background noise is generally based on signal-to-noise ratio [SNR; SNR = (signal mean – background mean)/background standard deviation]. However, other calculations can be used instead of SNR. He and Zhou (2008) developed a signal-to-both-standard-deviations ratio [SSDR; SSDR = (signal mean – background mean)/(signal standard deviation – background standard deviation)] which resulted in fewer false positives and negatives than the SNR calculation.

This raw data is then uploaded to the GeoChip data analysis pipeline (http://ieg.ou.edu/) and evaluated. The quality of individual spots, evenness of control spot hybridization signals across the slide surface, and background levels are assessed. Poor and low quality spots are removed along with outliers. Outliers are determined based on the signal intensities of replicate arrays and are defined as those positive spots with (signal – mean signal intensity of all replicate spots) is greater than three times the replicate spots' signal standard deviation (He and Zhou 2008). The signal intensities are then normalized and the data is stored in an experiment database for further statistical analysis using the data analysis pipeline.

Data Analysis

The most difficult task with FGAs, especially GeoChip, is data analysis due to the seemingly overwhelming amount of data obtained. A few data analysis methods have been used frequently and include relative abundance of gene groups based on gene number or total signal intensity, richness and diversity indices based on gene number, percent of gene overlap between samples, and response ratios. Methods commonly used for statistical analysis of microarray data include principal component analysis (PCA), cluster analysis (CA), and neural network analysis (NNA) (He et al. 2008). PCA is a multivariate statistical method which reduces the dimensionality of variables to maximize the visible variability of the data. The major advantage of PCA is that it identifies outliers (e.g., genes) in the data set that behave differently from most of the genes across a set of experiments. CA is used to identify groups with similar gene profiles, and it can help establish functionally related groups of genes to gain insights into structure and function of a given microbial community. NNA is a relatively new analysis technique for FGA data but can be used to examine gene relationships. Response ratios, which compare community response (e.g., gene levels or signal intensity) between conditions (e.g., control versus treatment, contaminated versus uncontaminated) (Luo et al. 2006), has been used to compare the community response to varying levels of oil contamination (Liang et al., 2009a). In addition, if environmental variables are available, canonical correspondence analysis (CCA) (ter Braak 1986), variation partitioning analysis (VPA) (Økland and Eilertsen 1994; Ramette and Tiedje 2007), and other correlation analyses (e.g., Mantel test) can be used to correlate environmental conditions

with the community structure for further understanding of the relationship between the microbial community and ecosystem functioning. CCA has been used in several GeoChip-based studies to better understand how environmental factors are affecting community structure (Yergeau et al. 2007; Wu et al. 2008; Zhou et al. 2008; Waldron et al. 2009; Van Nostrand et al. 2009). VPA is used to determine the relative influence of environmental parameters on the microbial community structure and is based on results of the CCA. The Mantel test has been used to correlate environmental factors with functional genes detected with GeoChip (He et al. 2007; Wu et al. 2008; Van Nostrand et al. 2009; Waldron et al. 2009).

Important Issues for Microarray Application

A great deal of progress has been made over the past decade with regards to the development of microarray technology for studying environmental communities (Wu et al. 2001, 2004, 2006; Adey et al. 2002; Rhee et al. 2004; Leibich et al. 2006; Gao et al. 2007; He et al. 2007; He and Zhou 2008). However, several challenges and key issues remain.

Nucleic acid quality. One of the most important steps for successful FGA analysis is obtaining high-quality DNA or/and RNA from environmental samples. Our lab has successfully used an established freeze-grind extraction method (Zhou et al. 1996; Hurt et al. 2001) followed by agarose gel purification (Liang et al. 2009b). However, some samples are still difficult to purify to the necessary level. In addition, gel purification only works for fresh DNA and for samples which yield a relatively large amount of DNA, since some portion of DNA may be lost with this method. Use of mRNA presents an even greater challenge. Isolation of mRNA from environmental samples is difficult due to the low abundance and instability of the mRNA. Very few studies have used FGAs for environmental mRNA analysis (Hurt et al. 2001; Dennis et al. 2003; Gao et al. 2007).

Probe coverage. One of the drawbacks of earlier FGA versions is that they lacked a comprehensive probe set and focused on only a few genes or gene groups. GeoChip 2.0 provided the most comprehensive FGA currently available and covered >150 gene groups (He et al. 2007). GeoChip 3.0 provides coverage of 292 gene groups and about four times as many genes as GeoChip 2.0 (He et al. 2010a). However, regardless of how comprehensive the GeoChip is now, sequences are constantly being added to public databases, leading to an exponential increase in the number of functional genes as well as the number of sequences for each particular functional gene. As such, continual updates of GeoChip are necessary. The probe design system used in our lab has an automatic update feature using predetermined keywords and seed sequences. However, even with the advances in probe design software, this process is still time consuming due to the large number of sequences and probes that must be designed and tested.

Specificity. Specificity is an important attribute of gene probes, especially for those designed to analyze environmental samples. However, a difficulty in designing

specific probes is that so many environmental sequences are unknown, and that many homologous genes are highly similar. Careful design of oligonucleotide probes can provide highly specific hybridizations. Criteria based on similarity, stretch and free energy have been used to design specific probes (He et al. 2005a, 2007; Leibich et al. 2006). Our evaluation of GeoChip 2.0 probes designed using these criteria revealed that a minimal number of false positives (0.002–0.004%) were observed (He et al. 2007).

In addition to probe design criteria, specificity can be adjusted by changing the hybridization conditions to increase or decrease stringency. Hybridization stringency is generally controlled by temperature or formamide concentration. At 65°C, hybridization occurred for sequences with similarities >87% (Wu et al. 2001). At 45°C, hybridization occurred for sequences with similarities as low as 70–75% (Wu et al. 2001). Hybridizations using 50-mer oligonucleotide FGAs at 50°C and 50% formamide (effective temperature, 80°C) were able to discriminate sequences with <88 to 94% similarities (Rhee et al. 2004; Leibich et al. 2006; Deng et al. 2008).

Signal intensity can be affected by sequence divergence as well as by sequence abundance; therefore, strategies need to be developed to determine which condition is occurring. One option to determine a true signal is the use of mismatched probes. Deng et al. (2008) found that probes with mismatches (3–5) distributed evenly across the probe were better able to distinguish perfect-matched targets versus mismatched targets than randomly distributed mismatched nucleotides. In addition, using relative comparisons across samples (i.e., comparing signal intensities from test samples to a control or background sample) rather than absolute comparisons will minimize or eliminate the effects of potential cross-hybridization (He et al. 2007). Assuming that test and control samples have similar community composition, using the ratios of test to control samples will cancel out any cross-hybridization (He et al. 2007).

Sensitivity. Another important aspect of microarrays is sensitivity, especially for environmental samples which often have complex communities with many strains in low abundance. Based on current FGA technology, the detection limit is 5% of the microbial community (Bodrossy et al. 2003), which provides coverage for only the dominant community members. PCR-based probes had sensitivity of 1 ng of pure genomic DNA or 25 ng of community DNA (Wu et al. 2001). Similar detection limits were observed for 50 mer oligonucleotide probes (Rhee et al. 2004; Tiquia et al. 2004). Several strategies could be used to increase sensitivity although these also decrease specificity. For example, increasing the length of probes, increases sensitivity (Denef et al. 2003; He et al. 2005a), but at the cost of specificity (Relógio et al. 2002). Another strategy that has been suggested is to increase the amount of probe per spot (Cho and Tiedje 2002; Relógio et al. 2002; Zhou and Thompson 2002) since membrane based arrays are generally more sensitive, due to the higher probe concentrations on the array surface (>1 μg/spot for membranes; <20 pg/spot for glass slides) (Cho and Tiedje 2002). Although increasing the probe concentration may result in a lower signal intensity (Denef et al. 2003), which would effectively counteract any gain in sensitivity.

In addition to probe or array design strategies, sensitivity can also be increased by utilizing several sample preparation and hybridization strategies. Whole community genome amplification (WCGA) can increase the concentration of all community

DNA; including low abundance sequences (Wu et al. 2006). WCGA is able to representatively amplify 1 to 250 ng of community DNA (Wu et al. 2006). This amplification method showed a detection limit of 10 pg of the original DNA, although with a much higher amplification bias than observed with 1 ng of DNA. Amplification via multiplex PCR, using primers for all genes contained on an array, has been used to increase the amount of array-specific DNA for a pathogen array (Palka-Santini et al. 2009). While this strategy may work for some FGAs, especially those focused on relatively small numbers of genes, this would not work for GeoChip because of the difficulty in designing primers for all genes on the array. Another option is the use of more sensitive labeling techniques. Using cyanine dye-doped nanoparticles (Zhou and Zhou 2004) or tyramide signal amplification labeling (Denef et al. 2003) can increase sensitivity up to tenfold. A final strategy is to develop more sensitive signal detection systems (Cho and Tiedje 2002; Zhou and Thompson 2002).

Quantitative applications. A major goal for microarray analysis is to be able to provide quantitative information. Some studies have shown a correlation between signal intensity and DNA concentration. PCR probes showed a correlation ($r = 0.94$) between signal intensity and DNA quantity over the concentration range of 0.5–100 ng (Wu et al. 2001). Oligonucleotide probes (50-mer) provided a linear relationship ($r = 0.98$–0.99) over a concentration range of 8–1,000 ng (Tiquia et al. 2004). Hybridizations with RNA have also been shown to be linear over the range of 50–100 ng (Gao et al. 2007).

Activity. Using DNA with FGAs provides information on population changes which can be used to infer microbial activity; but does not provide absolute evidence of that activity. The use of community mRNA would provide information on which community members or functional processes are active, similar to transcriptional arrays for pure cultures (Dennis et al. 2003; Bodrossy et al. 2006; Gao et al. 2007). However, our current difficulties with community mRNA are the low amount and the stability of environmental mRNA. WCRA overcomes the problem of low abundance (Gao et al. 2007) although the current amplification protocol is complex and time consuming. Therefore, improved methods for RNA extraction from environmental samples and mRNA amplification are needed. Another option for determining microbial activity is SIP analysis (Leigh et al. 2007).

Application of GeoChip for Microbial Community Analysis

The GeoChip has been used in numerous studies to examine the functional community structure and dynamics of microbial communities. Most of these studies have utilized community DNA to measure gene abundance although RNA or stable isotope probing (SIP) can also be used to examine gene expression. These studies have shown the power of GeoChip to link microbial community functional structure to biogeochemical, ecological, and environmental processes.

U(VI) contaminated environments. GeoChip has been used in several studies to examine U-contaminated groundwater at the US DOE's Field Research Center

(FRC) in Oak Ridge, TN. GeoChip 1.0, a prototype array containing 2006 (50-mer) oligonucleotide probes (Rhee et al. 2004; Tiquia et al. 2004), was used to examine communities within the FRC, which is contaminated with nitrate, uranium, and organic compounds. This study examined samples from contaminated and uncontaminated areas and observed higher gene numbers in uncontaminated sites compared to the contaminated sites, indicating the deleterious effect of the contaminants on the microbial communities (Wu et al. 2006). Using GeoChip 1.0, another study examined samples with a range of contamination levels and found that more genes were detected from the uncontaminated control site compared to the numbers detected from contaminated wells (Waldron et al. 2009). Microbial communities were examined within a pilot-scale test system established for the biostimulation of U(VI) reduction in the subsurface by injection of ethanol. A significant correlation ($r = 0.73$, $p < 0.05$) was observed between the U(VI) concentration and the amount of cytochrome genes detected, indicating the importance of cytochrome containing microorganisms in U(VI) reduction at this site (He et al. 2007). In the same bioremediation system, the effects of dissolved oxygen (DO) and ethanol amendment on the microbial community were examined, and the results showed that ethanol was a much stronger driver in controlling community structure than U(VI) or DO (Van Nostrand et al. 2009).

Hydrocarbon contaminated sites. GeoChip 2.0 was used to examine the microbial community of a bioremediation system designed for the remediation of diesel fuel in Vega Baja, Puerto Rico (Rodríguez-Martínez et al. 2006). Genes involved in the degradation of diesel and organic contaminants (phthalate, biphenyl, cyclohexanol, benzoate, and naphthalene degradation genes) were detected. The amount of anaerobic degradation genes increased over time, suggesting that, consistent with other evidence, the system shifted to an anaerobic process. Liang et al. (2009b) used GeoChip 2.0 to examine the microbial community of contaminated oil fields before and after bioremediation treatment. Ozonation treatment resulted in a decrease in most functional gene categories, including carbon, nitrogen, and sulfur cycling and organic contaminant degradation genes, but all gene categories recovered after treatment. Another study examined microbial communities from a contaminated oil field in China and found that higher levels of oil contamination resulted in lower diversity and a decreased number of functional genes detected (Liang et al., 2009a). In addition, genes associated with oil degradation, including catechol, protocatechuate, biphenyl degradation, increased under a moderate oil contamination level.

Soil communities. GeoChip 1.0 was used to examine functional changes of microbial communities under different land use strategies and found that diversity and functional gene numbers increased as soil organic carbon increased (Zhang et al. 2007). In another study, Yergeau et al. (2007) examined Antarctic sediments and found that cellulose degradation and denitrification genes were positively correlated with soil temperature. Additionally, Zhou et al. (2008) used GeoChip 2.0 to assess the gene-area relationship of microbial communities of forest soils, and the results suggest that the forest soil microbial community demonstrated a relatively flat gene-area relationship with less turnover than observed for plants and animals (Zhou et al. 2008).

Marine environments. Sediments from the Gulf of Mexico were examined using GeoChip 1.0 (Wu et al. 2008). Genes for carbon degradation, nitrification, denitrification, nitrogen fixation, sulfur reduction, phosphorus utilization, contaminant degradation, and metal resistance were detected and communities become more distinctive as depth increased. The environmental parameters, depth, porosity, and concentrations of ammonium, phosphate, Mn(II), and silicic acid appeared to be important drivers in determining the structure of the microbial communities in this environment. Another study characterized microbial communities from deep sea hydrothermal vents, including a mature chimney and the inner and outer portions of a 5-day-old chimney with GeoChip 2.0, and the results showed communities from the inner chimney were less diverse than those from the outer portion of the 5-day-old chimney or the mature chimney (Wang et al. 2009). GeoChip 2.0 has also been used to examine microbial communities from deep sea basalt and genes involved in carbon fixation, methane oxidation, methanogenesis, and nitrogen fixation, processes not previously associated with this environment, were detected (Mason et al. 2009).

Climate change. The latest version of GeoChip (GeoChip 3.0) has been used to study the effects of elevated CO_2 on microbial communities at a multifactor grassland experiment site, BioCON (Biodiversity, CO_2, and Nitrogen deposition) at the Cedar Creek Ecosystem Science Reserve in Minnesota, USA. This study suggests that elevated CO_2 significantly affects the soil microbial community (He et al., 2010b). In addition, the soil microbial community appears to influence global carbon and nitrogen cycling and may moderate the effects of climate change. These studies show the power and utility of GeoChip for analyzing microbial community functional structure from a variety of environments.

Other. In addition to these community DNA-based studies, several other studies have used GeoChip 2.0 to examine pure cultures and microbial activity within communities. Van Nostrand et al. (2007) used GeoChip 2.0 to probe four metal-resistant (Ni, Co, Cd, Zn) actinomycetes for metal resistance genes. Multiple metal resistance genes were detected including some implicated in Ni, Co, Cd, and Zn resistance. Leigh et al. (2007) used stable isotope probing (biphenyl) in conjunction with GeoChip to detect active PCB-degrading microbial populations within a hydrocarbon-contaminated aquifer. Gao et al. (2007) used amplified community mRNA to examine the activity of microbial communities from a denitrifying fluidized bed reactor at a uranium contaminated site. Genes expected to be active at this site were detected including nitrate reduction genes and several organic contaminant degradation genes.

Summary

Over the past decade great advances have been made in microarray technology and in FGA development. The GeoChip 2.0 has garnered a great deal of attention, and numerous studies over the last couple of years have demonstrated its applications

in the study of microbial ecology and linking microbial communities and geochemistry. These studies have demonstrated GeoChip's ability to provide sensitive, specific, and potentially quantitative information concerning microbial communities from a wide range of environments. This high-throughput, cutting edge technology is expected to revolutionize the field of microbial ecology and the study of microbial community functional structure and dynamics.

However, there are still technical, experimental and analysis challenges that need to be overcome. These include increasing sensitivity either through new technologies and methods in array printing or novel and/or improved target labeling methods to better detect functional genes at a low abundance. Strategies must be developed to improve the quantitative accuracy of FGA hybridizations. Bioinformatic tools and techniques are needed to assist in analysis, evaluation, and interpretation of the vast amounts of data resulting from FGA analysis. New tools are also needed for sequence retrieval, evaluation, and probe design. Novel analytical techniques are needed to fully utilize the FGA data. In addition, strategies and techniques must be developed in order to be able to compare data sets across samples, experiments, and labs.

Acknowledgements The effort for preparing this review was supported by the Virtual Institute for Microbial Stress and Survival (http://VIMSS.lbl.gov) supported by the U. S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomics Program:GTL through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U. S. Department of Energy, Environmental Remediation Science Program (ERSP), Office of Biological and Environmental Research, Office of Science, and Oklahoma Applied Research Support (OARS), Oklahoma Center for the Advancement of Science and Technology (OCAST), the State of Oklahoma through the Project AR062-034.

References

Adey NB, Lei M, Howard MT et al (2002) Gains in sensitivity with a device that mixes microarray hybridization solution in a 25-μm-think chamber. Anal Chem 74:6413–6417

Bodrossy L, Stralis-Pavese N, Murrell JC, Radajewski S, Weilharter A, Sessitsch A (2003) Development and validation of a diagnostic microbial microarray for methanotrophs. Environ Microbiol 5:566–582

Bodrossy L, Stralis-Pavese N, Konrad-Köszler M, Weilharter A, Reichenauer TG, Schöfer SA (2006) mRNA-based parallel detection of active methanotroph populations by use of a diagnostic microarray. Appl Environ Microbiol 72:1672–1676

Bontemps C, Goldier G, Gris-Liebe C, Carere S, Talini L, Boivin-Masson C (2005) Microarray-based detection and typing of the rhizobium nodulation gene *nodC*: potential of DNA arrays to diagnose biological functions of interest. Appl Environ Microbiol 71:8042–8048

Braker G, Fesefeldt A, Witzel KP (1998) Development of PCR primer systems for amplification of nitrite reductase genes (*nirK* and *nirS*) to detect denitrifying bacteria in environmental samples. Appl Environ Microbiol 64:3769–3775

Brodie EL, DeSantis TZ, Joyner DC et al (2006) Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. Appl Environ Microbiol 72:6288–6298

Burgmann H, Widmer F, Sigler WV, Zeyer J (2003) mRNA extraction and reverse transcription-PCR protocol for detection of *nifH* gene expression by *Azotobacter vinelandii* in soil. Appl Environ Microbiol 69:1928–1935

Call DR, Bakko MK, Krug MJ, Roberts MC (2003) Identifying antimicrobial resistance genes with DNA microarrays. Antimicrob Agents Chemother 47:3290–3295

Cho JC, Tiedje JM (2002) Quantitative detection of microbial genes by using DNA microarrays. Appl Environ Microbiol 58:1425–1430

Denef VJ, Park J, Rodrigues JLM, Tsoi TV, Hashsham SA, Tiedje JM (2003) Validation of a more sensitive method for using spotted oligonucleotide DNA microarrays for functional genomics studies on bacterial communities. Environ Micrbiol 5:933–943

Deng Y, He Z, Van Nostrand JD, Zhou J (2008) Design and analysis of mismatch probes for long oligonucleotide microarrays. BMC Genomics 9:491–503

Dennis P, Edwards EA, Liss SN, Fulthorpe R (2003) Monitoring gene expression in mixed microbial communities by using DNA microarrays. Appl Environ Microbiol 69:769–778

DeSantis TZ, Brodie EL, Moberg JP, Zubieta IX, Piceno YM, Andersen GL (2007) High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. Microb Ecol 53:371–383

Ehrenreich A (2006) DNA microarray technology for the microbiologist: an overview. Appl Microbiol Biotechnol 73:255–273

Gao H, Yang ZK, Gentry TJ, Wu L, Schadt CW, Zhou J (2007) Microarray-based analysis of microbial community RNAs by whole-community RNA amplification. Appl Environ Microbiol 73:563–571

Gentry TJ, Wickham GS, Schadt CW, He Z, Zhou J (2006) Microarray application in microbial ecology research. Microb Ecol 52:159–175

Guschin DY, Mobarry BK, Proudnikov D, Stahl DA, Rittmann BE, Mirzabekov AD (1997) Oligonucleotide microchips as genosensors for determinative and environmental studies in microbiology. Appl Environ Microbiol 63:2397–2402

He Z, Zhou J (2008) Empirical evaluation of a new method for calculating signal to noise ratio (SNR) for microarray data analysis. Appl Environ Microbiol 74:2957–2966

He Z, Wu L, Fields MW, Zhou J (2005a) Use of microarrays with different probe sizes for monitoring gene expression. Appl Environ Microbiol 71:5154–5162

He Z, Wu LY, Li XY, Fields MW, Zhou JZ (2005b) Empirical establishment of oligonucleotide probe design criteria. Appl Environ Microbiol 71:3753–3760

He Z, Gentry TJ, Schadt CW et al (2007) GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. ISME J 1:67–77

He Z, Van Nostrand JD, Wu L, Zhou J (2008) Development and application of functional gene arrays for microbial community analysis. T Nonferr Metal Soc 18:1319–1327

He Z, Deng Y, Van Nostrand JD, Tu Q, Xu M, Hemme CL, Li X, Wu L, Gentry TJ, Yin Y, Leibich J, Hazen TC, Zhou J (2010a) GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure, and functional activity. ISME J doi:10.1038/ismej.2010.46

He Z, Xu M, Deng Y, Kang S, Kellogg L, Wu L, Van Nostrand JD, Hobbie SE, Reich PB, Zhou J (2010b) Metagenomic analysis reveals a marked divergence in the functional structure of belowground microbial communities at elevated CO2. Ecol Lett 13:564–575

Hugenholtz P, Goebel BM, Pace NR (1998) Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. J Bacteriol 180:4765–4774

Hurt RA, Qiu X, Wu L, Roh Y, Palumbo AV, Tiedje JM, Zhou Z (2001) Simultaneous recovery of RNA and DNA from soils and sediments. Appl Environ Microbiol 67:4495–4503

Lee DY, Shannon K, Beaudette LA (2006) Detection of bacterial pathogens in municipal wastewater using an oligonucleotide microarray and real-time quantitative PCR. J Microbiol Meth 65:453–467

Leigh MB, Pellizari VH, Uhlík O, Sutka R, Rodrigues J, Ostrom NE, Zhou J, Tiedje JM (2007) Biphenyl-utilizing bacteria and their functional genes in a pine root zone contaminated with polychlorinated biphenyls (PCBs). ISME J 1:134–148

Li X, He Z, Zhou J (2005) Selection of optimal oligonucleotide probes for microarrays using multiple criteria, global alignment and parameter estimation. Nucleic Acids Res 33:6114–6123

Liang Y, Li G, Van Nostrand JD, He Z, Wu L, Deng Y, Zhang X, Zhou J (2009a) Microarray-based analysis of microbial functional diversity along an oil contamination gradient in oil field. FEMS Microbiol Ecol 70:324–333

Liang Y, Wang J, Van Nostrand JD, Zhou J, Zhang X, Li G (2009b) Microarray-based functional gene analysis of soil microbial communities in ozonation and biodegradation of crude oil. Chemosphere 75:193–199

Leibich J, Schadt CW, Chong SC, He Z, Rhee SK, Zhou J (2006) Improvement of oligonucleotide probe design criteria for functional gene microarrays in environmental applications. Appl Environ Microbiol 72:1688–1691

Loy A, Lehner A, Lee N, Adamczyk J, Meier H, Ernst J, Schleifer KH, Wagner M (2001) Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. Appl Environ Microbiol 68:5064–5081

Lueders T, Friedrich MW (2003) Evaluation of PCR amplification bias by terminal restriction fragment length polymorphism Analysis of small-subunit rRNA and *mcrA* genes by using defined template mixtures of methanogenic pure cultures and soil DNA extracts. Appl Envrion Microbiol 69:320–326

Luo Y, Hui D, Zhang D (2006) Elevated CO_2 stimulates net accumulations of carbon and nitrogen in land ecosystems: a meta-analysis. Ecology 87:53–63

Mason OU, DiMeo-Savoie CA, Van Nostrand JD, Zhou J, Fisk MR, Giovannoni SJ (2009) Prokaryotic diversity, distribution, and preliminary insights into their role in biogeochemical cycling in marine basalts. ISME J 3:231–242

McDonald IR, Kenna EM, Murrell JC (1995) Detection of methanotrophic bacteria in environmental samples with the PCR. Appl Environ Microbiol 61:116–121

McGrath KC, Thomas-Hall SR, Cheng CT, Leo L, Alexa A, Schmidt S, Schenk PM (2008) Isolation and analysis of mRNA from environmental microbial communities. J Microbiol Meth 75:172–176

McKenna P, Hoffmann C, Minkah N, Aye PP, Lackner A, Liu Z, Lozupone CA, Hamady M, Knight R, Bushman FD (2008) The Macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis. PLoS Pathog 4:e20

Miller SM, Tourlousse DM, Stedtfeld RD, Baushke SW, Herzog AB, Wick LM, Rouillard JM, Gulari E, Tiedje JM, Hashsham SA (2008) In situ-synthesized virulence and marker gene biochip for detection of bacterial pathogens in water. Appl Environ Microbiol 74: 2200–2209

Murray AE, Lies D, Li G, Nealson K, Zhou J, Tiedje JM (2001) DNA–DNA hybridization to microarrays reveals gene-specific differences between closely related microbial genomes. Proc Natl Acad Sci USA 98:9853–9858

Økland RH, Eilertsen O (1994) Canonical correspondence analysis with variation partitioning: some comments and an application. J Veg Sci 5:117–126

Palka-Santini M, Cleven BE, Eichinger L, Krönke M, Krut O (2009) Large scale multiplex PCR improves pathogen detection by DNA microarrays. BMC Microbiol 9:1

Palmer C, Bik EM, Digiulio DB, Relman DA, Brown PO (2007) Development of the human infant intestinal microbiota. PLoS Biol 5:e177

Ramette A, Tiedje JM (2007) Multiscale responses of microbial life in spatial distance and environmental heterogeneity in a patchy ecosystem. Proc Natl Acad Sci USA 104:2761–2766

Relógio A, Schwager C, Richter A, Ansorge W, Valcárcel J (2002) Optimization of oligonucleotide-based DNA microarrays. Nucleic Acids Res 30:e51

Rhee SK, Liu X, Wu L, Chong SC, Wan X, Zhou J (2004) Detection of genes involved in biodegradation and biotransformation in microbial communities by using 50-mer oligonucleotide microarrays. Appl Environ Microbiol 70:4303–4317

Rodríguez-Martínez EM, Pérez EX, Schadt CW, Zhou J, Massol-Deyá AA (2006) Microbial diversity and bioremediation of a hydrocarbon-contaminated aquifer (Vega Baja, Puerto Rico). Int J Environ Res Public Health 3:292–300

Sarkar SF, Gordon JS, Martin GB, Guttman DS (2006) Comparative genomics of host-specific virulence in *Pseudomonas syringae*. Genetics 147:1041–1056

Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 270:467–470

Schena M, Shalon D, Heller R, Chai A, Brown PO (1996) Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. Proc Natl Acad Sci USA 93:10614–10619

Sebat JL, Colwell FS, Crawford RL (2003) Metagenomic profiling: microarray analysis of an environmental genomic library. Appl Environ Microbiol 69:4927–4934

Small J, Call DR, Brockman FJ, Straub TM, Chandler DP (2001) Direct detection of 16S rRNA in soil extracts by using oligonucleotide microarrays. Appl Environ Microbiol 67:4708–4716

Steward GF, Jenkins BD, Ward BB, Zehr JP (2004) Development and testing of a DNA macroarray to assess nitrogenase (*nifH*) gene diversity. Appl Environ Microbiol 70:1455–1465

Stralis-Pavese N, Sessitsch A, Weilharter A, Reichenauer T, Riesing J, Csontos J, Murrell JC, Bodrossy L (2004) Optimization of diagnostic microarray for application in analysing landfill methanotroph communities under different plant covers. Environ Microbiol 6:347–363

Suzuki MT, Giovannoni SJ (1996) Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl Envrion Microbiol 62:625–630

Taroncher-Oldenburg G, Griner EM, Francis CA, Ward BB (2003) Oligonucleotide microarray for the study of functional gene diversity in the nitrogen cycle in the environment. Appl Environ Microbiol 69:1159–1171

ter Braak CJF (1986) Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67:1167–1179

Tiquia SM, Wu L, Chong SC, Passovets S, Xu D, Ying Xu, Zhou J (2004) Evaluation of 50-mer oligonucleotide arrays for detecting microbial populations in environmental samples. Biotechniques 36:1–8

Van Nostrand JD, Khijniak TV, Gentry TJ, Novak MT, Sowder AG, Zhou JZ, Bertsch PM, Morris PJ (2007) Isolation and characterization of four gram-positive nickel-tolerant microorganisms from contaminated sediments. Microb Ecol 53:670–682

Van Nostrand JD, Wu WM, Wu L et al (2009) GeoChip-based analysis of functional microbial communities during the reoxidation of a bioreduced uranium-contaminated aquifer. Environ Microbiol 11:2611–2626

Wagner M, Smidt H, Loy A, Zhou Z (2007) Unravelling microbial communities with DNA-microarrays: challenges and future directions. Microb Ecol 53:498–506

Waldron PJ, Wu L, Van Nostrand JD, Schadt C, Watson D, Jardine P, Palumbo T, Hazen TC, Zhou J (2009) Functional gene array-based analysis of microbial community structure in groundwaters with a gradient of contaminant levels. Environ Sci Technol 43:3529–3534

Wang F, Zhou H, Meng J, Peng X, Jiang L, Sun P, Zhang C, Van Nostrand JD, Deng Y, He Z, Wu L, Zhou J, Xiao X (2009) GeoChip-based analysis of metabolic diversity of microbial communities at the Juan de Fuca Ridge hydrothermal vent. Proc Natl Acad Sci USA 106:4840–4845

Warnecke PM, Stirzaker C, Melki JR, Millar DS, Paul CL, Clark SJ (1997) Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. Nucleic Acids Res 25:4422–4426

Whitman WB, Coleman DC, Wiebe WJ (1998) Prokaryotes: the unseen majority. Proc Natl Acad Sci USA 95:6578–6583

Wilson KH, Blitchington RB, Greene RC (1990) Amplification of bacterial 16S ribosomal DNA with polymerase chain reaction. J Clin Microbiol 28:1942–1946

Wilson M, DeRisi J, Kristensen HH, Imboden P, Rane S, Brown PO, Schoolnik GK (1999) Exploring drug-induced alterations in gene expression in mycobacterium tuberculosis by microarray hybridization. Proc Natl Acad Sci USA 96:12833–12838

Wilson WJ, Strout CL, DeSantis TZ, Stilwell JL, Carrano AV, Andersen GL (2002) Sequence-specific identification of 18 pathogenic microorganisms using microarray technology. Mol Cell Probes 16:119–127

Wu L, Thompson DK, Li G, Hurt RA, Tiedje JM, Zhou J (2001) Development and evaluation of functional gene arrays for detection of selected genes in the environment. Appl Environ Microbiol 67:5780–5790

Wu L, Thompson DK, Liu X, Fields MW, Bagwell CE, Tiedje JM, Zhou J (2004) Development and evaluation of microarray-based whole genome hybridization for detection of microorganisms within the context of environmental applications. Environ Sci Technol 38:6775–6782

Wu L, Liu X, Schadt CW, Zhou J (2006) Microarray-based analysis of submicrogram quantities of microbial community DNAs by using whole-community genome amplification. Appl Environ Microbiol 72:4931–4941

Wu L, Kellogg L, Devol AH, Tiedje JM, Zhou J (2008) Microarray-based characterization of microbial community functional structure and heterogeneity in marine sediments from the Gulf of Mexico. Appl Environ Microbiol 74:4516–4529

Yamamoto S, Harayama S (1995) PCR amplification and direct sequencing of gyrB genes with universal primers and their application to the detection and taxonomic analysis of Pseudomonas putida strains. Appl Environ Microbiol 61:1104–1109

Yergeau E, Kang S, He Z, Zhou J, Kowalchuk GA (2007) Functional microarray analysis of nitrogen and carbon cycling genes across an Antarctic latitude transect. ISME J 1:1–17

Yin H, Cao L, Qiu G, Wang D, Kellogg L, Zhou J, Dai Z, Liu X (2007) Development and evaluation of 50-mer oligonucleotide arrays for detecting microbial populations in Acid Mine Drainages and bioleaching systems. J Microbiol Meth 70:165–178

Zhang L, Srinivasan U, Marrs CF, Ghosh D, Gilsdorf JR, Foxman B (2004) Library on a slide for bacterial comparative genomics. BMC Microbiol 4:12–18

Zhang K, Martiny AC, Reppas NB, Barry KW, Malek J, Chisholm SW, Church GM (2006) Sequencing genomes from single cells by polymerase cloning. Nat Biotechnol 24:680–686

Zhang Y, Zhang X, Liu X, Xiao Y, Qu L, Wu L, Zhou J (2007) Microarray-based analysis of changes in diversity of microbial genes involved in organic carbon decomposition following land use/cover changes. FEMS Lett 266:144–151

Zhou J (2003) Microarrays for bacterial detection and microbial community analysis. Curr Opin Microbiol 6:288–294

Zhou J, Thompson DK (2002) Challenges in applying microarrays to environmental studies. Curr Opin Biotech 13:204–207

Zhou X, Zhou J (2004) Improving the signal sensitivity and photostability of DNA hybridizations on microarrays by using dye-doped core-shell silica nanoparticles. Anal Chem 76:5302–5312

Zhou J, Bruns MA, Tiedje JM (1996) DNA recovery from soils of diverse composition. Appl Environ Microbiol 62:316–322

Zhou J, Kang S, Schadt CW, Garten CT Jr (2008) Spatial scaling of functional gene diversity across various microbial taxa. Proc Natl Acad Sci USA 105:7768–7773