Use of functional gene arrays for elucidating *in situ* biodegradation

Joy D. Van Nostrand[1,2], Zhili He[1,2] and Jizhong Zhou[1,2]*

[1] Institute for Environmental Genomics, University of Oklahoma, Norman, OK, USA
[2] Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK, USA

Microarrays have revolutionized the study of microbiology by providing a high-throughput method for examining thousands of genes with a single test and overcome the limitations of many culture-independent approaches. Functional gene arrays (FGA) probe a wide range of genes involved in a variety of functions of interest to microbial ecology (e.g., carbon degradation, N fixation, metal resistance) from many different microorganisms, cultured and uncultured. The most comprehensive FGA to date is the GeoChip array, which targets tens of thousands of genes involved in the geochemical cycling of carbon, nitrogen, phosphorus, and sulfur, metal resistance and reduction, energy processing, antibiotic resistance and contaminant degradation as well as phylogenetic information (*gyrB*). Since the development of GeoChips, many studies have been performed using this FGA and have shown it to be a powerful tool for rapid, sensitive, and specific examination of microbial communities in a high-throughput manner. As such, the GeoChip is well-suited for linking geochemical processes with microbial community function and structure. This technology has been used successfully to examine microbial communities before, during, and after *in situ* bioremediation at a variety of contaminated sites. These studies have expanded our understanding of biodegradation and bioremediation processes and the associated microorganisms and environmental conditions responsible. This review provides an overview of FGA development with a focus on the GeoChip and highlights specific GeoChip studies involving *in situ* bioremediation.

Keywords: functional gene microarray, GeoChip, microbial communities, bioremediation, biodegradation

INTRODUCTION

As the most phylogenetically and functionally diverse group of organisms on the planet (estimated 2000–50,000 microbial species per gram of soil; Torsvik et al., 1990; Hong et al., 2006; Schloss and Handelsman, 2006; Roesch et al., 2007), microorganisms are critical to ecosystem functioning and are involved in the biogeochemical cycling of carbon, nitrogen, sulfur, phosphorus, and metals, as well as degradation or stabilization of contaminants in the environment. However, because a vast majority (>99%) of microorganisms remain uncultured (Amann et al., 1995; Fuhrman and Campbell, 1998; Whitman et al., 1998), culture-independent approaches must be used to gain a comprehensive picture of microbial communities. However, many of the culture-independent methods, such as 16S rRNA gene-based cloning or quantitative PCR, require a PCR amplification step, which introduces well-known biases (Suzuki and Giovannoni, 1996; Warnecke et al., 1997; Lueders and Friedrich, 2003). In addition, since many functional genes have too much variance or too few sequences available, conserved PCR primers cannot be designed for many functional genes. Even if primers could be designed for many functional genes, performing PCR with many different primer sets would be cost- and time-prohibitive.

Microarrays allow the examination of thousands of genes at one time without the need for PCR amplification of each gene. Since microarrays were first shown to be valuable for the study of

microbial communities (Guschin et al., 1997), several types have been designed to examine microbial communities. These include (i) phylogenetic oligonucleotide arrays (POA), designed to examine phylogenetic relatedness or community composition using 16S rRNA or other conserved phylogenetic genes (Small et al., 2001; Loy et al., 2002; Wilson et al., 2002; Brodie et al., 2006); (ii) community genome arrays (CGA), designed to examine the relatedness of microbial species or strains or to identify community members using whole-genomic DNA probes (Wu et al., 2004; Zhang et al., 2004); (iii) metagenomic arrays (MGA), designed as a high-throughput screening method using environmental clone library inserts as probes (Sebat et al., 2003; Mockler and Ecker, 2005; Gresham et al., 2008); (iv) whole-genome ORF arrays (WGA), designed to examine gene expression of individual microorganisms using probes for all ORFs in one or more genomes (Wilson et al., 1999), but can also be used for comparative genomics (Murray et al., 2001); and (v) functional gene arrays (FGAs), designed to examine multiple functional genes at one time using probes for key genes involved in microbial functional processes of interest (Wu et al., 2001; He et al., 2007, 2010a). This review will focus on FGAs.

MICROARRAYS VS. HIGH-THROUGHPUT SEQUENCING
While this review focuses on functional gene microarrays, high-throughput sequencing has become an increasingly popular choice

for examining and monitoring microbial communities and is even being used for metatranscriptome analysis (van Vliet, 2010). However, while many of the technical challenges of microarrays and high-throughput sequencing have been overcome, each still has some distinct advantages and disadvantages, which make them ideal as complementary approaches: (i) Random sampling errors. In most sequencing studies, only a small proportion of the microbial community is actually sampled (McKenna et al., 2008) and while theoretically with true random sampling the probability of sampling the same fraction of the community multiple times is low (Zhou et al., 2008), one would expect that dominant populations would have a greater chance of being sampled multiple times. These sampling errors can result in low reproducibility between technical replicates (17.2 ± 2.3% for two replicates; 8.2 ± 2.3% for three; Zhou et al., 2011). Microarrays, in contrast, interrogate all samples against the same set of sequences (probes), so that the same population is sampled each time. (ii) Relative abundance. Abundance of individual species will vary greatly within microbial communities. With sequencing-based approaches there will be a bias toward the most abundant sequences in the environment so that many of the obtained sequences will represent the most abundant species/sequences while possibly missing lesser abundant species/sequences. Microarrays are not affected in the same way since lesser abundance sequences will still hybridize to their corresponding probe and as long as it is above the detection limit, it will be detected. (iii) New sequence detection. One of the greatest advantages of sequencing is that new sequences are easily detected since any sequences in the sample can be sequenced (open system). Microarrays, in contrast, can detect only the limited number of sequences covered by the probe set on the array (closed system), as such, it is not able to detect new sequences. Some new microarray techniques have been developed to allow the discovery of new sequences. Capture microarrays have been developed, which use lower stringency conditions to hybridize or "capture" sequence variants (Albert et al., 2007; Okou et al., 2007). These captured sequences are then washed off and sequenced. An array with probes specific for viral families has been developed that uses the hybridization pattern to classify novel viruses (Wang et al., 2002; Ksiazek et al., 2003). As such, microarray and sequencing approaches could be used to maximize the benefits and minimize the deficiencies of each.

DEVELOPMENT OF FUNCTIONAL GENE ARRAYS

The first FGA developed used PCR-amplicon probes and targeted four N-cycling genes (*nirS*, *nirK*, *amoA*, and *pmoA*; Wu et al., 2001). However, since PCR-amplicons were used, only a limited number of genes could be included because conserved primers can only be designed for a few functional genes. In addition, it would be cost- and time-prohibitive to amplify genes from hundreds of microorganisms or clones in order to achieve a truly diverse probe set. Most microarrays now use oligonucleotide probes, which are more specific (Zhou, 2003), can be easily customized (Denef et al., 2003; Zhou, 2003; Gentry et al., 2006), and are relatively inexpensive.

Since then several different FGAs have been developed, although most of these cover only a limited number of genes and focus on specific functional groups or locations. For example, FGAs have been designed to examine methanotrophs (Bodrossy et al., 2003, 2006; Stralis-Pavese et al., 2004), N-cycling genes (Taroncher-Oldenburg et al., 2003; Jenkins et al., 2004; Steward et al., 2004; Zhang et al., 2007a), pathogens and virulence factors (Call et al., 2003; Kostić et al., 2005; Cleven et al., 2006; Miller et al., 2008; Palka-Santini et al., 2009), rhizobial isolates (Bontemps et al., 2005), and acid mine drainage (AMD) and bioleaching systems (Yin et al., 2007).

The most comprehensive FGAs reported to date are the GeoChip arrays. The GeoChip 1.0 had 2006 oligonucleotide probes (50-mers) for genes involved in nitrification, denitrification, nitrogen fixation, methane oxidation, sulfate reduction (Tiquia et al., 2004), organic contaminant degradation, and metal resistance (Rhee et al., 2004). This array was used in several studies examining microbial communities at uranium (U)-contaminated sites (Wu et al., 2006a; Waldron et al., 2009), in the Gulf of Mexico (Wu et al., 2008), and under different land use strategies (Zhang et al., 2007b) and showed FGAs to be useful for microbial community studies. GeoChip 2.0 was developed to provide a truly comprehensive probe set for multiple functional gene categories and to provide increased specificity for highly homologous gene variants (He et al., 2007). GeoChip 2.0 contains 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. This array has been used in numerous studies to examine microbial communities at metals contaminated sites (Gao et al., 2007; Van Nostrand et al., 2009, 2011), oil or diesel-contaminated sites (Rodríguez-Martínez et al., 2006; Liang et al., 2009a,b), coral mucus (Kimes et al., 2009), lake or river samples (Taş et al., 2009; Parnell et al., 2010), deep sea samples (Mason et al., 2009; Wang et al., 2009), Antarctic soils (Yergeau et al., 2007) and to examine the taxa–area relationship (Zhou et al., 2008).

GeoChip 3.0 covers 56,990 sequences from 292 gene families, greatly increasing the number of genes and categories covered compared to GeoChip 2.0 and added new control features (He et al., 2010a). New gene categories include antibiotic resistance, energy processing, and phylogenetic markers (i.e., *gyrB*). A set of 16S rRNA gene probes were added as positive controls, human, plant, or hyperthermophile gene probes were added as negative controls, and a common oligo reference standard (CORS) was added for data normalization and comparison. The CORS is composed of an artificial sequence probe that is co-spotted with each gene probe and the complementary CORS target, labeled with a contrasting fluorescent dye to the sample, which is then spiked into each sample prior to hybridization (Liang et al., 2010). The signal intensity of the CORS probe can then used to normalize the signal intensity of the sample and allows comparison of samples hybridized at different times. In addition, a computational pipeline has been developed for GeoChip probe design and data analysis. The GeoChip 3.0 has been used to examine microbial communities associated with elevated CO_2 (He et al., 2010b), to examine communities within coal formation production waters (Wawrik et al., 2012b) or rhizosphere communities in As-contaminated sites (Xiong et al., 2010).

GeoChip 4.0, the newest version, is synthesized by Nimblegen (Madison, WI, USA) in their 12-plex format and contains 83,992 probes targeting 152,414 genes in 410 gene categories (Lu et al., 2012a). In addition to added genes in most categories, new categories added include stress response, antibiotic resistance, and bacteriophage genes. It has been used to examine microbial communities during the 2010 Gulf oil spill (Lu et al., 2012a),

GeoChip DESIGN

PROBE DESIGN AND SELECTION

GeoChip covers a wide range of functional genes and currently includes sequences from bacteria, Achaea, fungi, and viruses. The first step in designing new probes for the array is deciding which processes should be included. Then genes for enzymes or proteins that are key to the process of interest are selected. These could be catalytic subunits or proteins with recognition sites or that provide functional specificity. Next, keywords are selected to search public sequence databases (e.g., GenBank). The keywords should be as broad as possible since proteins from different microorganisms may be annotated differently or have more general or specific annotations. Once the sequences are downloaded, they are confirmed by HMMER alignment[1] with preselected seed sequences. The seed sequences are those sequences for which the protein identity and function have been experimentally confirmed. This is a critical step in the design process and these sequences should be selected with care. The HMMER confirmed sequences are then used to design gene- or group-specific 50-mer oligonucleotide probes using new versions of the CommOligo software (Li et al., 2005) and experimentally determined criteria based on sequence homology ($\leq 90\%$ identity for gene-specific probes, and $\geq 96\%$ for group-specific probes), continuous stretch length (≤ 20 bases for gene-specific probes, and ≥ 35 for group-specific probes), and free energy (≥ 35 kJ mol^{-1} for gene-specific probes, and ≤ 60 kJ mol^{-1} for group-specific probes; He et al., 2005b; Liebich et al., 2006). The probes are then BLASTed against the GenBank database to confirm specificity. Keywords, downloaded sequences, seed sequences, HMMER confirmed sequences, and designed probes are stored in corresponding databases for use in future array updates.

The newly designed probe sets can then be commercially synthesized. Several options are available for producing arrays. Synthesized oligonucleotide probes can be spotted onto nylon membranes (Steward et al., 2004) or glass slides (Taroncher-Oldenburg et al., 2003; Rhee et al., 2004; Tiquia et al., 2004). Glass slides are more frequently used since they have less background fluorescence (Schena et al., 1995, 1996) and allow higher probe density (Ehrenreich, 2006). Probes can also be added to slides using bubble Jet printing (Okamoto et al., 2000), laser-induced forward transfer (Serra et al., 2004), or photolithography (Chen et al., 2009). In addition, a few companies, such as Agilent or Affymetrix, synthesize custom microarrays using a customer's probe set.

TARGET PREPARATION

GeoChip can be hybridized with either DNA or RNA. Most DNA samples used for GeoChip analysis are extracted using a

---

[1] http://hmmer.wustl.edu/

well-established freeze-grind method with detergent lysis (Zhou et al., 1996; Hurt et al., 2001) since it provides high molecular weight DNA, important for later amplification steps. The use of RNA presents some challenges as mRNA is unstable and has a low abundance in environmental samples. Several papers have described methods for extracting environmental RNA, including a protocol for the dual extraction of both DNA and RNA (Hurt et al., 2001; Burgmann et al., 2003) or RNA alone (McGrath et al., 2008; Poretsky et al., 2009a). Methods for mRNA enrichment include size separation by gel electrophoresis (McGrath et al., 2008) or use of commercial kits [MICROBExpress (Ambion) and/or mRNA-ONLY (Epicentre Biotechnologies); Poretsky et al., 2009b; Mettel et al., 2010]. Size separation obtained 115–155 ng mRNA from 4.6–5.3 μg total RNA (McGrath et al., 2008). Using commercial kits, Mettel et al. (2010) were able to obtain 140–530 ng of mRNA from 0.4–2.0 μg total RNA per 0.5 g soil.

Nucleic acid quality is of great importance for microarray analysis. DNA and RNA should have an A_{260} to A_{280} ratio ~ 1.8 and > 1.9, respectively and an A_{260} to $A_{320} \geq 1.7$. The A_{260} to A_{320} ratio is most important in determining microarray success (Ning et al., 2009). Some environmental samples, especially those with high humics, can be difficult to purify up to the necessary level. A gel purification strategy followed by a phenol–chloroform–butanol extraction (Xie et al., 2007; Liang et al., 2011) has been successful with a wide range of soil and sediment samples.

Large amounts of DNA (e.g., 1 μg) or RNA (e.g., 5 μg) are needed for GeoChip hybridization. However, it can be difficult to get sufficient quantities of nucleic acid from some types of samples (e.g., water) or the sample is too difficult or impossible to replace to use such large quantities of nucleic acid. In this case, amplification of DNA or RNA can be done using either whole community genome amplification (WCGA; Wu et al., 2006a) or whole community RNA amplification (WCRA; Gao et al., 2007). WCGA uses the Templiphi 500 amplification kit (phi 29 DNA polymerase, GE Healthcare, Piscataway, NJ, USA) with a modified amplification buffer and using 1–100 ng DNA provides a sensitive (10 fg detection limit) and representative amplification ($<0.5\%$ of amplified genes showed >2-fold difference from unamplified; Wu et al., 2006a). WCRA provides a representative amplification with 50–100 ng of starting material.

There are commercial kits available for microbial RNA amplification such as the MessageAmp^{TM} II-Bacteria RNA Amplification Kit (Life Technologies, Grand Island, NY, USA). There are also other commercially available methods for WCGA. Wang et al. (2011) compared two of these (*Bacillus stearothermophilus* DNA polymerase (*Bst*) and REPLI-g; Qiagen, Valencia, CA, USA) with the modified Templiphi kit (Wu et al., 2006a). The amplification bias for all methods was relatively low (<3-fold). Less bias was observed with REPLI-g and Templiphi for pure culture DNA and with REPLI-g for community DNA while *Bst* showed the least inhibition by lesser quality DNA.

The amplified (or unamplified) nucleic acids are directly labeled with a fluorescent dye (Cy3 or Cy5) using random priming with the Klenow fragment of DNA polymerase for DNA (Wu et al., 2006a) or Superscript^{TM} II/III RNase H-reverse transcriptase for RNA (He et al., 2005b). The labeled DNA/RNA is then purified and dried for hybridization.

HYBRIDIZATION AND IMAGE ANALYSIS

The labeled nucleic acids are then hybridized to the microarray at 42–50°C with 40–50% formamide (He et al., 2007, 2010a,b; Lu et al., 2012a). Hybridization specificity can be adjusted by varying the temperature or the formamide concentration (the effective hybridization temperature increases by 0.6°C for every 1% of formamide). Hybridized slides are then scanned and analyzed by quantifying the pixel density (intensity) of each spot using image analysis software. Commercial manufacturers often have their own analysis software or other microarray analysis software can be used, such as GenePix Pro (Molecular Devices, Sunnyvale, CA, USA), GeneSpotter (MicroDiscovery, San Diego, CA, USA), or ImaGene (BioDiscovery, El Segundo, CA, USA). For GeoChip data, there is a data analysis pipeline[2] for rapid preprocessing and data analysis. Poor and low quality spots and outliers, based on Grubbs' test of outliers (Grubbs, 1969), are removed and then the signal intensities of all spots are normalized. Positive spots can be determined using signal-to-noise ratio [SNR = (signal mean − background mean)/background standard deviation], signal-to-both-standard-deviations ratio [SSDR = (signal mean − background mean)/(signal standard deviation − background standard deviation)] (He and Zhou, 2008), or signal-to-background ratio (SBR = signal mean/background mean) (Loy et al., 2002).

DATA ANALYSIS

Due to the large volume of data obtained from GeoChip, data analysis can be very challenging. The data has a multivariate structure and the number of variables is much larger than the number of observations ($p \gg n$). To assist users with data analysis steps, a pipeline is available which performs many of the common analyses[3]. Some common descriptive statistics used include relative abundance of gene categories or subcategories, richness and diversity (α and β) indices, and percentages of gene overlap between samples. To compare the overall community structure, unconstrained ordination [principal component analysis (PCA) and correspondence analysis (CA)] to reduce the dimensionality of variables in order to maximize the visible variability of the data or hierarchical cluster analysis (HCA), which groups communities based on the similarity of their gene profiles, can be used. To compare communities, response ratios, which compare the signal intensity of genes between conditions (Luo et al., 2006; Liang et al., 2009a), t-tests, ANOVA, and dissimilarity tests can be used. Several methods can be used to examine the relationship between communities and environmental parameters. These include constrained ordination, such as canonical correspondence analysis (CCA; ter Braak, 1986), distance-based redundancy analysis (db-RDA; Legendre and Anderson, 1999), variation partitioning analysis (VPA; Økland and Eilertsen, 1994; Ramette and Tiedje, 2007), and Mantel test. A relatively new analysis method is the random matrix theory-based (Mehta, 1990) neural network analysis (NNA) used to examine gene relationships within microbial ecological networks (Zhou et al., 2010).

---

[2]http://ieg.ou.edu/

[3]ieg.ou.edu/microarray/

IMPORTANT CHALLENGES FOR MICROARRAY ANALYSIS

NUCLEIC ACID QUALITY

Having high-quality nucleic acids (non-degraded, large fragments to improve amplification yields, absence of inhibitors or contaminants which may impede subsequent amplification and labeling steps) is the most important criterion for successful microarray experiments. Nucleic acids can be purified using commercial kits although the presence of humic acids and other contaminants can still be a problem. If large amounts of DNA are present, an agarose gel purification followed by phenol–chloroform–butanol extraction (Xie et al., 2007; Liang et al., 2009b) can be used, but large amounts of DNA are lost with this method so it is not practical for low abundance samples. So, better purification methods with high recovery yields are needed.

SEQUENCE COVERAGE

One of the main objectives in developing FGAs was to provide a truly comprehensive probe set (He et al., 2007). Each new GeoChip version has expanded the coverage of gene variants and expands the number of genes covered (He et al., 2007, 2010a; Lu et al., 2012a). This continued expansion is challenging as the number of gene sequences available is constantly increasing as new sequences are being submitted to public databases. While the GeoChip design pipeline[2] has an automated update feature which uses the previously selected key words and seed sequences to search the NCBI database, downloading new sequences and designing probes is still time consuming due to the sheer volume of sequences available. As such, better and faster computation systems are needed. In addition, available microarray probe density limits are rapidly being approached as the number of GeoChip probes increases. So, new methods of array construction to increase probe density are needed.

SPECIFICITY AND SENSITIVITY

Two key issues for microarray hybridization of microbial communities are specificity and sensitivity since environmental communities can have such vast diversities. Both of these conditions can be improved at various stages of microarray design, construction, target preparation, or hybridization. During probe design, determining the best criteria for probe design, such as using experimentally determined design criteria (He et al., 2005b; Liebich et al., 2006) can improve specificity, thus decreasing the number of false positives (He et al., 2007). Probe length also affects specificity and sensitivity; longer probes are more sensitive, but less specific (Denef et al., 2003; He et al., 2005a).

The method of array synthesis can also affect sensitivity and specificity. Increasing the probe concentration per spot can increase sensitivity (Cho and Tiedje, 2002; Relógio et al., 2002; Zhou and Thompson, 2002). However, this may also decrease specificity by decreasing the overall probe signal intensity (Denef et al., 2003). The choice of array surface can also be important as use of unmodified array slides can decrease background fluorescence thus requiring a lower signal fluorescence for detection (Kumar et al., 2000; Gudnason et al., 2008).

Target preparation strategies can also affect these parameters. Amplification of community DNA can increase sensitivity. WCGA was able to representatively amplify 1–250 ng of community DNA

(Wu et al., 2006a), increasing the detection limit from 25 ng to 10 pg (2 bacterial cells); however, using such small quantities of starting material greatly increases the amplification bias compared to the bias observed with 1 ng of DNA. Labeling methods can also affect sensitivity. For example, cyanine dye-doped nanoparticles or tyramide signal amplification labeling are able to increase sensitivity 10-fold (Denef et al., 2003; Zhou and Zhou, 2004).

Hybridization conditions can also be used to increase specificity and sensitivity. Temperature and formamide concentration can be modified to adjust stringency thus altering specificity (Wu et al., 2001). A lower hybridization solution volume (Shalon et al., 1996) and mixing during hybridization (Adey et al., 2002) have both been shown to increase sensitivity. Decreasing ozone levels, which can degrade cy-dye signal (Branham et al., 2007), can also improve sensitivity.

MONITORING GENE ACTIVITY

Most GeoChip analysis has involved the use of DNA, so that only gene abundance can be determined. These changes can be used to infer microbial activity, but cannot provide direct proof of activity. mRNA can be used for FGA analysis to monitor activity (Dennis et al., 2003; Bodrossy et al., 2006; Gao et al., 2007; Wawrik et al., 2012a), although as mentioned above, working with environmental RNA can be challenging. Stable isotope probing (SIP) has also been used with GeoChip to monitor microbial activity (Leigh et al., 2007).

Gao et al. (2007) used amplified community mRNA from a denitrifying fluidized bed reactor to examine microbial activity. Genes for nitrate and nitrite reduction, organic contaminant degradation, sulfite reduction, and polyphosphate kinase were detected, consistent with reactor operation (Gao et al., 2007). Another study used amplified community mRNA to examine nitrate utilization in marine bacterial communities (Wawrik et al., 2012a). Hybridization results indicated activity by *ureC*, *nirS*, *nirK*, *narG*, *nosZ*, *napA*, *nrfA*, *amoA*, and *nifH* genes, indicating that urea cycling, denitrification, dissimilatory nitrate, nitrite reduction, and N fixation were occurring (Wawrik et al., 2012a).

Another method of monitoring microbial activity with GeoChip is to combine it with SIP (Leigh et al., 2007). Microcosms were set up from soil samples collected from the root zone of a tree growing in a PCB-contaminated site and fed ^{13}C-labeled or unlabeled biphenyl. Genes involved in biphenyl degradation were detected as were other organic contaminant degradation genes including those for degradation of benzoate, catechol, naphthalene, and phenol.

APPLICATION OF GeoChip TO BIOREMEDIATION STUDIES
METALS CONTAMINATED SITES
Several GeoChip-related studies have examined microbial communities from U-contaminated groundwater at the U.S. Department of Energy (DOE) Oak Ridge Integrated Field Research Challenge (OR-IFRC) site. Groundwater samples covering a range of contamination levels and an uncontaminated background sample were compared using GeoChip 1.0 (Wu et al., 2006a). Samples from the uncontaminated site and those with lower levels of contaminants had higher functional gene diversity and gene numbers. In addition, as expected based on the contaminants present at this

site, genes for denitrification, organic contaminant degradation, metal resistance, and sulfite reduction (*dsr*) were detected. A similar sample set using the same array was examined in greater detail in a later study (Waldron et al., 2009). In this study, sulfate, pH, U, and Tc were found to be the most important drivers in determining the microbial community structure, with pH and the combination of U and Tc explaining ∼21% of the variance observed or 29–40% when all four variables were included.

Another study at this site examined a pilot-scale field bioremediation system which used ethanol as an electron donor to stimulate microbial communities and immobilize U(VI) by reduction to U(IV) (Luo et al., 2006; Wu et al., 2006b,c). GeoChip 2.0 was used to examine the microbial communities during different phases of operation. A period of active U(VI) reduction occurred after initial start-up (days 137–304). During this period U(VI) reduction was relatively rapid and genes associated with denitrification, sulfate reduction, and Fe(III) reduction increased in abundance, suggesting that these populations were involved in U(VI) reduction (Van Nostrand et al., 2011). This active reduction was followed by a maintenance period during which the low level of U(VI) was maintained, and the denitrifying, sulfate- and Fe(III)-reducing communities remained in higher abundance. Next, the stability of the bioreduced U(IV) was examined by allowing the system to enter periods of starvation (ethanol injections were halted) and reoxidation (dissolved O_2 entered the system). The functional communities showed distinct clustering patterns based on whether the system received ethanol or not, indicating a shift in community structure with the return of ethanol injections (Van Nostrand et al., 2009). While total gene numbers increased once ethanol injection was restarted, the relative abundance of each gene group changed little during and after starvation, indicating a functionally diverse community which could be stimulated after adverse conditions. Chemical oxygen demand (COD, i.e., ethanol) was the most important driver in determining community structure, but temperature, sulfate, and U(VI) were also important.

In this same remediation system, the sediment microbial community was examined with GeoChip 2.0 after 2 years of operation (Xu et al., 2010). Sediment samples were collected from 11 wells, 5 from the outer loop and 6 from the inner loop. Results revealed significant differences between the microbial communities in the inner and outer loops. The inner loop communities had higher gene numbers and greater diversity than those in the outer loop and inner and outer loop samples were grouped separately based on hierarchical clustering and principle component analysis, indicating that the ethanol injections stimulated the microbial communities in the inner loop. In addition, genes important for U(VI) reduction such as cytochrome *c*, *dsr*, and denitrification as well as genes involved in metal resistance and organic contaminant degradation were enriched in the inner loop where electron donor was added. This study demonstrated the importance of U(VI)-reducing populations for the maintenance of reducing U(IV) in this bioremediation system.

Another GeoChip 2.0 study examined groundwater microbial communities at a field site examining the use of acetate to stimulate U(VI)-reducing microorganisms in the subsurface at the Old Rifle site, a former U ore processing facility in Rifle CO (Liang

et al., 2012). The study compared communities taken during a shift from sulfate to Fe(III)-reducing conditions. The overall community structure changed with the switch from Fe(III)- to sulfate-reducing conditions and were reflective of the redox conditions at the site. Sulfate-reducing and methane-generating microorganisms increased in abundance under sulfate-reducing conditions. Acetate, U(VI) and redox potential were important environmental variables in determining the microbial community structure.

Xie et al. (2011) examined five AMD sites in China using GeoChip 2.0 to determine the functional diversity and metabolic potential of microbial communities in these sites and to determine how the communities responded to environmental conditions. The sites showed a great deal of variability in regards to the microbial communities with ~150–1000 functional genes detected in each sample. Most of the genes represented on the GeoChip that were involved in C, N, S cycling and metal resistance were detected in all of the AMD sites. Results indicated that the immediate environmental conditions were important in forming the variations in the functional structure of microbial communities as opposed to spatial distance. There was a positive correlation between Zn resistance gene abundance and Zn concentration but not for other metals. However, the concentrations of B, Co, Cu, La, Mg, and S were significantly correlated with the community structure in these communities. Overall, results suggested that AMD microbial communities may not be as simple as previously thought.

GeoChip 2.0 has also been used to probe pure culture isolates for the presence of specific genes. Four Ni-resistant Gram-positive actinomycetes were hybridized to GeoChip to get a better idea of what metal resistance genes were present (Van Nostrand et al., 2007). Genes associated with resistance to Al, As, Cd, Cr, Cu, Hg, Ni, Te, and Zn were detected.

MICROBIAL COMMUNITIES ASSOCIATED WITH PHYTOREMEDIATION

Microbial communities from the rhizosphere of the arsenic-hyperaccumulating plant *Pteris vittata* and non-rhizosphere samples were examined using GeoChip 3.0 (Xiong et al., 2010). The functional gene diversity was significantly correlated with As concentration. Interestingly, As contaminated rhizosphere samples had higher functional gene diversity than non-rhizosphere samples even though the non-rhizosphere samples had a lower level of As. In addition, greater numbers of As resistance genes, with higher signal intensities, were detected in rhizosphere samples compared to non-rhizosphere samples and very few genes were detected in both environments, suggesting that the rhizosphere and non-rhizosphere microbial communities were distinct. Results suggested that the *P. vittata* rhizosphere may protect the microbial communities from As contamination.

Another study used GeoChip 2.0 to examine microbial communities in Zn- and Cd-contaminated soil microcosms with or without *Thlaspi caerulescens*, a Cd and Zn hyperaccumulator plant (Epelde et al., 2010). Higher numbers of functional genes were detected in the contaminated samples than in uncontaminated samples and in planted samples compared to unplanted. Thirty-five to forty-seven percent of the variation in community structure observed was explained by metal concentrations. All of the Cd and/or Zn resistance genes (12) were detected in the contaminated, planted samples while only 7 were detected in the contaminated/unplanted samples. Substrate-induced respiration, K concentration, and nitrate concentration were the most important environmental variables in determining the functional community structure.

OIL-CONTAMINATED SITES

The microbial community associated with a bioremediation system comprised of a fluidized bed reactor to clean diesel-contaminated groundwater in Vega Baja, Puerto Rico was examined with the GeoChip 1.0 (Rodríguez-Martínez et al., 2006). Genes involved in the degradation of diesel fuel and other organic contaminants (acetylene, aniline, benzoate, biphenyl, cyclohexanol, methyl tert-butyl ether, naphthalene, phthalate, protocatechuate, and toluene) were detected. Increased signal intensities for genes involved in anaerobic benzoate degradation indicated a shift toward anaerobiosis over time, a conclusion supported by other experimental evidence.

Liang et al. (2009b) examined the effect of different bioremediation treatments on microbial communities using laboratory scale bioremediation systems with sediment from contaminated oil fields and inoculated with oil degrading enrichment cultures. The systems were incubated 242 days, treated with ozone, and incubated an additional 125 days. Many oil degradation genes (benzene, benzoate, catechol, polyaromatic hydrocarbon aromatics, protocatechuate, phthalate) were detected with GeoChip 2.0. Ozonation treatment resulted in an almost 50% reduction in the number of functional genes detected. Gene numbers increased again after a recovery period and the community retained the ability to degrade oil.

Another study used GeoChip 2.0 to characterize microbial communities along an oil contaminant gradient and found a decreased number of functional genes as the contaminant levels increased although genes involved in the degradation of biphenyl, catechol, and protocatechuate increased in the more contaminated samples (Liang et al., 2009a). The most important environmental factors in determining the microbial community structure were oil concentration and soil available nitrogen.

Liang et al. (2011) collected contaminated and uncontaminated soils from five oil fields across China in order to determine whether oil contamination or geographic location played a larger role in determining the microbial community structure. Results from GeoChip 2.0 indicated that communities from uncontaminated sites had higher functional gene diversity than those from contaminated sites in the same geographical area. Overall, the microbial communities clustered based on geographic location; however, when only organic contaminant degradation genes were examined, the contaminated samples clustered together. Geographic location was able to explain ~33% of the microbial community variation observed, oil explained ~10% of the variation, and soil geochemistry explained another 12%, while the remainder (~41%) was unexplained.

GeoChip 4.0 was used to compare microbial communities in oil-contaminated water to those from uncontaminated water in order to understand the effects of the 2010 Gulf of Mexico oil spill (Hazen et al., 2010; Lu et al., 2012a). Results indicated that after only 40 days the presence of the hydrocarbon plume

(1100 m depth) caused a significant shift in the microbial community functional structure and composition and that indigenous microorganisms, similar to known petroleum degraders, were stimulated by the hydrocarbon plume. Many genes associated with hydrocarbon degradation were significantly enriched in plume samples (Hazen et al., 2010; Lu et al., 2012a). Genes that were enriched in plume samples included those for naphthalene 1,2-dioxygenase, β-oxidation of benzylsuccinate, cyclohexanone 1,2-monooxygenase, and alkene monooxygenase (Lu et al., 2012a). These findings suggest that the microbial communities in the Gulf of Mexico were capable of intrinsic bioremediation and that the presence of the oil stimulated the oil-degrading community members and were important in determining the fate of the deep-sea oil spill.

PESTICIDE CONTAMINATION

In a study using GeoChip 2.0 to examine three atrazine-contaminated aquifers and a background site, Liebich et al. (2009) detected more genes in the background site compared to the contaminated sites. The aquifer with the highest level of contamination had the highest number of genes, most involved in contaminant degradation, compared to the other contaminated samples. Atrazine-degradation genes were detected in all contaminated samples and verified by PCR. These results indicated that even small amounts of contaminant were enough to select for specific degrading populations.

River sediments from industrial pollutant and pesticide-contaminated sites were examined with GeoChip 2.0 and the results indicated that contaminant level was not a major driver in these systems (Taş et al., 2009). Instead, C/N ratio, depth, total Kjeldahl N, and location were the strongest drivers in determining the community structure. Most of the reductive dehalogenation genes detected were from *Dehalococcoides* spp., suggesting that this microorganism may play an important role in contaminant degradation in this system.

OTHER CONTAMINANTS

GeoChip 2.0 was used to examine phenanthrene-spiked soil microcosms to examine the effect of phenanthrene on microbial communities (Ding et al., 2012). Communities were examined after a 21-day incubation and compared with communities from day 0. A larger number of genes were detected in spiked soils compared to the control soils. Genes showing an increase in the spiked soils included dioxygenases involved in aromatic compound degradation, genes involved in the degradation of PAHs

(*nahA, rhda, nahQ, narR*), and genes involved in the degradation of one-ring aromatic compounds. In addition, an overall shift in community composition and structure was noted in spiked soils as determined by non-metric multidimensional scaling.

Another study examined microbial communities associated with a leachate-contaminated landfill using GeoChip 3.0 (Lu et al., 2012b). Groundwater samples were collected from wells along a flowpath of the landfill. Communities directly under the landfill and in the closest well had significantly lower functional gene diversity and richness. Genes involved in the anaerobic degradation of organic contaminates such as aromatic acids (*bclA, bbs, tutFDG*), phenoxyacetic acid herbicides (*ftdA*) atrazine (*atzABC, trzN, trzA, trzE*) were detected in all wells. Based on canonical correspondence analysis, the environmental variables (pH, sulfate, ammonia, and dissolved organic carbon) had significant effects on the community structure.

SUMMARY

The GeoChip arrays have been shown to be powerful tools in linking microbial function to ecosystem processes and are able to provide sensitive, specific, and potentially quantitative information. Use of this array in bioremediation studies have expanded our understanding of the microbial processes and communities at work in these sites and provide information necessary for the successful improvement and application of bioremediation strategies. Over the past decade, great improvements have been made in regards to microarray technology, design, and application. However, there are still technical hurdles that need to be overcome to further improve sensitivity and specificity in addition to better methods of nucleic acid extraction and purification. Improved bioinformatics tools are also needed to assist with data processing and analysis.

ACKNOWLEDGMENTS

REFERENCES

Adey, N. B., Lei, M., Howard, M. T., Jenson, J. B., Mayo, D. A., Butel, D. L., Coffin, S. C., Moyer, T. C., Slade, D. E., Spute, M. K., Hancock, A. M., Eisenhoffer, G. T., Dalley, B. K., and McNeely, M. R. (2002). Gains in sensitivity with a device that mixes microarray hybridization solution in a 25-μm-think chamber. *Anal. Chem.* 74, 6413–6417.

Albert, T. J., Molla, M. N., Muzny, D. M., Nazareth, L., Wheeler, D.,

Song, X., Richmond, T. A., Middle, C. M., Rodesch, M. J., Packard, C. J., Weinstock, G. M., and Gibbs, R. A. (2007). Direct selection of human genomic loci by microarray hybridization. *Nat. Methods* 4, 903–905.

Amann, R. I., Ludwig, W., and Schleifer, K. H. (1995). Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59, 143–169.

Bodrossy, L., Stralis-Pavese, N., Konrad-Köszler, M., Weilharter, A., Reichenauer, T. G., Schöfer, D., and 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, J. C., Radajewski, S., Weilharter, A., and Sessitsch, A. (2003). Development and validation of a diagnostic microbial microarray for

methanotrophs. *Environ. Microbiol.* 5, 566–582.

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

Branham, W. S., Melvin, C. D., Han, T., Gesai, V. G., Moland, C. L., Scully, A. T., and Fuscoe,

J. C. (2007). Elimination of laboratory ozone leads to a dramatic improvement in the reproducibility of microarray gene expression measurements. *BMC Biotechnol.* 7, 8. doi: 10.1186/1472-6750-7-8

Brodie, E. L., DeSantis, T. Z., Joyner, D. C., Baek, S. M., Larsen, J. T., Andersen, G. L., Hazen, T. C., Richardson, P. M., Herman, D. J., Tokunaga, T. K., Wan, J. M., and Firestone, M. K. (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, W. V., and, 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, D. R., Bakko, M. K., Krug, M. J., and Roberts, M. C. (2003). Identifying antimicrobial resistance genes with DNA microarrays. *Antimicrob. Agents Chemother.* 47, 3290–3295.

Chen, Q., Huaqun, Y., Luo, H., Xie, M., Qiu, G., and Liu, X. (2009). Micro-array based whole-genome hybridization for detection of microorganisms in acid mine drainage and bioleaching systems. *Hydrometallurgy* 95, 96–103.

Cho, J. C., and Tiedje, J. M. (2002). Quantitative detection of microbial genes by using DNA microarrays. *Appl. Environ. Microbiol.* 58, 1425–1430.

Cleven, B. E. E., Palka-Santini, M., Gielen, J., Meembor, S., Krönke, M., and Krut, O. (2006). Identification and characterization of bacterial pathogens causing bloodstream infections by DNA microarray. *J. Clin. Microbiol.* 44, 2389–2397.

Denef, V. J., Park, J., Rodrigues, J. L. M., Tsoi, T. V., Hashsham, S. A., and Tiedje, J. M. (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, E. A., Liss, S. N., and Fulthorpe, R. (2003). Monitoring gene expression in mixed microbial communities by using DNA microarrays. *Appl. Environ. Microbiol.* 69, 769–778.

Ding, G. C., Heuer, H., He, Z., Xie, J., Zhou, J., and Smalla, K. (2012). More functional genes and convergent overall functional patterns detected by GeoChip in phenanthrene spiked soils. *FEMS*

Microbiol. Ecol. doi: 10.1111/j.1574-6941.2012.01413.x [Epub ahead of print].

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

Epelde, L., Becerril, J. M., Kowalchuk, G. A., Deng, Y., Zhou, J.-Z., and Garbisu, C. (2010). Impact of metal pollution and *Thlaspi caerulescens* growth on soil microbial communities. *Appl. Environ. Microbiol.* 76, 7843–7853.

Fuhrman, J. A., and Campbell, L. (1998). Microbial Microdiversity. News and Views. *Nature* 393, 410–411.

Gao, H., Yang, Z. K., Gentry, T. J., Wu, L., Schadt, C. W., and Zhou, J. (2007). Microarray-based analysis of microbial community RNAs by whole-community RNA amplification. *Appl. Environ. Microbiol.* 73, 563–571.

Gentry, T. J., Wickham, G. S., Schadt, C. W., He, Z., and Zhou, J. (2006). Microarray application in microbial ecology research. *Microbial. Ecol.* 52, 159–175.

Gresham, D., Dunham, M. J., and Botstein, D. (2008). Comparing whole genomes using DNA microarrays. *Nat. Rev.* 9, 291–302.

Grubbs, F. (1969). Procedures for detecting outlying observations in samples. *Technometrics* 11, 469–479.

Gudnason, H., Dufva, M., Bang, D. D., and Wolff, A. (2008). An inexpensive and simple method for thermally stable immobilization of DNA on an unmodified glass surface: UV linking of poly(T)10-poly(C)10-tagged DNA probes. *Biotechniques* 45, 261–271.

Guschin, D. Y., Mobarry, B. K., Proudnikov, D., Stahl, D. A., Rittmann, B. E., and Mirzabekov, A. D. (1997). Oligonucleotide microchips as genosensors for determinative and environmental studies in microbiology. *Appl. Environ. Microbiol.* 63, 2397–2402.

Hazen, T. C., Dubinsky, E., DeSantis, T., Andersen, G., Piceno, Y., Singh, N., Jansson, J., Probst, A., Borglin, S., Fortney, J., Stringfellow, W., Bill, M., Conrad, M., Tom, L., Chavarria, K., Alusi, T., Lamendella, R., Joyner, D., Spier, C., Baelum, J., Auer, M., Zemla, M., Chakraborty, R., Sonnenthal, E., D'Haeseleer, P., Holman, H.-Y., Osman, S., Lu, Z. M., Van Nostrand, J., Deng, Y., Zhou, J. Z., and Mason, O. U. (2010). Deep-sea oil plume enriches Indigenous oil-degrading bacteria. *Science* 330, 204–208.

He, Z., Deng, Y., Van Nostrand, J. D., Tu, Q., Xu, M., Hemme, C. L., Li, X., Wu, L., Gentry, T. J., Yin, Y., Leibich, J., Hazen, T. C., and Zhou, J. (2010a). GeoChip 3.0 as a high-throughput tool for analyzing microbial community structure, composition and functional activity. *ISME J.* 4, 1167–1179.

He, Z., Xu, M., Deng, Y., Kang, S., Kellogg, L., Wu, L., Van Nostrand, J. D., Hobbie, S. E., Reich, P. B., and Zhou, J.-Z. (2010b). Metagenomic analysis reveals a marked divergence in the functional structure of belowground microbial communities at elevated CO_2. *Ecol. Lett.* 13, 564–575.

He, Z., Gentry, T. J., Schadt, C. W., Wu, L., Liebich, J., Chong, S. C., Wu, W. M., Gu, B., Jardine, P., Criddle, C., and Zhou, J. Z. (2007). GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *ISME J.* 1, 67–77.

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

He, Z., Wu, L. Y., Li, X. Y., Fields, M. W., and Zhou, J. Z. (2005b). Empirical establishment of oligonucleotide probe design criteria. *Appl. Environ. Microbiol.* 71, 3753–3760.

He, Z., and 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.

Hong, S.-H., Bunge, J., Jeon, S.-O., and Epstein, S. S. (2006). Predicting microbial species richness. *Proc. Natl. Acad. Sci. U.S.A.* 103, 117–122.

Hurt, R. A., Qiu, X., Wu, L., Roh, Y., Palumbo, A. V., Tiedje, J. M., and Zhou, J. (2001). Simultaneous recovery of RNA and DNA from soils and sediments. *Appl. Environ. Microbiol.* 67, 4495–4503.

Jenkins, B. D., Steward, G. F., Short, S. M., Ward, B. B., and Zehr, J. P. (2004). Fingerprinting diazotroph communities in the Chesapeake Bay. *Appl. Environ. Microbiol.* 70, 1767–1776.

Kimes, N. E., Van Nostrand, J. D., Weil, E., Zhou, J., and Morris, P. J. (2009). The microbial functional structure of *Montastraea faveolata*, an important Caribbean reef-building coral, differs between healthy and yellow-band diseased (YBD) colonies. *Environ. Microbiol.* 12, 541–556.

Kostić, T., Weilharter, A., Sessitsch, A., and Bodrossy, L. (2005). High-sensitivity, polymerase chain reaction-free detection of microorganisms and their functional genes

using 70-mer oligonucleotide diagnostic microarray. *Anal. Biochem.* 346, 333–335.

Ksiazek, T. G., Erdman, D., Goldsmith, C. S., Zaki, S. R., Peret, T., Emery, S., Tong, S., Urbani, C., Comer, J. A., Lim, W., Rollin, P. E., Dowell, S. F., Ling, A. E., Humphrey, C. D., Shieh, W. J., Guarner, J., Paddock, C. D., Rota, P., Fields, B., DeRisi, J., Yang, J. Y., Cox, N., Hughes, J. M., LeDuc, J. W., Bellini, W. J., Anderson, L. J., and the SARS Working Group. (2003). A novel coronavirus associated with severe acute respiratory syndrome. *N. Engl. J. Med.* 348, 1953–1966.

Kumar, A., Larsson, O., Parodi, D., and Liang, Z. (2000). Silanized nucleic acids: a general platform for DNA immobilization. *Nucleic Acids Res.* 28, e71.

Legendre, P., and Anderson, M. J. (1999). Distance-based redundancy analysis: testing multi-species responses in multi-factorial ecological experiments. *Ecol. Monogr.* 69, 1–24.

Leigh, M. B., Pellizari, V. H., Uhlík, O., Sutka, R., Rodrigues, J., Ostrom, N. E., Zhou, J., and Tiedje, J. M. (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., and 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., He, Z., Wu, L., Deng, Y., Li, G., and Zhou, J. (2010). Development of a common oligo reference standard (CORS) for microarray data normalization and comparison across different microbial communities. *Appl. Environ. Microbiol.* 76, 1088–1094.

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

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

Liang, Y. T., Van Nostrand, J. D., Deng, Y., He, Z. H., Wu, L. Y., Zhang, X., Li, G., and Zhou, J. (2011). Functional gene diversity of soil microbial communities from five oil-contaminated

fields in China. *ISME J.* 5, 403–413.

Liang, Y., Van Nostrand, J. D., N'Guessan, L. A., Peacock, A. D., Deng, Y., Long, P. E., Resch, C. T., Wu, L., He, Z., Li, G., Hazen, T. C., Lovley, D. R., and Zhou, J. (2012). Microbial functional gene diversity with a shift of subsurface redox condition during *in situ* uranium reduction. *Appl. Environ. Microbiol.* doi: 10.1128/AEM.06528-11. [Epub ahead of print].

Liebich, J., Schadt, C. W., Chong, S. C., He, Z., Rhee, S. K., and Zhou, J. (2006). Improvement of oligonucleotide probe design criteria for functional gene microarrays in environmental applications. *Appl. Environ. Microbiol.* 72, 1688–1691.

Liebich, J., Wachtmeister, T., Zhou, J., and Burauel, P. (2009). Degradation of diffuse pesticide contaminants: screening for microbial potential using a functional gene microarray. *Vadose Zone J.* 8, 703–710.

Loy, A., Lehner, A., Lee, N., Adamcsyk, J., Meier, H., Ernst, J., Schleifer, K. H., and Wagner, M. (2002). Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. *Appl. Environ. Microbiol.* 68, 5064–6081.

Lu, Z.-M., Deng, Y., Van Nostrand, J. D., He, Z., Voordeckers, J., Zhou, A., Lee, Y. J., Mason, O. U., Dubinsky, E. A., Chavarria, K. L., Tom, L. M., Fortney, J. L., Lamendella, R., Jansson, J. K., D'Haeseleer, P., Hazen, T. C., and Zhou, J. (2012a). Microbial gene functions enriched in the Deepwater Horizon deep-sea oil plume. *ISME J.* 6, 451–460.

Lu, Z., He, Z., Parisi, V. A., Kang, S., Deng, Y., Van Nostrand, J. D., Masoner, J. R., Cozzarelli, I. M., Suflita, J. M., and Zhou, J. (2012b). GeoChip-based analysis of microbial functional gene diversity in a landfill leachate-contaminated aquifer. *Environ. Sci. Technol.* 46, 5824–5833.

Lueders, T., and Friedrich, M. W. (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. Environ. Microbiol.* 69, 320–326.

Luo, J., Wu, W., Fienen, M. N., Jardine, P. M., Mehlhorn, T. L., Watson, D. B., Cirpka, O. A., Criddle, C. S., and Kitanidis, P. K. (2006). A nested-cell approach for *in situ* remediation. *Ground Water* 44, 266–274.

Mason, O. U., DiMeo-Savoie, C. A., Van Nostrand, J. D., Zhou, J., Fisk, M. R., and Giovannoni, S. J. (2009). Prokaryotic diversity, distribution, and preliminary insights into their role in biogeochemical cycling in marine basalts. *ISME J.* 3, 231–242.

McGrath, K. C., Thomas-Hall, S. R., Cheng, C. T., Leo, L., Alexa, A., Schmidt, S., and Schenk, P. M. (2008). Isolation and analysis of mRNA from environmental microbial communities. *J. Microbiol. Methods* 75, 172–176.

McKenna, P., Hoffmann, C., Minkah, N., Aye, P. P., Lackner, A., Liu, Z. (2008). The Macaque gut microbiome in health, lentiviral infection, and chronic enterocolitis. *PLoS Pathog.* 4, e20. doi: 10.1371/journal.ppat.0040020

Mehta, M. L. (1990). *Random Matrix Theory*. New York: Springer.

Mettel, C., Kim, Y., Shrestha, P. M., and Liesack, W. (2010). Extraction of RNA from soil. *Appl. Environ. Microbiol.* 76, 5995–6000.

Miller, S. M., Tourlousse, D. M., Stedtfeld, R. D., Baushke, S. W., Herzog, A. B., Wick, L. M., Rouillard, J. M., Gulari, E., Tiedje, J. M., and Hashsham, S. A. (2008). *In situ*-synthesized virulence and marker gene biochip for detection of bacterial pathogens in water. *Appl. Environ. Microbiol.* 74, 2200–2209.

Mockler, T. C., and Ecker, J. R. (2005). Applications of DNA tiling arrays for whole-genome analysis. *Genomics* 85, 1–15.

Murray, A. E., Lies, D., Li, G., Nealson, K., Zhou, J., and Tiedje, J. M. (2001). DNA–DNA hybridization to microarrays reveals gene-specific differences between closely related microbial genomes. *Proc. Natl. Acad. Sci. U.S.A.* 98, 9853–9858.

Ning, J., Liebich, J., Kästner, M., Zhou, J., Schäffer, A., and Burauel, P. (2009). Different influences of DNA purity indices and quantity on PCR-based DGGE and functional gene microarray in soil microbial community study. *Appl. Microbiol. Biotechnol.* 82, 983–993.

Okamoto, T., Suzuki, T., and Yamamoto, N. (2000). Microarray fabrication with covalent attachment of DNA using Bubble Jet technology. *Nat. Biotechnol.* 18, 438–441.

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

Okou, D. T., Steinberg, K. M., Middle, C., Cutler, D. J., Albert, T. J., and Zwick, M. E. (2007). Microarray-based genomic selection for high-throughput resequencing. *Nat. Methods* 4, 907–909.

Palka-Santini, M., Cleven, B. E., Eichinger, L., Krönke, M., and Krut, O. (2009). Large scale multiplex PCR improves pathogen detection by DNA microarrays. *BMC Microbiol.* 9, 1. doi: 10.1186/1471-2180-9-1

Parnell, J. J., Rompato, G., Latta, L. C. IV, Pfrender, M. E., Van Nostrand, J. D., He, Z., Zhou, J., Andersen, G., Champine, P., Ganesan, B., and Weimer, B. C. (2010). Functional biogeography as evidence of gene transfer in hypersaline microbial communities. *PLoS ONE* 5, e12919. doi: 10.1371/journal.pone.0012919

Poretsky, R. S., Gifford, S., Rinta-Kanto, J., Vila-Costa, M., and Moran, M. A. (2009a). Analyzing gene expression from marine microbial communities using environmental transcriptomics. *J. Vis. Exp.* 24, 1086.

Poretsky, R. S., Hewson, I., Sun, S., Allan, A. E., Zehr, J. P., and Moran, M. A. (2009b). Comparative day/night metatranscriptomic analysis of microbial communities in the North Pacific subtropical gyre. *Environ. Microbiol.* 11, 1358–1375.

Ramette, A., and Tiedje, J. M. (2007). Multiscale responses of microbial life in spatial distance and environmental heterogeneity in a patchy ecosystem. *Proc. Natl. Acad. Sci. U.S.A.* 104, 2761–2766.

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

Rhee, S. K., Liu, X., Wu, L., Chong, S. C., Wan, X., and 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, E. M., Pérez, E. X., Schadt, C. W., Zhou, J., and Massol-Deyá, A. A. (2006). Microbial diversity and bioremediation of a hydrocarbon-contaminated aquifer (Vega Baja, Puerto Rico). *Int. J. Environ. Res. Public Health* 3, 292–300.

Roesch, L. F. W., Fulthorpe, R. R., Riva, A., Casella, G., Hadwin, A. K. M., Kent, A. D., Daroub, S. H., Camargo, F. A., Farmerie, W. G., and Triplett, E. W. (2007). Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J.* 1, 283–290.

Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467–470.

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

Schloss, P. D., and Handelsman, J. (2006). Toward a census of bacteria in soil. *PLoS Comput. Biol.* 2, 786–793. doi: 10.1371/journal.pcbi.0020092

Sebat, J. L., Colwell, F. S., and Crawford, R. L. (2003). Metagenomic profiling: Microarray analysis of an environmental genomic library. *Appl. Environ. Microbiol.* 69, 4927–4934.

Serra, P., Colina, M., Fernández, J. M., Sevilla, L., and Morenza, J. L. (2004). Preparation of functional DNA microarrays through laser-induced forward transfer. *Appl. Phys. Lett.* 85, 1639–1641.

Shalon, D., Smith, S. J., and Brown, P. O. (1996). A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res.* 6, 639–645.

Small, J., Call, D. R., Brockman, F. J., Straub, T. M., and Chandler, D. P. (2001). Direct detection of 16S rRNA in soil extracts by using oligonucleotide microarrays. *Appl. Environ. Microbiol.* 67, 4708–4716.

Steward, G. F., Jenkins, B. D., Ward, B. B., and Zehr, J. P. (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., Csonton, J., Murrell, J. C., and Bodrossy, L. (2004). Optimization of diagnostic microarray for application in analysing landfill methanotroph communities under different plant covers. *Environ. Microbiol.* 6, 347–363.

Suzuki, M. T., and Giovannoni, S. J. (1996). Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl. Environ. Microbiol.* 62, 625–630.

Taroncher-Oldenburg, G., Griner, E. M., Francis, C. A., and Ward, B. B. (2003). Oligonucleotide microarray for the study of functional gene diversity in the nitrogen cycle in the

environment. *Appl. Environ. Microbiol.* 69, 1159–1171.

Taş, N., van Eekert, M. H. A., Schraa, G., Zhou, J., de Vos, W. M., and Smidt, H. (2009). Tracking functional guilds: *Dehalococcoides* spp. in European river basins contaminated with hexachlorobenzene. *Appl. Environ. Microbiol.* 75, 4696–4704.

ter Braak, C. J. F. (1986). Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67, 1167–1179.

Tiquia, S. M., Wu, L., Chong, S. C., Passovets, S., Xu, D., Xu, Y., and Zhou, J. (2004). Evaluation of 50-mer oligonucleotide arrays for detecting microbial populations in environmental samples. *Biotechniques* 36, 664–670, 672, 674–675.

Torsvik, V., Goksoyr, J., and Daae, F. L. (1990). High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol.* 56, 782–787.

Van Nostrand, J. D., Khijniak, T. V., Gentry, T. J., Novak, M. T., Sowder, A. G., Zhou, J. Z., Bertsch P. M., and Morris P. J. (2007). Isolation and characterization of four Grampositive nickel-tolerant microorganisms from contaminated sediments. *Microb. Ecol.* 53, 670–682.

Van Nostrand, J. D., Wu, L., Wu, W. M., Gentry, T. J., Huang, Z., Deng, Y., Carley, J., Carroll, S., He, Z., Gu, B., Luo, J., Criddle, C. S., Watson, D. B., Jardine, P. M., Marsh, T. L., Tiedje, J. M., Hazen, T. C., and Zhou, J. (2011). Dynamics of microbial community composition and function correlate with changes in site geochemistry during the *in situ* bioremediation of a uranium-contaminated aquifer. *Appl. Environ. Microbiol.* 77, 3860–3869.

Van Nostrand, J. D., Wu, W. M., Wu, L., Deng, Y., Carley, J., Carroll, S., He, Z., Gu, B., Luo, J., Criddle, C. S., Watson, D. B., Jardine, P. M., Marsh, T. L., Tiedje, J. M., Hazen, T. C., and Zhou, J. (2009). GeoChip-based analysis of functional microbial communities during the reoxidation of a bioreduced uranium-contaminated aquifer. *Environ. Microbiol.* 11, 2611–2626.

van Vliet, A. H. M. (2010). Next generation sequencing of microbial transcriptomics: challenges and opportunities. *FEMS Microbiol. Lett.* 302, 1–7.

Waldron, P. J., Wu, L., Van Nostrand, J. D., Schadt, C., Watson, D., Jardine, P., Palumbo, T., Hazen, T. C., and 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, D., Coscoy, L., Zylberberg, M., Avila, P. C., Boushey, H. A., Ganem, D., and DeRisi, J. L. (2002). Microarray-based detection and genotyping of viral pathogens. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15687–15692.

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

Wang, J., Van Nostrand, J. D., Wu, L., He, Z., Li, G., and Zhou, J.-Z. (2011). Microarray-based evaluation of whole-community genome DNA amplification methods. *Appl. Environ. Microbiol.* 77, 4241–4245.

Warnecke, P. M., Stirzaker, C., Melki, J. R., Millar, D. S., Paul, C. L., and Clark, S. J. (1997). Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic Acids Res.* 25, 4422–4426.

Wawrik, B., Boling, W., Van Nostrand, J., Xie, J. P., Zhou, J.-Z., and Bronk, D. A. (2012a). Assimilatory nitrate utilization by bacteria on the West Florida shelf as determined by stable isotope probing and functional microarray analysis. *FEMS Microbiol. Ecol.* 79, 400–411.

Wawrik, B., Mendivelso, M., Parisi, V. A., Suflita, J. M., Davidova, I. A., Marks, C. R., Van Nostrand, J. D., Liang, Y., Zhou, J., Huizinga, B. J., Strapoć, D., and Callaghan, A. V. (2012b). Field and laboratory studies on the bioconversion of coal to methane in the San Juan Basin. *FEMS Microbiol. Ecol.* 81, 26–42.

Whitman, W. B., Coleman, D. C., and Wiebe, W. J. (1998). Prokaryotes: the unseen majority. *Proc. Natl. Acad. Sci. U.S.A.* 95, 6578–6583.

Wilson, M., DeRisi, J., Kristensen, H. H., Imboden, P., Rane, S., Brown, P. O., and Schoolnik, G. K. (1999). Exploring drug-induced alterations in gene expression in Mycobacterium tuberculosis by microarray hybridization. *Proc. Natl. Acad. Sci. U.S.A.* 96, 12833–12838.

Wilson, W. J., Strout, C. L., DeSantis, T. Z., Stilwell, J. L., Carrano, A. V., and Andersen, G. L. (2002). Sequence-specific identification of 18 pathogenic microorganisms using microarray technology. *Mol. Cell. Probes* 16, 119–127.

Wu, L., Kellogg, L., Devol, A. H., Tiedje, J. M., and 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, C. W., and Zhou, J. (2006a). Microarray-based analysis of submicrogram quantities of microbial community DNAs by using whole-community genome amplification. *Appl. Environ. Microbiol.* 72, 4931–4941.

Wu, W. M., Carley, J., Fienen, M., Mehlhorn, T., Lowe, K., Nyman, J., Luo, J., Gentile, M. E., Rajan, R., Wagner, D., Hickey, R. F., Gu, B., Watson, D., Cirpka, O. A., Kitanidis, P. K., Jardine, P. M., and Criddle, C. S. (2006b). Pilot-scale *in situ* bioremediation of uranium in a highly contaminated aquifer 1: conditioning of a treatment zone. *Environ. Sci. Technol.* 40, 3978–3985.

Wu, W. M., Carley, J., Gentry, T., Ginder-Vogel, M. A., Fienen, M., Mehlhorn, T., Yan, H., Caroll, S., Pace, M. N., Nyman, J., Luo, J., Gentile, M. E., Fields, M. W., Hickey, R. F., Gu, B., Watson, D., Cirpka, O. A., Zhou, J., Fendork, S., Kitanidis, P. K., Jardine, P. M., and Criddle, C. S. (2006c). Pilot-scale *in situ* bioremediation of uranium in a highly contaminated aquifer. 2: U(VI) reduction and geochemical control of U(VI) bioavailability. *Environ. Sci. Technol.* 40, 3986–3995.

Wu, L., Thompson, D. K., Li, G., Hurt, R. A., Tiedje, J. M., and 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, D. K., Liu, X., Fields, M. W., Bagwell, C. E., Tiedje, J. M., and 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.

Xie, J. P., He, Z. H., Liu, X. X., Liu, X. D., Van Nostrand, J. D., Deng, Y., Wu, L., Qiu, G., and Zhou, J. (2011). GeoChip-based analysis of the functional gene diversity and metabolic potential of microbial communities in acid mine drainage. *Appl. Environ. Microbiol.* 77, 991–999.

Xie, J., Wu, L., Liu, X., Qiu, G., and Zhou, J. (2007). "Improved procedure for DNA extraction and purification from soil," in *108th General

Meeting of the American-Society-for-Microbiology, June 1–5*, Boston, MA, USA.

Xiong, J. B., Wu, L. Y., Tu, S. X., Van Nostrand, J. D., He, Z. H., Zhou, J.-Z., and Wang, G. J. (2010). Microbial communities and functional genes associated with soil arsenic contamination and the rhizosphere of the arsenic-hyperaccumulating plant *Pteris vittata* L. *Appl. Environ. Microbiol.* 76, 7277–7284.

Xu, M., Wu, W. M., Wu, L., He, Z., Van Nostrand, J. D., Deng, Y. Luo, J., Carley, J., Ginder-Vogel, M., Gentry, T. J., Gu, B., Watson, D., Jardine, P. M., Marsh, T. L., Tiedje, J. M., Hazen, T., Criddle, C. S., Zhou, J. (2010). Responses of microbial community functional structures to pilot-scale uranium *in situ* bioremediation. *ISME J.* 4, 1060–1070.

Yergeau, E., Kang, S., He, Z., Zhou, J., and Kowalchuk, G. A. (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., and 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, L., Hurek, T., and Reihold-Hurek, B. (2007a). A *nifH*-based oligonucleotide microarray for functional diagnostics of nitrogen-fixing microorganisms. *Microbial Ecol.* 53, 456–70.

Zhang, Y., Zhang, X., Liu, X., Xiao, Y., Qu, L., Wu, L., and Zhou, J. (2007b). 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.

Zhang, L., Srinivasan, U., Marrs, C. F., Ghosh, D., Gilsdorf, J. R., and Foxman, B. (2004). Library on a slide for bacterial comparative genomics. *BMC Microbiol.* 4, 12–18. doi: 10.1186/1471-2180-4-12

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

Zhou, J., Bruns, M. A., and Tiedje, J. M. (1996). DNA recovery from soils of diverse composition. *Appl. Environ. Microbiol.* 62, 316–322.

Zhou, J.-Z., Deng, Y., Luo, F., He, Z.-H., Tu, Q., and Zhi, X. Y. (2010). Functional molecular ecological networks. *mBio* 1, e00169-10.

Zhou, J., Kang, S., Schadt, C. W., and Garten, C. T. Jr. (2008). Spatial scaling of functional gene diversity across various microbial taxa. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7768–7773.

Zhou, J., and Thompson, D. K. (2002). Challenges in applying microarrays to environmental studies. *Curr. Opin. Biotechnol.* 13, 204–207.

Zhou, J., Wu, L., Deng, Y., Zhi, X., Jiang, Y., Tu, Q., Xie, J., Van Nostrand, J. D., He, Z., and Zhou, J.

(2011). Reproducibility and quantitation of amplicon sequencing-based detection. *ISME J.* 5, 1303–1313.

Zhou, X., and 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.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any

commercial or financial relationships that could be construed as a potential conflict of interest.