20 GeoChip: A High Throughput Genomic Tool for Linking Community Structure to Functions

J. D. Van Nostrand[1,2] · *Y. Liang*[1,3] · *Z. He*[1,2] · *G. Li*[3] · *J. Zhou*[1,2]
[1]Institute for Environmental Genomics, Department of Botany and Microbiology, University of Oklahoma, Norman, OK, USA
[2]Virtual Institute for Microbial Stress and Survival
joy.vannostrand@ou.edu
[3]Department of Environmental Science and Engineering, Tsinghua University, Beijing, China

K. N. Timmis (ed.), *Handbook of Hydrocarbon and Lipid Microbiology*, DOI 10.1007/978-3-540-77587-4_198,

Abstract: GeoChip is a comprehensive functional gene array that targets key functional genes involved in the geochemical cycling of N, C, and P, sulfate reduction, metal resistance and reduction, and contaminant degradation. Studies have shown the GeoChip to be a sensitive, specific, and high-throughput tool for microbial community analysis that has the power to link geochemical processes with microbial community structure. However, several challenges remain regarding the development and applications of microarrays for microbial community analysis.

1 Introduction

Microarrays allow for the examination of thousands of genes at one time in a rapid, high-throughput, and parallel manner. The first microarray reported was designed to examine gene expression in *Arabidopsis thaliana* (Schena et al., 1995). Since then, numerous types of microarrays have been developed and widely used for gene expression studies. The utility and power of microarrays was further expanded when the potential for microarrays to study microorganisms in the environment was tested (Guschin et al., 1997). Currently, three main types of arrays are being used for environmental studies (Gentry et al., 2006; Zhou, 2003; Zhou and Thompson, 2002). (1) Phylogenetic oligonucleotide arrays (POA) contain probes derived from 16S rRNA or other conserved genes designed to examine the phylogenetic relatedness of microorganisms or determine community composition. (2) Community genome arrays (CGA) use whole genomic DNA from microbial isolates to determine the relatedness of pure culture isolates or to probe for specific microorganisms or genes (Wu et al., 2004, 2008; Zhang et al., 2004). (3) Functional gene arrays (FGA) contain probes for genes that encode enzymes or proteins involved in specific functions of interest (Gentry et al., 2006; He et al., 2007; Wu et al., 2001). The primary advantage of FGAs is that they allow for the simultaneous examination of numerous functional gene groups at one time (Gentry et al., 2006; He et al., 2007; Wagner et al., 2007; Wu et al., 2001, 2006; Zhou et al., 2008; Zhou and Thompson, 2002) unlike traditional PCR-based molecular techniques, which allow examination of only a few genes at a time. Because FGAs provide information regarding the potential functional capabilities of microbial communities, this type of array is ideally suited for linking microbial communities with geochemical processes. In addition, compared to 16S rRNA gene-based phylogenetic arrays, FGAs can provide higher taxonomic resolution at the species-strain level (Tiquia et al., 2004). Due to their usefulness and versatility, this chapter will focus on the development and application of FGAs for examining microbial communities. Several recent reviews are available on the use of other types of microarrays for environmental studies (Gentry et al., 2006; Sessitsch et al., 2006).

2 Overview of GeoChip Analysis

The first generation of FGA, a prototype containing 100 functional genes (*nirS*, *nirK*, *amoA*, and *pmoA*) probes, was reported by Wu et al. (2001). Since this report, several other functional gene microarrays have been produced (Bodrossy et al., 2003; Cho and Tiedje, 2002; Gentry et al., 2006; Rhee et al., 2004; Zhang et al., 2006), including some designed for specific groups of interest such as N cycling (Tiquia et al., 2004), contaminant degradation and metal

resistance (Rhee et al., 2004), antibiotic resistance (Call et al., 2003), or for specific locations, such as acid mine drainage sites (Yin et al., 2007). The most comprehensive FGA reported to date is the GeoChip 2.0 (He et al., 2007), a high density FGA, which targets ~10,000 functional genes (~24,000 50-mer gene probes) from 150 gene families involved in the geochemical cycling of C, N, and S, metal reduction and resistance, and organic contaminant degradation. GeoChip 2.0 has been used to investigate biogeochemical, ecological and environmental processes (He et al., 2007; Leigh et al., 2007; Wu et al., 2008; Yergeau et al., 2007; Zhou et al., 2008). A new generation GeoChip (GeoChip 3.0) has been designed and is currently being tested and validated. This new GeoChip will have approximately 47,000 probes, representing 292 gene families (unpublished data). This array will target not only many important functional genes already on GeoChip 2.0, but also other functional genes such as the phylogenetic marker *gyrB*.

A diagram illustrating the protocol for array design, sample preparation, and data analysis is shown in *Fig. 1*. The first step is determining functional genes to be included on the array. These genes should encode for an enzyme or protein critical to the processes of interest, should be conserved while still providing enough sequence divergence to allow design of specific probes, and should have a large set of sequences available in public databases (Gentry et al., 2006). Selected keywords are used to automatically search public sequence databases (e.g., GenBank) and resulting sequences are downloaded based on the key words. To ensure that all included sequences are correct, the downloaded sequences are confirmed using HMMER alignment (http://hmmer.wustl.edu/) with seed sequences, which have had protein function and identity confirmed through experimental methods. The selected seed sequences are then stored in a database for future use. Those sequences which pass the HMMER alignment are deposited to a local sequence database. Single or multiple probes for each

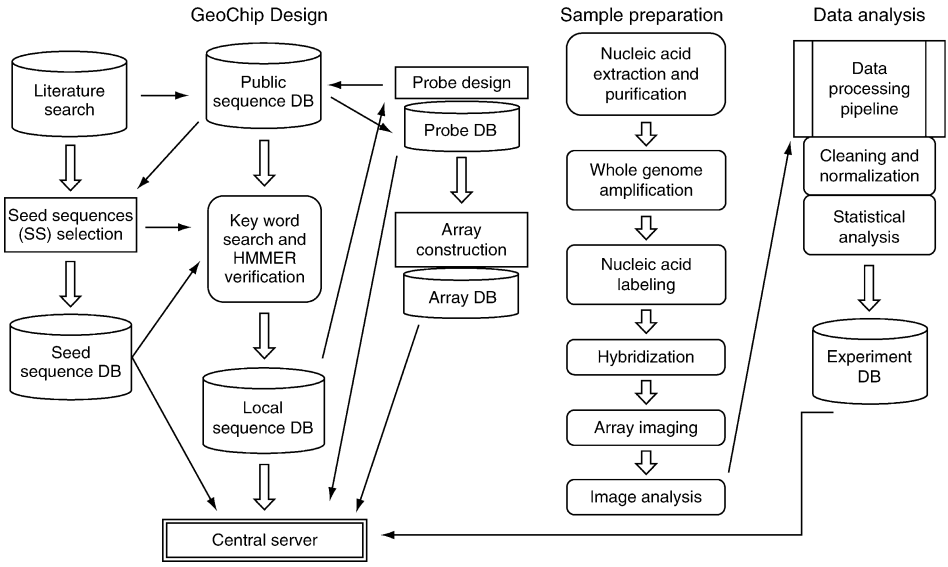

Figure 1

Design and use of GeoChip. See text for full explanation of all steps. DB, database.

selected sequence are designed with CommOligo (Li et al., 2005) by considering sequence homology, continuous stretches and free energy (He et al., 2005b). To ensure probe specificity, all designed probes are checked against the GenBank database, and the best probes are then synthesized and used to construct the array.

Community DNA is generally extracted and purified from environmental samples using well established methods (Zhou et al., 1996). If environmental samples do not yield sufficient DNA for GeoChip analysis, whole community genome amplification (Wu et al., 2006) is used to obtain enough DNA (usually 1–100 ng is used for amplification) for hybridization. The amplified DNA is then labeled with fluorescent dyes (e.g., Cy3, Cy5) and then hybridized to the array at 42°C and 50% formamide. After hybridization, the array is imaged and then digitally analyzed using microarray analysis software, which measures signal and background intensities. The resulting raw data is uploaded to the GeoChip data analysis pipeline (http://ieg.ou.edu/). The data is then evaluated for quality, cleaned by removing poor and low intensity spots, and normalized. The normalized data is stored in an experiment database and can then be used for further statistical analysis using the data analysis pipeline. A more detailed description of method is available in Chapter 51, Vol. 5, Part 3.

Several of these steps are critical in terms of sequence inclusion, probe design, and success of hybridization. Because they are used to validate inclusion of downloaded sequences, the selection of seed sequences is critical and chosen sequences must be carefully evaluated. Another critical step in array design is determining probe design criteria, as this will greatly affect specificity. For GeoChip 2.0, He et al. (2007) used gene-specific probe criteria of $\leq$90% sequence identity, $\leq$20-base continuous stretch and ≥ -35 kcal mol^{-1} free energy. Group-specific probes required the above requirements and $\geq$96% sequence identity, $\geq$35-base continuous stretch and ≤ -60 kcal mol^{-1} of free energy. Based on these criteria, all probes are expected to be highly specific to their corresponding targets and minimal false positive hybridizations occurred at 45–50°C and 50% formamide. Finally, DNA quality and purity are of utmost importance in obtaining quality hybridization results.

3 Key Issues for Microarray Application

Much progress has been made over the past decade regarding the use of microarray technology for environmental studies (Adey et al., 2002; Gao et al., 2007; He et al., 2007; Liebich et al., 2006; Rhee et al., 2004; Wu et al., 2001, 2004, 2006). Several key issues are discussed below.

Quality of nucleic acids. Obtaining high-quality DNA or/and RNA from environmental samples is the key to successful FGA analysis. While DNA extraction and purification has been successful using an established freeze-grind method (Hurt et al., 2001; Zhou et al., 1996), some samples are still difficult to purify to the necessary level. Extraction and purification of mRNA is more of a challenge. The instability and low abundance of environmental mRNA makes isolation difficult. Only a few studies have examined community mRNA from environmental samples using FGAs (Dennis et al., 2003; Gao et al., 2007; Hurt et al., 2001). Peplies et al. (2006) used community rRNA for a 16S rRNA array and Parro et al. (2007) used community RNA for a *Leptospirillum ferrooxidans* array.

Probe coverage. The number of sequences available in the public databases is increasing exponentially. In order to maintain sufficient coverage of known sequences and keep GeoChip up-to-date and relevant, continual updates are necessary. Our probe design system (*Fig. 1*) allows automatic updates using the predetermined keywords and seed sequences. However,

with the currently available probe design software, it is still difficult and time consuming to design and test such large numbers of sequences and probes. In addition, with the increasing number of sequences available, the number of probes on the GeoChip will increase as well. As such, the maximum capacity of arrays is still a limiting factor.

Specificity. Designing specific gene probes for community samples is difficult since environmental sequences could display high sequence divergence and a majority of environmental sequences are unknown. Oligo-based probes generally provide higher specificity than PCR products due to the ability to customize the oligo sequence and omit conserved areas of the sequence (Denef et al., 2003; Rhee et al., 2004). Specificity can be adjusted by changing the hybridization stringency. Hybridizations performed at 65°C required sequence similarities of >87% for hybridization to occur (Wu et al., 2001). Under low stringency conditions (45°C), detection of sequences divergent (70–75%) from the probe sequence was possible, but no cross-hybridization occurred (Wu et al., 2001). Rhee et al. (2004) showed that 50-mer oligonucleotide FGAs were able to distinguish sequences <88–94% identity at 50°C and 50% formamide. As mentioned previously, probe design criteria are important in determining specificity. He et al. (2007) used criteria for similarity, stretch and free energy to increase specificity (discussed above).

Sensitivity. The detection limit for FGAs based on current technology is 5% of the microbial community (Bodrossy et al., 2003). So, only the dominant community members can be detected. The relatively low sensitivity of slide-based arrays is primarily the result of the amount of probe (<20 pg/spot for glass slides vs. >1 μg/spot for membranes) (Cho and Tiedje, 2002). So, one strategy for increasing sensitivity is to increase probe concentration (Cho and Tiedje, 2002; Relógio et al., 2002; Zhou and Thompson, 2002), although this could also result in loss of signal intensity (Denef et al., 2003). Increasing the probe length increases sensitivity (Denef et al., 2003; He et al., 2005a), but also decreases specificity (Relógio et al., 2002). Several sample preparation and hybridization strategies can be applied which can increase sensitivity. A relatively new amplification technique, whole community genome amplification (WCGA), can be used to increase the concentration of all community DNA; including low abundance sequences (Wu et al., 2006). With WCGA, 1–250 ng of community DNA could be representatively amplified. The use of different types of dyes or labeling products can increase sensitivity. For example, the use of cyanine dye-doped nanoparticles for labeling has been shown to increase sensitivity tenfold (Zhou and Zhou, 2004) as has tyramide signal amplification labeling methods (Denef et al., 2003). An additional strategy to increase sensitivity include development of more sensitive signal detection systems (Cho and Tiedje, 2002; Zhou and Thompson, 2002).

Quantitative applications. One goal of microarray analysis is to correlate signal intensity with gene abundance. Several studies have shown correlation between these two variables. For example, a linear relationship ($r = 0.94$) between signal intensity and DNA concentration has been observed for both pure culture DNA and community DNA over a concentration range of 0.5–100 ng using PCR probes (Wu et al., 2001). With 50-mer oligo probes, a linear relationship ($r = 0.98–0.99$ for each probe) was observed over a concentration range of 8–1,000 ng (Tiquia et al., 2004). However, while several studies have suggested that microarray results can be quantitatively accurate, there remains the difficulty of accurately determining when differences in signal intensity arise from gene abundance or from sequence divergence.

Statistical analysis. GeoChip produces a massive amount of data. The sheer volume of this data makes it difficult to determine the best way to analyze and utilize the data. While several statistical methods have been used with GeoChip data (e.g., diversity indices, cluster analysis,

principle component analysis, DCA, CCA), additional tools and techniques are needed to rapidly analyze and visualize such huge datasets.

4 Application of FGAs for Environmental Study

GeoChip 2.0 is the most comprehensive FGA currently published for studying various biogeochemical, ecological, and environmental processes (He et al., 2007). This and earlier FGAs have been used to study the functional potential and diversity of microbial communities from a variety of samples and have shown the power of GeoChip to link microbial community functional structure to environmental processes and as a generic tool for profiling the differences of microbial community structure. Recent microbial community studies include: (1) Examining community changes during bioremediation of U(VI). GeoChip 2.0 was used to examine U-contaminated groundwater and showed a significant correlation between U(VI) concentrations and the amount of cytochrome genes detected ($r = 0.73$, $p < 0.05$) indicating the importance of this group of microorganisms in U(VI) reduction (He et al., 2007). (2) Monitoring microbial community shifts in a diesel bioremediation system. Rodríguez-Martínez et al. (2006) examined a fluidized bed reactor fed diesel fuel contaminated-groundwater in Vega Baja, Puerto Rico. Several genes involved in the degradation of diesel were detected including those involved in the degradation of phthalate, biphenyl, cyclohexanol, benzoate, and naphthalene. An increase in the amount of genes involved in anaerobic degradation of organic contaminants was also observed over time. These results were consistent with other evidence that suggested that the system shifted to a predominantly anaerobic process over time. (3) Investigating microbial N and C cycling in Antarctic sediments. Yergeau et al. (2007) examined Antarctic sediments using GeoChip 2.0 and found that cellulose degradation and denitrification genes positively correlated with soil temperature. (4) Examining the impact of different land use strategies on microbial soil communities. Zhang and colleagues (2007) used GeoChip 1.0 to examine microbial functional changes under different land use strategies and found that diversity and functional gene numbers increased as soil organic carbon increased. (5) Determining spatial scaling of microbial communities in forest soil. GeoChip 2.0 was used to assess the gene-area relationship of microbial communities in soils and found it to be relatively flat with less turnover than observed for plants and animals (Zhou et al., 2008). (6) Examining the structure of marine sediment communities. Examination of sediments from the Gulf of Mexico using GeoChip 1.0 revealed increasingly unique communities with depth (Wu et al., 2008). Genes involved in several of the key processes of this environment were detected. In addition, it appeared that several environmental parameters (depth, porosity, and concentrations of ammonium, phosphate, Mn(II), and silicic acid) may be important in shaping the structure of microbial communities in this environment. (7) Characterizing the structure of deep sea hydrothermal vent communities. Another study examined microbial communities from a mature deep-sea hydrothermal vent chimney and the inner and outer portions of a 5-day old chimney (Wang et al., 2009). The results indicated that microbial functional diversity is much lower in the inner chimney than the outer or mature chimneys. These findings were confirmed by real-time PCR and clone library results. (8) Other studies. GeoChip 2.0 has also been used to probe pure culture isolates for metal resistance genes (Van Nostrand et al., 2007), to detect active microbial populations in stable isotope probing experiments (Leigh et al., 2007), and to examine gene expression in microbial communities

(Gao et al., 2007). These studies indicated that the GeoChip is a powerful tool for analyzing microbial community functional structure.

We have recently analyzed four soil samples from two oilfields, Daqing and Shengli, located in northeastern and eastern China, respectively (Liang et al., unpublished data). Both sites had been contaminated with crude oil for more than a decade. The samples were collected from uncontaminated (0.5% oil), low (4.9%), medium (6.0%), and high (16.2%) contaminated areas. Interestingly, when all detected functional genes were used for hierarchical cluster analysis, the communities grouped based on geographical location with all samples from the Daqing oilfield grouping together while the community from the Shengli oilfield was separate (*Fig. 2*), suggesting that distance was the most important variable in deciding functional community structure. Canonical correspondence analysis (CCA) was used to further examine the relationship between functional genes and environmental variables (ter Braak, 1986). Using the two most significant geochemical parameters (pH and oil concentration) for these samples showed a very clear sample gradient based on oil concentration along the oil vector. The results of this study also showed the power of GeoChip to profile microbial community differences related to hydrocarbon contamination and degradation.

5 Research Needs

While FGAs allow for the examination of functional potential and the presence of specific genes, they do not provide information regarding the actual activity of microbial populations. However, a few studies have reported techniques which allow for the examination of microbial activity using FGAs. Stable isotope probing was used to detect active community members involved in PCB degradation (Leigh et al., 2007). Using mRNA for FGA hybridization, similar to the use of pure culture arrays, would allow for the determination of active community members (Bodrossy et al., 2006; Dennis et al., 2003; Gao et al., 2007). Since all probes are selected from functional gene coding sequences, it is possible to detect microbial activity using RNA targets. However, one difficulty with the use of mRNA in community studies, unlike for pure culture arrays, is the low abundance of mRNA in the environment. A whole community RNA amplification procedure was recently published that overcomes this obstacle and allows for the examination of active microbial processes, even in samples with low abundance (Gao et al., 2007). However, this protocol is complex and time consuming, so improved techniques for extraction of community RNA and subsequent amplification are needed.

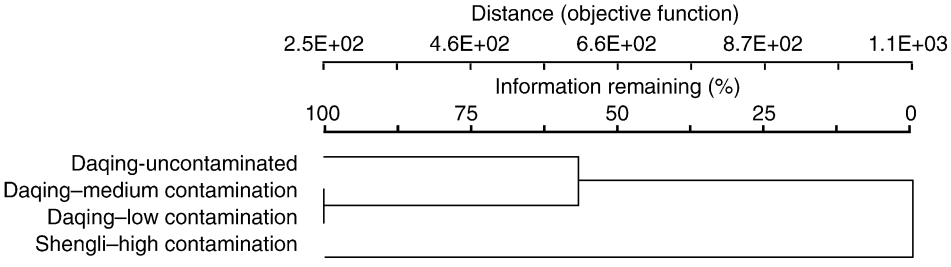

Figure 2

Hierarchical clustering of functional genes.

Acknowledgments

The effort for preparing this review was supported by the Virtual Institute for Microbial Stress and Survial (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-ACO2-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy, the 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, and the National Natural Scientific Foundation of China (No. 40730738).

References

Adey NB, Lei M, Howard MT, Jenson JB, Mayo DA, Butel DL, Coffin SC, Moyer TC, Slade DE, Spute MK, Hancock AM, Eisenhohoffer GT, Dalley BK, McNeely MR (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, Konrad-Köszler M, Weilharter A, Reichenauer TG, Schöfer D, Sessitsch A (2006) mRNA-based parallel detection of active methanotroph populations by use of a diagnostic microarray. Appl Environ Microbiol 72: 1672–1676.

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.

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 Microbiol 5: 933–943.

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.

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. Microbial 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, Gentry TJ, Schadt CW, Wu L, Liebich J, Chong SC, Huang Z, Wu W, Gu B, Jardine P, Criddle C, Zhou J (2007) GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. ISME J 1: 67–77.

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.

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.

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.

Liebich 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.

Parro V, Moreno-Paz M, González-Toril E (2007) Analysis of environmental transcriptomes by DNA microarrays. Environ Microbiol 9: 453–464.

Peplies J, Lachmund C, Glöckner FO, Manz W (2006) A DNA microarray platform based on direct detection or rRNA for characterization of freshwater sediment-related prokaryotic communities. Appl Environ Microbiol 72: 4829–4838.

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.

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

Sessitsch A, Hackl E, Wenzl P, Kilian A, Kostic T, Stralis-Pavese N, Tankouo Sandjong B, Bodrossy L (2006) Diagnostic microbial microarrays in soil ecology. New Phytol 171: 719–736.

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, Xu Y, 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.

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

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. PNAS 106: 4840–4845.

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.

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, 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.

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 Methods 70: 165–178.

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 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 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, 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. PNAS 105: 7768–7773.

Zhou J, Thompson DK (2002) Challenges in applying microarrays to environmental studies. Curr Opin Biotechnol 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.