Detection and Characterization of Uncultivated Microorganisms Using Microarrays

Terry J. Gentry, Zhili He, and Jizhong Zhou (✉)

Contents

Abstract Microarrays have unprecedented potential for the high-throughput detection and characterization of uncultivated microorganisms. Several different types of arrays have been developed or adapted for the interrogation of microbial genomes and monitoring microbial population dynamics and/or activity in relation to various microbial processes such as bioremediation and biogeochemical cycling.

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

T.J. Gentry
Department of Soil & Crop Sciences, Texas A&M University, 2474 TAMU, 370 Olsen Blvd., College Station, TX 77843

Microbiol Monogr, doi:10.1007/7171_2008_6

Even though the number of such microarray studies has increased dramatically over the last few years, microarray analysis of uncultivated microorganisms still poses several challenges including a lack of sequence information for many organisms and related issues regarding the sensitivity and specificity of detection. As research continues to address these difficulties, and with further technological advances, microarrays will undoubtedly find even broader application to the investigation of uncultivated microorganisms, thus greatly increasing our understanding of the genetics, physiology, and distribution of these heretofore largely uncharacterized microorganisms.

1 Introduction

Research over the last few years has begun to reveal the incredible diversity of microorganisms that inhabit many environments, with one study suggesting as many as a million distinct genomes per gram of soil (Gans et al. 2005). While there is some debate over the accuracy of that specific estimation (Bunge et al. 2006; Volkov et al. 2006), it is now well established that there are enormous numbers and diversity of environmental microorganisms that we know little to nothing about – largely because the inability to cultivate these organisms has precluded their phenotypic characterization (Handelsman 2004). The development of molecular methods (e.g., polymerase chain reaction, PCR) has reduced the need for cultivation as a prerequisite for microbial characterization and has resulted in the discovery of many novel organisms; however, most of these methods are relatively of low throughput and reveal only limited information, usually phylogenetic, about the organisms they detect. Other methods such as metagenomic sequencing can reveal enormous amounts of information about uncultivated microorganisms or communities but are typically only used on a small number of samples. Microarray-based analyses have the potential to bridge this analytical gap.

Microarrays have unprecedented capability as a specific, sensitive, quantitative, and high-throughput technology for the detection and characterization of uncultivated microorganisms. Microarrays are currently available that can simultaneously detect thousands of different target organisms or genes, and this capacity continues to increase as the technology rapidly progresses (Brodie et al. 2006; He et al. 2007). Novel types of microarrays are continually being developed, but most arrays can be classified in one of five groups on the basis of the design and targets of the array: phylogenetic oligonucleotide arrays (POAs), functional gene arrays (FGAs), community genome arrays (CGAs), metagenomic arrays (MGA), and whole-genome open reading frame arrays (WGA). These microarrays are rapidly being developed and/or adapted to environmental microbiology-related research with nearly 40 such papers being published in 2006 alone (Wagner et al. 2007). Researchers have used different approaches for microarray-based analyses (Fig. 1) to investigate pertinent topics ranging from the elucidation of microorganisms involved in bioremediation

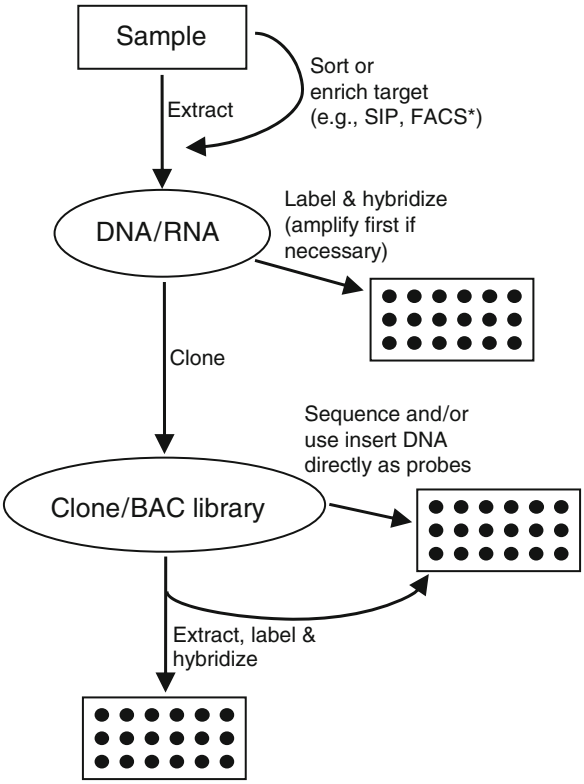

Fig. 1 Possible approaches for microarray analysis of uncultivated microorganisms (**SIP* Stable isotope probing, *FACS* Fluorescence-activated cell sorting)

and biogeochemical cycling processes to the detection of potential bioterrorism agents. The use of microarrays for studying uncultivated microorganisms still poses several challenges with respect to target sequence information, sensitivity, specificity, and data analysis, but applications will undoubtedly continue to increase as advances in microarray technology are made.

2 Types of Microarrays

2.1 Phylogenetic Oligonucleotide Arrays

Most POAs are based on specific regions of the ssu-rRNA (i.e., 16S rRNA) gene (Table 1). The rRNA gene has several characteristics that make it useful as a phylogenetic marker: e.g., (1) it exists in all organisms; (2) it appears to be rarely transferred between organisms; (3) it contains both conserved and variable regions of sequence;

Table 1 Microarrays for characterization of uncultivated microorganisms[a]

	Type of Array				
Property	POA	FGA	CGA	MGA	WGA
Probe template	Ribosomal rRNA genes	Functional genes	Whole genome	Environmental DNA or RNA	Open reading frames in whole genome
Probe target	rRNA or rDNA	mRNA or DNA	DNA	DNA or RNA	DNA or mRNA
Probe length	~18–25 nt	~50–70 nt oligos or ~200–1,000 nt PCR products	Whole genome	~1,000 nt	~25 nt oligos or ~200–3,000 nt PCR products
Information provided	Phylogenetic	Phylogenetic and functional	Phylogenetic	Functional	Phylogenetic and functional
Specificity	Species level or single nucleotide difference	< 80–90% sequence homology	Species – strain	≥Strain	Strain or single nucleotide difference
Sensitivity (ng of DNA)	~500	~1–8 (~0.01[b])	~0.2 (~0.1[b])	Undetermined	~0.1 (~0.04[b]) [PCR-based probes]

POA Phylogenetic oligonucleotide array, *FGA* Functional gene array, *CGA* Community genome array, *MGA* Metagenomic array, *WGA* Whole genome open reading frame array

Change Poa, fga, cga, mga, and wga to all caps like the table header and the Figure 1 caption.

[a]Adapted from (Zhou 2003)

[b]Quantitative detection limits using whole-community genome amplification methods are indicated in parentheses (Wu et al. 2006)

and (4) a vast amount of sequences exist in public databases such as the Ribosomal Database Project (RDP) (Cole et al. 2007).

One challenge for POAs targeting rRNA; (or DNA) is the formation of secondary structures in the target, which could result in decreased binding to the arrayed probes and thus reduce detection sensitivity. Several software programs exist that can identify self-complementarity in oligonucleotide probes, but it is difficult to predict the effects of target secondary structure on hybridization efficiency. Since the secondary structure of rRNA occurs in specific regions, it is possible to reduce the problem of target accessibility by designing probes complementary to regions less likely to form secondary structures. However, in contrast to the nonspecific binding of oligonucleotide probes to rRNA targets (false positives), which may be eliminated by stringent hybridization conditions, it is more difficult to eliminate instances where the probe and target cannot bind as a consequence of secondary structure in the target molecule (false negatives) (Peplies et al. 2003). One potential option to reduce this problem is to fragment or shear the target prior to hybridization (Chandler et al. 2003; Liu et al. 2007). Another potential option is the use of oligonucleotide helper probes designed to disrupt local secondary structure in the target by binding the target adjacent to the probe binding site (Peplies et al. 2003). This approach, however, can potentially be problematic if the helper probe binds too closely to either end of the capture probe binding site (Chandler et al. 2003). Furthermore, reduction of secondary structure in one region may result in the

formation of secondary structure(s) in other regions of the target, potentially affecting the binding sites of other probes (Peplies et al. 2003). It is also difficult to use the helper probe approach with high-density arrays since the helper probes also need to have a level of specificity similar to the capture probes. Recent research suggests that it may be possible to overcome the limitations imposed on probe design by the secondary structure of the rRNA target as long as the thermodynamic affinity of the probe for the target is sufficient and the incubation period of the hybridization is extended to optimize the kinetics of target unfolding and probe binding (Yilmaz and Noguera 2004; Yilmaz et al. 2006).

The conserved nature of rRNA genes may necessitate the use of short oligonucleotides (~20-mers) as POA probes, especially for high density arrays, in order to achieve the required level of specificity. A common POA format for high-density arrays consists of using multiple probes that perfectly match a specific target along with corresponding probes containing a single mismatch, usually at the central position (Wilson et al. 2002; El Fantroussi et al. 2003; Peplies et al. 2003; Urakawa et al. 2003; Brodie et al. 2006; Brodie et al. 2007). Larger signal intensity for a perfectly matched probe relative to the mismatched probe indicates detection of the target sequence.

The POAs are among the most commonly used microarrays owing to the well-established use of rRNA as a tool for characterizing microbial communities. Researchers have used POAs to examined microbial populations in a diverse set of environments including air (Brodie et al. 2007), lake water and sediments (Rudi et al. 2000; Castiglioni et al. 2004), estuary sediments (El Fantroussi et al. 2003) and enrichments (Koizumi et al. 2002), soil (Brodie et al. 2006), activated sludge (Adamczyk et al. 2003), and hypersaline cyanobacterial mats (Loy et al. 2002). The largest POA developed to date, dubbed the PhyloChip, used the perfect match/mismatch approach and contained over half million 25-mer oligonucleotide probes (Brodie et al. 2006). Approximately 300,000 of the probes were perfect match/mismatch probes targeting 16S rRNA genes and the rest were for image orientation, controls, or other purposes. The PhyloChip was designed on the basis of >30,000 16S rRNA gene sequences. These sequences were grouped into ~9,000 operational taxonomic units (OTUs). The researchers attempted to design a set of 11 or more probes specific to each OTU, which were then screened for potential cross-hybridization. Detection was considered positive if (1) the perfect match probe signal was >1.3 times the mismatch probe signal and (2) the difference in the signal between the two probes (perfect match – mismatch) was at least 500 times greater than the squared noise value. The PhyloChip has been successfully used to characterize microbial communities in urban aerosols and to monitor subsurface microbial populations during in situ reduction of uranium (discussed in more detail in a later section) (Brodie et al. 2006, 2007).

2.2 Functional Gene Arrays (GeoChips)

In contrast to POAs, which are primarily designed on the basis of phylogenetic genes (e.g., rRNA), FGAs are based on genes that encode for specific functional processes or phenotypic characteristics. (While rRNA genes are technically functional

genes critical to microbial growth, the term *functional gene* in this context has traditionally been used to indicate genes that encode proteins that directly impact some externally measurable process besides cell growth.) This enables FGAs to both provide information on the genetic capacity for and/or activity of a specific microbial process and also provide a level of phylogenetic classification. However, FGAs are not as well suited as POAs for the design of broad-group and universal probes because of the variability of many functional genes. Therefore, individual FGA probes are typically designed to detect only specific or small groups of organisms.

Several different groups of genes encoding key enzymes in various metabolic processes have been used for FGAs, including those involved in biogeochemical cycles (Wu et al. 2001; Cho and Tiedje 2002; Bodrossy et al. 2003; Taroncher-Oldenburg et al. 2003; Jenkins et al. 2004; Rhee et al. 2004; Steward et al. 2004; Stralis-Pavese et al. 2004; Tiquia et al. 2004; He et al. 2007), contaminant remediation (Cho and Tiedje 2002; Denef et al. 2003; Dennis et al. 2003; Rhee et al. 2004; He et al. 2007), and pathogen detection/characterization (Call et al. 2003; Grimm et al. 2004; Korczak et al. 2005; Perreten et al. 2005). Important characteristics that any gene should possess in order to be potentially be included in an FGA include that (1) it encodes an important enzyme or protein in the process of interest; (2) it is evolutionarily conserved yet has enough sequence divergence in different organisms to allow probe design for individual species; and (3) it has substantial sequence data from various isolates and environmental samples available in public databases. Because of enormous environmental sequence variability, it may be beneficial in many instances to initially do clone libraries for the gene and environment of interest in order to verify that the sequences used for FGA design adequately represent sequence diversity in a specific set of samples. Caution should be exercised when attempting to link FGA results based on "key" genes with the capacity for a specific function since the presence of a particular gene may not indicate the presence of all the genes necessary to complete the process of interest.

The FGA probes can be either PCR amplicons of various functional genes (Wu et al. 2001; Cho and Tiedje 2002) or synthesized oligonucleotides (Denef et al. 2003; Rhee et al. 2004). One advantage of PCR amplicon probes is that they can be constructed using conserved primers designed from known sequence data in other organisms and therefore do not require prior sequence information for the specific target organism/sample. It can, however, be practically impossible to obtain all the requisite isolates and environmental clones from their divergent sources in order to produce a comprehensive PCR amplicon FGA. In contrast, synthesized, oligonucleotide probes can be designed directly from available sequence data thereby eliminating the need to physically obtain (and maintain) the original isolates or clones. Oligonucleotide probes are also usually much shorter, being designed on the basis of only a portion of a particular gene. This potentially enables the design of better probes by allowing the avoidance of highly conserved regions or problematic regions of a gene. Additional probe design criteria also impact the specificity and sensitivity of both PCR amplicon and oligonucleotide probes as discussed in subsequent sections.

The development of comprehensive FGAs is currently limited only by the availability of essential isolates and/or sequence data and the capital resources needed for array construction. The largest FGA constructed to date, dubbed the GeoChip, contains over twenty-four thousand 50-mer oligonucleotide probes representing >10,000 genes involved in the C, N, and S cycles; organic contaminant degradation; and metals resistance and reduction (He et al. 2007). The GeoChip has been already been used in a variety of studies that delineated the microbial populations involved in bioremediation of uranium (He et al. 2007) and polychlorinated biphenyls (Leigh et al. 2007), examined the diversity of C and N cycling genes across an Antarctic transect (Yergeau et al. 2007), and interrogated the genomes of nickel-resistant bacteria (Van Nostrand et al. 2007).

2.3 Community Genome Arrays

The community genome array(CGA) is similar to membrane-based reverse sample genome probing (Greene and Voordouw 2003), but instead uses a nonporous hybridization surface and fluorescence-based detection. These features enable high throughput analyses but reduce the sensitivity of detection. Wu et al. (2004) developed a CGA which contained the genomic DNA of 67 different bacteria including α-, β-, and γ-*Proteobacteria* and Gram-positive bacteria. The array contained one species' genome per spot and could achieve species- to-strain-level differentiation depending on the hybridization conditions. The primary disadvantage of the CGAs designed so far is that only cultivated microorganisms have been included. However, with recent advances in the creation of large-insert, metagenomic libraries, it may also be possible to use DNA from uncultivated organisms for microarray construction.

2.4 Metagenomic Arrays

As evidenced by the publication of this book, the vast majority of environmental microorganisms are yet to be, and possibly cannot be, isolated using current techniques. This has led to recent development of the field of metagenomics (see Miao and Davies (2008)) (Handelsman 2004). Metagenomic approaches have been used to sequence entire communities in an acid mine drainage site (Tyson et al. 2004) and a portion of a more complex community from the Sargasso Sea (Venter et al. 2004). While it is currently difficult to assemble even the dominant genomes from microbially diverse sites such as surface soils and sediments (Torsvik and Ovreas 2002; Venter et al. 2004; Tringe et al. 2005), the combination of microarray and metagenomic technologies has the potential to reveal considerable information on these yet-uncultivated microorganisms.

MGAs are still in the early stages of development but have enormous potential for characterizing both the genetic capacity and activity of microbial communities. One of the few MGAs published so far was based on ~1-kb inserts amplified from 672 cosmids derived from a groundwater enrichment and was used as a high-throughput method to screen groundwater isolates and community DNA (Sebat et al. 2003). While the above array contained probes that were only ~1 kb, larger fragments from fosmid or bacterial artificial chromosome libraries could potentially be used to provide higher genomic throughput (Berry et al. 2003; Wu et al. 2004). Furthermore, it may be possible to create MGAs for measuring microbial activity if enough mRNA can be obtained to generate a cDNA library from which probes could be designed. The use of MGAs has great promise for the study of uncultivated microorganisms since one of the major current limitations for microarray analysis is the lack of sequence information.

2.5 Whole-Genome Open Reading Frame Arrays

WGAs contain probes for all (or at least many) of the open reading frames (ORFs) in a genome. Like FGAs, MGAs can be either based on PCR amplicon or oligonucleotide probes. One application of WGAs is to study the environmental transcriptomics of relatively simple systems – containing a single or few different types of organisms (Barnett et al. 2004). These arrays can also be used for comparative genomics with specific potential for studying lateral gene transfer and microevolution (Behr et al. 1999; Salama et al. 2000; Dong et al. 2001; Murray et al. 2001; Dziejman et al. 2002; Ochman et al. 2005; Ochman and Santos 2005). Custom WGAs can be designed, based only on sequence data, through companies such as NimbleGen Systems Inc. (http://www.nimblegen.com/products/cgr/index.html) for rapidly surveying microbial genomes through comparative genome sequencing via microarrays.

2.6 Isotope Arrays

While microarrays have unprecedented power for the high-throughput detection and characterization of microorganisms, their greatest utility lies perhaps in combined use with other complementary methods (Wagner et al. 2007). This approach can be used not only to validate microarray results but also to produce robust, synergistic tools for investigating microbial processes. One of the potentially most powerful combined approaches is the use of microarrays in conjunction with isotope techniques. By using labeled substrates (e.g., ^{13}C or ^{14}C), this system can achieve high-throughput differentiation between microorganisms which are actively involved in a process of interest and those that are not, based on incorporation of the labeled substrate by the active organisms. Both radioisotopes (e.g., ^{14}C) and stable isotopes (e.g., ^{13}C) have been successfully used for these purposes (Radajewski et al. 2000; Adamczyk et al. 2003; Polz et al. 2003; Wagner et al. 2006; Cebron et al. 2007; Leigh et al. 2007). One

distinction between radioisotopes and stable isotopes is that methods are currently available to detect radioisotopes directly on an array. The use of the stable isotopes requires separation of labeled and unlabeled nucleic acids, usually by centrifugation, prior to microarray hybridization, which can limit the sensitivity of stable isotope detection. However, this limitation may be eliminated in the future if methods are developed for directly detecting ^{13}C on arrays.

2.7 Other Types of Arrays

Numerous other types of microarrays have been used for microbial characterization – primarily of cultivated isolates. Random or digested genomic fragments have been used as probes on arrays for applications where the genome sequences of the target organisms were unknown (Cho and Tiedje 2001; Parro and Moreno-Paz 2003; Kim et al. 2004). For some applications, mixed-genome arrays containing random gene fragments from multiple organisms may enable greater discrimination (Wan et al. 2007). Likewise, randomly selected oligonucleotide arrays (consisting of 47 nonamer probes randomly generated based on the *E. coli* K-12 genome) have been used to fingerprint bacteria (Kingsley et al. 2002). In another approach researchers generated an array based on a library of 9,600 rRNA clones (Bent et al. 2006). The rRNA inserts were PCR amplified, arrayed, and then hybridized with a series of different oligonucleotide probes to generate a fingerprint for each clone. These fingerprints were then used to classify new, unidentified clones.

As the technology continues to progress, new array formats will undoubtedly expand the utility of microarrays for analysis of uncultivated microorganisms. While advances in extraction and amplification methods are making microarray analysis of environmental RNA more practical, this is only an indirect measure of microbial activity. The development and adaption of protein arrays would potentially allow direct measurement of enzymatic expression in environmental samples (Duburcq et al. 2004; Ro et al. 2005; Bjornstad et al. 2006; Gregson et al. 2006). Additional research may ultimately improve microarray automation and make it possible to construct microarrays that are field-deployable and can achieve real-time characterization of microbial communities (Chandler and Jarrell 2004; Liu et al. 2004; Liu and Zhu 2005).

3 Applications of Microarrays

3.1 Microbial Detection

The high-throughput potential of microarrays can be advantageous for several applications where it is necessary to simultaneously detect many different microorganisms, such as assessing microbial diversity and/or biogeography or screening

for potential bioterrorism agents (Table 2). In one of the most descriptive ecological applications of POAs to date, a POA containing 132 probes (18-mer) was used to characterize sulfate-reducing bacteria at four depths (ranging from 0 to 30 cm) in two acidic, low-sulfate fens (wetland soils) in Germany (Loy et al. 2004). The POA consisted of probes specific to the rRNAs of individual and groups of organisms, covering all known lineages of sulfate-reducing bacteria. The fens differed in iron content, vegetation, acidity, and to some degree, seasonal water saturation. The POA results indicated that stable sulfate-reducing populations varied little with depth within each of the two sites but were different between the sites. Members of the S*yntrophobacteraceae* were detected in the upper 30 cm of both sites, but *Desulfomonile* spp. were only found in one soil which also contained a more diverse sulfate-reducing community. These results were confirmed by direct PCR amplification with the appropriate group-specific rRNA primers and by the detection of the corresponding *dsrAB* genes from the samples. The same research group also developed a POA targeting all of the cultured and uncultured members of the *Rhodocyclales* (Loy et al. 2005). The array detected *Rhodocyclales* populations representing <1% of the total community, following *Rhodocyclales*-selective PCR amplification. The POA indicated the presence of several uncultured *Zoogloea*-, *Ferribacterium/Dechloromonas*-, and *Sterolibacterium*-like organisms in the activated sludge from an industrial wastewater treatment plant, which was corroborated by the results from a 16S rRNA gene clone library. The results also demonstrated that the *Rhodocyclales* community in the reactor, thought to represent the major denitrifiers in the system, had dramatically changed, possibly as a result of alterations in treatment plant operations.

One of the most comprehensive POAs developed so far contained 31,179 perfectly matched hierarchical 20-mer probes (with a corresponding number of single mismatch probes as negative, mismatch controls) targeting 1,945 prokaryotic and 431 eukaryotic sequences (Wilson et al. 2002). The array was used to investigate microorganisms collected from a 1.4 million liter air sample. The POA results generally agreed with those from an rRNA gene clone library, but could only resolve differences to the third level of phylogenetic rank, as defined by RDP, and could not identify individual species. Eight of 10 phylogenetic clusters detected by the array were represented in the rRNA gene clone libraries, and the organisms not detected had relatively low signals on the array. Approximately 7% of the clones were not detected by the POA, but these were from novel organisms not represented in the RDP or on the array. In contrast, there was no good correlation between the relative numbers of clones in each group and the signal intensity of that group detected by the array, indicating a potential limitation with respect to microbial quantitation with this system. Subsequent research has indicated that this array format does enable quantitative detection and may actually reveal more diversity than a normal-sized clone library (Desantis et al. 2005, 2007). A larger, more recent version of the array was used to monitor urban aerosols in two U.S. cities over a 17-week period (Brodie et al. 2007). The POA analysis indicated that the aerosols contained at least 1,800 diverse bacterial types including several organisms related to known pathogens and potential bioterrorism agents.

Table 2 Examples of studies that used microarrays for microbial characterization

Application	Sample	Type of array	Target	Amplification or selection	Results	Reference
Bioremediation	Soil	POA	rDNA	PCR	Demonstrated that populations of uranium-reducing bacteria were maintained following uranium reduction and were not the likely cause of subsequent uranium re-oxidation	Brodie et al. (2006)
Bioremediation	Groundwater	FGA	DNA	Whole-genome community amplification	Identified microorganisms whose population dynamics were correlated with in situ reduction of uranium in a contaminated aquifer	He et al. (2007)
Bioremediation	Mixed culture, reactor	FGA	mRNA	None	Detected increased expression of specific functional genes during degradation of a herbicide (2,4-D)	Dennis et al. (2003)
Bioremediation	Enrichment, soil	FGA	DNA, mRNA	None	Detection of naphthalene degradation genes in soils and those up-regulated in enrichments during degradation	Rhee et al. (2004)
Bioremediation	Soil	FGA	mRNA	Stable isotope probing	Identification of functional genes expressed during degradation of ^{13}C-labeled polychlorinated biphenyls	Leigh et al. (2007)
Biogeochemical cycling	Activated sludge	POA	rRNA	None	Combined with ^{14}C-based isotope techniques for parallel detection of ammonia-oxidizing bacteria and CO_2 fixation in nitrifying activated sludge	Adamczyk et al. (2003)
Biogeochemical cycling	Soil	POA	rDNA	PCR	Detection of sulfate-reducing bacteria in low-sulfate, acid fens	Loy et al. (2004)
Biogeochemical cycling	Soil	FGA	DNA	Stable isotope probing	Identified microorganisms actively consuming $^{13}CH_4$	Cebron et al. (2007)
Biogeochemical cycling	Soil	FGA	DNA	PCR	Linked differences in and methane oxidation with differences in vegetation and populations of specific groups of methanotrophs	Stralis-Pavese et al. (2004)
Environmental transcriptomics	River water	Other	mRNA	T7 RNA polymerase – amplification	Measured in situ gene expression for a dominant member of a highly acidic and metal contaminated ecosystem	Parro et al. (2007)
Metagenomic profiling	Riverine biofilm enrichment	MGA	DNA	None	Detection, isolation, and characterization of genes from uncultivated microorganisms	Sebat et al. (2003)
Microbial detection	Air	POA	rDNA	PCR	Characterization of airborne bacteria and detection of potential bioterrorism agents	Brodie et al. (2007)

POA Phylogenetic oligonucleotide array, *FGA* Functional gene array, *MGA* Metagenomic array

3.2 Microbial Population Dynamics and Activity

Several studies have used FGAs to investigate microbial involvement in biogeochemical processes including nitrogen fixation, nitrification, denitrification, sulfate reduction (Wu et al. 2001; Taroncher-Oldenburg et al. 2003; Jenkins et al. 2004; Steward et al. 2004; Tiquia et al. 2004), methane oxidation (Bodrossy et al. 2003; Stralis-Pavese et al. 2004), and remediation of organic and metal contaminants (Denef et al. 2003; Rhee et al. 2004; He et al. 2007; Leigh et al. 2007). In one of the largest-scale FGA applications to date, Stralis-Pavese et al. (2004) used an array containing 68 different 17- to 27-mer probes, primarily targeting the particulate methane monooxygenase (*pmoA*) genes of several methanotrophs, to investigate the impact of five different plant covers on methanotrophic activity in lysimeters under landfill-simulating conditions. The authors linked the methanotrophic community structure in the vegetated lysimeters, in which type II methanotrophs had a competitive advantage over type Ia methanotrophs, with increased methane oxidation relative to the nonvegetated lysimeters. Furthermore, the relative abundances of methanotrophs were lower in the lysimeters that did not receive biogas.

One of the first applications of the GeoChip FGA was to monitor groundwater microbial populations during in situ uranium reduction (He et al. 2007). The array detected several different c-type cytochrome and *dsrAB* (dissimilatory sulfite reductase) genes that were significantly correlated with uranium reduction. This included genes from *Geobacter*-type iron-reducing bacteria and *Desulfovibrio*-type sulfate-reducing bacteria, suggesting that these organisms actively or indirectly played a role in uranium reduction. Demonstrating the potential for biogeographical applications, Yergeau et al. (2007) used the GeoChip to examine soil microbial communities across an Antarctic latitudinal transect. The GeoChip revealed that cellulose degradation and denitrification genes were correlated with soil temperature, and N-fixation genes were linked to plots mainly vegetated by lichens. The GeoChip has also been used to detect biphenyl-degrading organisms in polychlorinated biphenyl–contaminated soil (Leigh et al. 2007).

One of the major potential benefits of FGAs is that they can also potentially be used to determine environmental gene expression by measuring mRNA. However, only a handful of studies have used FGAs for mRNA analysis to date (Dennis et al. 2003; Rhee et al. 2004; Bodrossy et al. 2006; Gao et al. 2007). Dennis et al. (2003) constructed a PCR amplicon FGA (271–1,300 bp fragments) containing probes for 64 genes including several from the 2,4-dichlorophenoxyacetic acid (2,4-D)-degradation pathway of *Ralstonia eutropha* JMP134 and related organisms. The FGA was used to test mixed cultures consisting of four isolates from a batch reactor treating pulp mill effluent and varying concentrations of *R. eutropha* JMP134. The cultures were amended with 2 mM of 2,4-D and incubated 6 h prior to mRNA extraction. Significant induction of 2,4-D degradation genes was detected from populations as low as 0.0037% (3.7×10^3 cells in 10^8 total community) to 3.7%, depending on the specific genes detected and sequence similarity of the probes that were used.

Rhee et al. (2004) used a 50-mer oligonucleotide FGA (a predecessor to the GeoChip that contained 1,662 probes) to determine both the presence and expression of naphthalene-degradation genes in soil enrichments. Previously contaminated soil was enriched with naphthalene. Four different naphthalene-degradation genes, three of which were from *Rhodococcus* spp., were detected at higher levels in the naphthalene-amended enrichment, based on DNA analysis. Likewise, FGA analysis of mRNA detected three different *Rhodococcus* sp. genes involved in naphthalene degradation that were upregulated (40- to 100-fold) in the naphthalene-amended enrichment, including two of the genes detected by DNA analysis. The results also indicated that other potential naphthalene-degrading organisms, whose genes were detected in the enrichments by the DNA hybridizations, were not the primary naphthalene degrading organisms under the tested conditions. An FGA has also been used to detect active methanotrophs in lysimeters simulating landfill covers (Bodrossy et al. 2006).

Most of the above studies that analyzed mRNA were relatively simple systems – mixed cultures or enrichments. In many cases, it remains challenging to use FGAs to analyze mRNAs from more complex environmental samples because of difficulties in extracting sufficient quantities of high-quality mRNA (Saleh-Lakha et al. 2005). While advances in extraction techniques (Hurt et al. 2001; Sessitsch et al. 2002; Burgmann et al. 2003) and newly available commercial kits (e.g., TruRNA from Atom Sciences, Inc. and FastRNA from Qbiogene) are helping with this process, it may be necessary to further purify many samples which may decrease mRNA yields. This can be especially problematic for low-biomass environments that already do not contain sufficient mRNA for FGA analysis. New approaches such as (1) the addition of a poly(A) tail to bacterial RNA for subsequent amplification (Botero et al. 2005) or (2) the amplification of whole community mRNAs using random primers with an attached T7 promoter sequence for subsequent amplification (Bodrossy et al. 2006; Moreno-Paz and Parro 2006; Gao et al. 2007) will likely expand the applications of microarray analyses for determinations of microbial activity in environmental samples.

Besides FGAs, other types of arrays also can be used to study different microbial processes. The PhyloChip POA containing >500,000 probes based on 16S rRNA gene sequences was used to examine microbial populations in subsurface material during uranium remediation (Brodie et al. 2006). PhyloChip analysis revealed that populations of metal-reducing bacteria (e.g., *Geobacteraceae*) were abundant during uranium reduction (induced by lactate biostimulation) and did not decrease during subsequent uranium re-oxidation. These results indicated that decreases in the populations of uranium-reducing bacteria were likely not responsible for the re-oxidation.

3.3 Microbial Characterization

Several different types of arrays (e.g., CGAs, FGAs, and WGAs) have been used to characterize microbial genomes. While many of the examples below are for

isolated microorganisms, similar analyses could be done for large insert metagenomic DNA. The main use of CGAs is to determine the genomic relatedness of isolated bacteria (or metagenomic DNA) to each other and also the organisms represented on the array. A CGA that contained the entire genomic DNA of 67 different bacteria (one species' genome per spot) including α-, β-, and γ-*Proteobacteria* and Gram-positive bacteria – most being *Azoarcus*, *Pseudomonas*, or *Shewanella* spp. – was designed as a tool for detecting specific microorganisms within a natural microbial community (Wu et al. 2004). The CGA could achieve species- to strain-level differentiation depending on the hybridization temperature used. The CGA was also capable of correlating differences in microbial communities in different environmental samples with differences in biogeochemical and physical properties between the sites.

Similarly, WGAs can be used to characterize microbial isolates or metagenomic DNA. Dong et al. (2001) used a WGA containing 96% of the annotated ORFs in *Escherichia coli* K-12 to comparatively interrogate the genome of a related bacterium, *Klebsiella pneumoniae* 342. Only 70% of *E. coli* K-12 ORFs were found in *K. pneumoniae* 342. Both bacteria shared many highly conserved genes including those for energy, amino acid, and fatty acid metabolism along with cofactor synthesis, cell division, DNA replication, transcription, translation, transport, and regulatory proteins. In contrast, many *E. coli* K-12 hypothetical and putative regulatory proteins, chaperones, and enzymes, in addition to genes thought to have been acquired from phage, plasmids, or transposons via lateral transfer, were not found in *K. pneumoniae*. Other researchers have used a WGA to find evidence for horizontal gene transfer in *Shewanella* spp. (Murray et al. 2001).

Van Nostrand et al. (2007) used the GeoChip FGA to characterize the genomes of four nickel-tolerant bacteria isolated from contaminated sediment at the U.S. Department of Energy Savannah River Site in Aiken, SC. Phenotypic assays had indicated that the bacteria were resistant to several metals and antibiotics. Subsequent GeoChip analysis detected between 20 and 86 metal-related genes, including some for arsenic, chromium, copper, mercury, and nickel in the genomes of the different isolates.

4 Challenges and Limitations of Microarray Analysis

4.1 Reliance on Cultivated Organisms and/or Known Sequences

Many types of microarrays require isolated organisms or prior knowledge of target sequences for the construction of probes. This represents one of the greatest challenges for the analysis of environmental samples given the vast number of unknown DNA sequences in many samples. Microorganisms critical to the process of interest may be completely overlooked if they do not have corresponding probes on the array. Furthermore, probes that are designed to be specific to known sequences can also cross-hybridize to similar, unknown sequences from related or unrelated genes.

This can potentially result in either an underestimated signal due to weak binding of a slightly divergent sequence or a completely misleading signal due to binding of a nontarget sequence. In any case, microarray data from environmental samples should be cautiously interpreted.

4.2 Sensitivity

Another challenge for microarray analysis is the limited sensitivity of detection, with usually only dominant organisms being identified. Most microarrays are printed on nonporous, planar glass slide platforms because this enables high-density printing. However, this reduces detection sensitivity several orders of magnitude as compared to membrane-based hybridizations, most likely due to less probe being attached to the nonporous surfaces (Cho and Tiedje 2002). New slide chemistries, such as ultra thin or three-dimensional platforms, may ultimately help to increase detection sensitivity while maintaining the capacity for high-throughput printing (Guschin et al. 1997; Urakawa et al. 2003; Zhou et al. 2004; Hesse et al. 2006; Sunkara et al. 2007).

The choice of probe format can also impact sensitivity. Oligonucleotide probes have several advantages for probe design, but they are typically ~10- to 100-fold less sensitive than longer PCR amplicon or CGA probes (Wu et al. 2001; Denef et al. 2003; Rhee et al. 2004; Wu et al. 2004). For example, detection limits of 0.2 ng of target genomic DNA for a CGA (Wu et al. 2004), 1 ng for a PCR amplicon FGA (Wu et al. 2001), and 5–8 ng for a 50-mer oligonucleotide FGA (Rhee et al. 2004) have been reported for hybridization in the absence of background DNA. In the presence of background DNA, sensitivity is further reduced about 10-fold (Rhee et al. 2004; Tiquia et al. 2004; Wu et al. 2004). For one 50-mer FGA, this detection limit corresponded to ~10^7 cells or 5% of the total community, which is similar to other published studies (Cho and Tiedje 2002). One study recently directly compared the detection sensitivity of 50-, 60-, and 70-mer oligonucleotide probes with PCR amplicon probes (He et al. 2005a). The oligonucleotide probes had detection limits of 25–100 ng of genomic DNA, while the PCR amplicon probes had a detection limit of 5 ng of genomic DNA. These limits roughly corresponded to 10^7 and 10^6 gene copies for the oligonucleotide and PCR amplicon probes, respectively. The 70-mer probes produced results most similar to the PCR amplicons. The detection sensitivities of other arrays (MGAs and WGAs) will depend on the specific probe design.

Advances in nucleic acid labeling methods may also help to increase sensitivity (Denef et al. 2003; Steward et al. 2004; Zhou and Zhou 2004). One study reported that tyramide signal amplification labeling increased the signal intensity of a 70-mer FGA ~10-fold compared to commonly used Cy dye-labeling techniques (Denef et al. 2003). Use of new labels such as nanoparticles (Zhou and Zhou 2004) or even the development of label-free arrays may ultimately further enhance detection sensitivity (Thompson et al. 2005; Liu and Bazan 2006).

While the approaches discussed above may help to detect dominant members of relatively high biomass communities, improved methods are needed for detecting less abundant microorganisms. Potential approaches for selectively enriching specific targets include (1) PCR amplification of specific genes using conserved primers (Bodrossy et al. 2003), (2) use of capture techniques such as magnetic beads (Tsai et al. 2003), and (3) directed cell separation using fluorescence in situ hybridization and flow cytometry (Podar et al. 2007). For low-biomass samples, or small numbers of sorted cells, nonspecific amplification of whole community DNA can be used to generate sufficient material for microarray analysis (see Podar et al. 2008).

A whole-community genome amplification (WCGA) procedure has been developed, based on phi 29 DNA polymerase, that can amplify nanogram quantities of DNA to microgram quantities with a linear relationship between the starting template and the final product (Wu et al. 2006). The method was capable of quantitatively detecting 0.04–125 ng (r^2 = 0.65–0.99) of pure culture DNA using a WGA, 0.1–1,000 ng (r^2 = 0.91) of genomic DNA using a CGA, and 0.01–250 ng (r^2 = 0.96–0.98) of environmental DNA using a 50-mer FGA. The approach allowed representative detection of individual genes from 1 to 500 ng of initial environmental (groundwater) DNA template. Smaller amounts of template, as low as 10 fg, could be detected via WCGA, but this affected the representativeness of amplification.

4.3 Specificity

The specificity of microarray probes is governed by several factors including probe design and hybridization conditions. While probes can be designed to meet specific binding characteristics, at least theoretically, there is always potential for cross-hybridization if used with environmental samples that contain unsequenced microbial populations (Chandler and Jarrell 2004). This issue can be reduced by obtaining more sequence data (i.e., clone libraries) for specific samples prior to microarray analysis, but will likely remain a concern for most environments.

It is, however, possible to improve probe specificity for specific targets. This is especially true for oligonucleotide probes that can be designed to avoid conserved regions of genes or those producing stable secondary structures. Arrays can differentiate a single mismatch in a probe–target hybridization using shorter oligonucleotide probes (~20-mers) (Wilson et al. 2002; Urakawa et al. 2003; Zhou et al. 2004). One of the more common formats for POAs uses at least one probe for a given target that perfectly matches a target sequence and a corresponding probe(s) that contains one mismatched nucleotide, usually at a central position (Brodie et al. 2006). Detection of the target sequence is usually determined on the basis of greater signal intensity (using some predetermined level) for the perfectly matched probe relative to the mismatched probe. Owing to the potential for unexpected probe behavior (higher signal for the mismatched probe), several different sets of probes are commonly included on these arrays for each target sequence. Thermal

dissociation curves for individual probe–target duplexes have been used to further improve the discriminating power of these arrays (Liu et al. 2001; El Fantroussi et al. 2003; Urakawa et al. 2003; Li et al. 2004; Eyers et al. 2006).

While short oligonucleotide probes are often preferred for POAs in order to achieve the desired level of specificity, it is possible to use longer oligonucleotide probes (~40- to 70-mers) for FGAs owing to the more variable nature of most functional genes. This allows for an increase in detection sensitivity while still achieving species-level specificity. Different FGAs have been reported to be capable of differentiating sequences with less than 87–94% sequence identity using oligonucleotide probes (50 or 70-mers). This is slightly higher than the 80–85% reported for a PCR amplicon FGA (Wu et al. 2001; Taroncher-Oldenburg et al. 2003; Rhee et al. 2004). Besides percent sequence identity, other factors including long stretches of a probe that perfectly match a nontarget sequence (Kane et al. 2000; Chen et al. 2006), the position of mismatches (Letowski et al. 2004), and the free energy of probe–target duplexes can also affect specificity (Li and Stormo 2001; Held et al. 2003). While it may be difficult to design probes that optimally match all these criteria using a sequential test and select process for each factor, it may be possible to use more relaxed criteria but yet produce specific probes if the factors are simultaneously considered during probe design (Liebich et al. 2006). This would increase the number of probes that can be designed for a given target and could be very important when designing probes from very similar sequence data (i.e., clone libraries) for which only a limited number of specific probes can be designed.

Several different software programs are currently available for the design of oligonucleotide microarray probes; however, many of these were developed for use with individual genomes and may design nonspecific probes when used with environmental sequence data (Li et al. 2005). A relatively new software program, called CommOligo, is now available that is capable of designing specific probes from orthologous gene sequences such as those produced by clone libraries (Li et al. 2005). CommOligo can design single or multiple unique probes for each sequence using multiple, user-specified criteria, including maximal sequence identity, maximal length of continuous perfectly matched nucleotides, free energy, self-binding, melting temperature, and GC content. Another feature of CommOligo is the ability to design group-specific probes for clusters of sequences whose sequence identities are too similar to each other to allow the design of unique probes for each sequence.

The specificity of microarray hybridizations can also be affected by increasing or decreasing the stringency of the hybridization conditions (e.g., temperature, formamide concentration, etc.) (Wu et al. 2001; Letowski et al. 2004; He et al. 2005b). For example, the specificity of one CGA could be increased from species-level to strain-level detection by increasing the hybridization temperature from 55 to 65 or 75°C (Wu et al. 2004). This may be useful for some applications in order to achieve more or less specificity depending on the research objective, but it also illustrates the need for careful use of microarrays in order to achieve the appropriate level of hybridization stringency for which the array was designed.

4.4 Quantitation

Several studies have indicated that microarrays can quantitatively detect target sequences in environmental samples (Wu et al. 2001, 2004; Rhee et al. 2004; Desantis et al. 2005). It has been demonstrated that short oligonucleotide arrays (e.g., POAs, WGAs) based on perfect match/mismatch probes pairs can quantitatively detect targets (r = 0.917) in the presence of background DNA (Desantis et al. 2005). However, POAs of this format are commonly used with PCR-amplified nucleic acids, which may complicate quantitative detection. Another study reported a good linear relationship (r^2 = 0.94) between a mixture of 11 different genes, varying in concentration from 1 pg to 1 ng, that were hybridized to a PCR amplicon FGA (Wu et al. 2001). Similarly, Rhee et al. (2004) found a linear relationship (r^2 = 0.95–0.99) for multiple genes, over a range of 75–1,000 ng genomic DNA, detected by a 50-mer FGA in the presence of background DNA. This study also demonstrated that FGA detection of mRNA was linear (r^2 = 0.96–0.99) over a range of 5.0×10^7–1.6×10^9 cells in the presence of background RNA. Linear detection (r^2 = 0.98), from 25 to 1,000 ng of DNA, has been reported for at least one CGA (Wu et al. 2004). Even with these examples of successful quantitative detection by microarrays, caution should be used when conducting experiments and interpreting data because of the potential variability introduced by multiple analytical steps including DNA extraction, amplification, labeling, hybridization, and analysis.

4.5 Data Analysis and Standardization

Another challenge for microarray analysis of uncultivated microbial communities is the availability and standardization of methods for data analysis. Many researchers currently adapt statistical methods that were initially developed for functional genomics of pure cultures. In some cases, it may be possible to use traditional statistical methods to analyze data, but these methods may not be adequate for analyzing the complex data sets that are frequently generated from environmental samples (Eyers et al. 2006). These problems will only be magnified as more comprehensive arrays are developed that will require more involved methods such as artificial neural networks in order to interpret the results (Pozhitkov et al. 2005). Furthermore, researchers commonly use different analytical methods for microarray construction, labeling, and hybridization, which can make it difficult to compare data between experiments and laboratories (Wilkes et al. 2007). The development and consistent use of universal standards for microarray experiments would help address these issues. One approach that appears promising is the use of labeled oligonucleotides or DNA fragments, spiked into the hybridization solution, as a control (Cho and Tiedje 2002; Dudley et al. 2002). This allows microarray results to be normalized on the basis of the signal intensity resulting from hybridization of these controls with corresponding control probes on the array.

5 Concluding Remarks

Comprehensive microarrays are now available for the study of uncultivated microorganisms. As the technology continues to progress rapidly, new advances will undoubtedly expand the potential of microarrays for microbial characterization even further. However, the true power of these arrays will begin to be realized only as they are used to investigate complex microbial communities and interactions in a variety of environments.

Acknowledgments The authors' efforts in preparing this chapter were supported by the United States Department of Energy Office of Science as part of its Genomics:GTL program through the Virtual Institute of Microbial Stress and Survival (VIMSS; http://vimss.lbl.gov) and Environmental Remediation Science Program of the Office of Biological and Environmental Research, Office of Science.

References

Adamczyk J, Hesselsoe M, Iversen N, Horn M, Lehner A, Nielsen PH, Schloter M, Roslev P, Wagner M (2003) The isotope array, a new tool that employs substrate-mediated labeling of rRNA for determination of microbial community structure and function. Appl Environ Microbiol 69:6875–6887

Barnett MJ, Toman CJ, Fisher RF, Long SR (2004) A dual-genome Symbiosis Chip for coordinate study of signal exchange and development in a prokaryote-host interaction. Proc Natl Acad Sci USA 101:16636–16641

Behr MA, Wilson MA, Gill WP, Salamon H, Schoolnik GK, Rane S, Small PM (1999) Comparative genomics of BCG vaccines by whole-genome DNA microarray. Science 284:1520–1523

Bent E, Yin B, Figueroa A, Ye J, Fu Q, Liu Z, McDonald V, Jeske D, Jiang T, Borneman J (2006) Development of a 9600-clone procedure for oligonucleotide fingerprinting of rRNA genes: utilization to identify soil bacterial rRNA genes that correlate in abundance with the development of avocado root rot. J Microbiol Methods 67:171–180

Berry AE, Chiocchini C, Selby T, Sosio M, Wellington EM (2003) Isolation of high molecular weight DNA from soil for cloning into BAC vectors. FEMS Microbiol Lett 223:15–20

Bjornstad A, Larsen BK, Skadsheim A, Jones MB, Andersen OK (2006) The potential of ecotoxicoproteomics in environmental monitoring: biomarker profiling in mussel plasma using ProteinChip array technology. J Toxicol Environ Health A 69:77–96

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

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

Botero LM, D'Imperio S, Burr M, McDermott TR, Young M, Hassett DJ (2005) Poly(A) polymerase modification and reverse transcriptase PCR amplification of environmental RNA. Appl Environ Microbiol 71:1267–1275

Brodie EL, Desantis TZ, Joyner DC, Baek SM, Larsen JT, Andersen GL, Hazen TC, Richardson PM, Herman DJ, Tokunaga TK, Wan JM, Firestone MK (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

Brodie EL, DeSantis TZ, Parker JP, Zubietta IX, Piceno YM, Andersen GL (2007) Urban aerosols harbor diverse and dynamic bacterial populations. Proc Natl Acad Sci USA 104:299–304

Bunge J, Epstein SS, Peterson DG (2006) Comment on "Computational improvements reveal great bacterial diversity and high metal toxicity in soil". Science 313:918; author reply 918

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

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

Castiglioni B, Rizzi E, Frosini A, Sivonen K, Rajaniemi P, Rantala A, Mugnai MA, Ventura S, Wilmotte A, Boutte C, Grubisic S, Balthasart P, Consolandi C, Bordoni R, Mezzelani A, Battaglia C, De Bellis G (2004) Development of a universal microarray based on the ligation detection reaction and 16S rRNA gene polymorphism to target diversity of cyanobacteria. Appl Environ Microbiol 70:7161–7172

Cebron A, Bodrossy L, Stralis-Pavese N, Singer AC, Thompson IP, Prosser JI, Murrell JC (2007) Nutrient amendments in soil DNA stable isotope probing experiments reduce the observed methanotroph diversity. Appl Environ Microbiol 73:798–807

Chandler DP, Jarrell AE (2004) Automated purification and suspension array detection of 16S rRNA from soil and sediment extracts by using tunable surface microparticles. Appl Environ Microbiol 70:2621–2631

Chandler DP, Newton GJ, Small JA, Daly DS (2003) Sequence versus structure for the direct detection of 16S rRNA on planar oligonucleotide microarrays. Appl Environ Microbiol 69:2950–2958

Chen YA, Chou CC, Lu X, Slate EH, Peck K, Xu W, Voit EO, Almeida JS (2006) A multivariate prediction model for microarray cross-hybridization. BMC Bioinformatics 7:101

Cho JC, Tiedje JM (2001) Bacterial species determination from DNA-DNA hybridization by using genome fragments and DNA microarrays. Appl Environ Microbiol 67:3677–3682

Cho JC, Tiedje JM (2002) Quantitative detection of microbial genes by using DNA microarrays. Appl Environ Microbiol 68:1425–1430

Cole JR, Chai B, Farris RJ, Wang Q, Kulam-Syed-Mohideen AS, McGarrell DM, Bandela AM, Cardenas E, Garrity GM, Tiedje JM (2007) The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. Nucleic Acids Res 35:D169–D172

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

Desantis TZ, Stone CE, Murray SR, Moberg JP, Andersen GL (2005) Rapid quantification and taxonomic classification of environmental DNA from both prokaryotic and eukaryotic origins using a microarray. FEMS Microbiol Lett 245:271–278

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

Dong Y, Glasner JD, Blattner FR, Triplett EW (2001) Genomic interspecies microarray hybridization: rapid discovery of three thousand genes in the maize endophyte, *Klebsiella pneumoniae* 342, by microarray hybridization with *Escherichia coli* K-12 open reading frames. Appl Environ Microbiol 67:1911–1921

Duburcq X, Olivier C, Malingue F, Desmet R, Bouzidi A, Zhou F, Auriault C, Gras-Masse H, Melnyk O (2004) Peptide-protein microarrays for the simultaneous detection of pathogen infections. Bioconjug Chem 15:307–316

Dudley AM, Aach J, Steffen MA, Church GM (2002) Measuring absolute expression with microarrays with a calibrated reference sample and an extended signal intensity range. Proc Natl Acad Sci USA 99:7554–7559

Dziejman M, Balon E, Boyd D, Fraser CM, Heidelberg JF, Mekalanos JJ (2002) Comparative genomic analysis of *Vibrio cholerae*: genes that correlate with cholera endemic and pandemic disease. Proc Natl Acad Sci U S A 99:1556–1561

El Fantroussi S, Urakawa H, Bernhard AE, Kelly JJ, Noble PA, Smidt H, Yershov GM, Stahl DA (2003) Direct profiling of environmental microbial populations by thermal dissociation analysis of native rRNAs hybridized to oligonucleotide microarrays. Appl Environ Microbiol 69:2377–2382

Eyers L, Smoot JC, Smoot LM, Bugli C, Urakawa H, McMurry Z, Siripong S, El-Fantroussi S, Lambert P, Agathos SN, Stahl DA (2006) Discrimination of shifts in a soil microbial community associated with TNT-contamination using a functional ANOVA of 16S rRNA hybridized to oligonucleotide microarrays. Environ Sci Technol 40:5867–5873

Gans J, Wolinsky M, Dunbar J (2005) Computational improvements reveal great bacterial diversity and high metal toxicity in soil. Science 309:1387–1390

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

Greene EA, Voordouw G (2003) Analysis of environmental microbial communities by reverse sample genome probing. J Microbiol Methods 53:211–219

Gregson BP, Millie DF, Cao C, Fahnenstiel GL, Pigg RJ, Fries DP (2006) Simplified enrichment and identification of environmental peptide toxins using antibody-capture surfaces with subsequent mass spectrometry detection. J Chromatogr A 1123:233–238

Grimm V, Ezaki S, Susa M, Knabbe C, Schmid RD, Bachmann TT (2004) Use of DNA microarrays for rapid genotyping of TEM beta-lactamases that confer resistance. J Clin Microbiol 42:3766–3774

Guschin D, Yershov G, Zaslavsky A, Gemmell A, Shick V, Proudnikov D, Arenkov P, Mirzabekov A (1997) Manual manufacturing of oligonucleotide, DNA, and protein microchips. Anal Biochem 250:203–211

Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms. Microbiol Mol Biol Rev 68:669–685

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

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

Held GA, Grinstein G, Tu Y (2003) Modeling of DNA microarray data by using physical properties of hybridization. Proc Natl Acad Sci USA 100:7575–7580

Hesse J, Jacak J, Kasper M, Regl G, Eichberger T, Winklmayr M, Aberger F, Sonnleitner M, Schlapak R, Howorka S, Muresan L, Frischauf AM, Schutz GJ (2006) RNA expression profiling at the single molecule level. Genome Res 16:1041–1045

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

Jenkins BD, Steward GF, Short SM, Ward BB, Zehr JP (2004) Fingerprinting diazotroph communities in the Chesapeake Bay by using a DNA macroarray. Appl Environ Microbiol 70:1767–1776

Kane MD, Jatkoe TA, Stumpf CR, Lu J, Thomas JD, Madore SJ (2000) Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. Nucleic Acids Res 28:4552–4557

Kim BC, Park JH, Gu MB (2004) Development of a DNA microarray chip for the identification of sludge bacteria using an unsequenced random genomic DNA hybridization method. Environ Sci Technol 38:6767–6774

Kingsley MT, Straub TA, Call DR, Daly DS, Wunschel SC, Chandler DP (2002) Fingerprinting closely related *Xanthomonas* pathovars with random nonamer oligonucleotide microarrays. Appl Environ Microbiol 68:6361–6370

Koizumi Y, Kelly JJ, Nakagawa T, Urakawa H, El-Fantroussi S, Al-Muzaini S, Fukui M, Urushigawa Y, Stahl DA (2002) Parallel characterization of anaerobic toluene- and ethylbenzene-degrading microbial consortia by PCR-denaturing gradient gel electrophoresis, RNA-DNA membrane hybridization, and DNA microarray technology. Appl Environ Microbiol 68:3215–3225

Korczak B, Frey J, Schrenzel J, Pluschke G, Pfister R, Ehricht R, Kuhnert P (2005) Use of diagnostic microarrays for determination of virulence gene patterns of Escherichia coli K1, a major cause of neonatal meningitis. J Clin Microbiol 43:1024–1031

Leigh MB, Pellizari VH, Uhlik 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

Letowski J, Brousseau R, Masson L (2004) Designing better probes: effect of probe size, mismatch position and number on hybridization in DNA oligonucleotide microarrays. J Microbiol Meth 57:269–278

Li ES, Ng JK, Wu JH, Liu WT (2004) Evaluating single-base-pair discriminating capability of planar oligonucleotide microchips using a non-equilibrium dissociation approach. Environ Microbiol 6:1197–1202

Li F, Stormo GD (2001) Selection of optimal DNA oligos for gene expression arrays. Bioinformatics 17:1067–1076

Li XY, He ZL, Zhou JZ (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

Liu B, Bazan GC (2006) Synthesis of cationic conjugated polymers for use in label-free DNA microarrays. Nat Protoc 1:1698–1702

Liu RH, Yang J, Lenigk R, Bonanno J, Grodzinski P (2004) Self-contained, fully integrated biochip for sample preparation, polymerase chain reaction amplification, and DNA microarray detection. Anal Chem 76:1824–1831

Liu WT, Zhu L (2005) Environmental microbiology-on-a-chip and its future impacts. Trends Biotechnol 23:174–179

Liu WT, Mirzabekov AD, Stahl DA (2001) Optimization of an oligonucleotide microchip for microbial identification studies: a non-equilibrium dissociation approach. Environ Microbiol 3:619–629

Liu WT, Guo H, Wu JH (2007) Effects of target length on the hybridization efficiency and specificity of rRNA-based oligonucleotide microarrays. Appl Environ Microbiol 73:73–82

Loy A, Lehner A, Lee N, Adamczyk J, Meier H, Ernst J, Schleifer KH, 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 68:5064–5081

Loy A, Kusel K, Lehner A, Drake HL, Wagner M (2004) Microarray and functional gene analyses of sulfate-reducing prokaryotes in low-sulfate, acidic fens reveal cooccurrence of recognized genera and novel lineages. Appl Environ Microbiol 70:6998–7009

Loy A, Schulz C, Lucker S, Schopfer-Wendels A, Stoecker K, Baranyi C, Lehner A, Wagner M (2005) 16S rRNA gene-based oligonucleotide microarray for environmental monitoring of the betaproteobacterial order "*Rhodocyclales*". Appl Environ Microbiol 71:1373–1386

Miao V, Davies J (2008) Metagenomics and antibiotic discovery from uncultivated bacteria. Microbil Monogr. doi: 10.1007/7171_2008_8

Moreno-Paz M, Parro V (2006) Amplification of low quantity bacterial RNA for microarray studies: time-course analysis of Leptospirillum ferrooxidans under nitrogen-fixing conditions. Environ Microbiol 8:1064–1073

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

Ochman H, Santos SR (2005) Exploring microbial microevolution with microarrays. Infect Genet Evol 5:103–108

Ochman H, Lerat E, Daubin V (2005) Examining bacterial species under the specter of gene transfer and exchange. Proc Natl Acad Sci USA 102 Suppl 1:6595–6599

Parro V, Moreno-Paz M (2003) Gene function analysis in environmental isolates: the nif regulon of the strict iron oxidizing bacterium *Leptospirillum ferrooxidans*. Proc Natl Acad Sci USA 100:7883–7888

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

Peplies J, Glockner FO, Amann R (2003) Optimization strategies for DNA microarray-based detection of bacteria with 16S rRNA-targeting oligonucleotide probes. Appl Environ Microbiol 69:1397–1407

Perreten V, Vorlet-Fawer L, Slickers P, Ehricht R, Kuhnert P, Frey J (2005) Microarray-based detection of 90 antibiotic resistance genes of gram-positive bacteria. J Clin Microbiol 43:2291–2302

Podar, Keller, Hugenholtz (2008) Single cell whole genome amplification of uncultivated organisms. Microbil Monogr. doi: 10.1007/7171_2008_10

Podar M, Abulencia CB, Walcher M, Hutchison D, Zengler K, Garcia JA, Holland T, Cotton D, Hauser L, Keller M (2007) Targeted access to the genomes of low-abundance organisms in complex microbial communities. Appl Environ Microbiol 73:3205–3214

Polz MF, Bertilsson S, Acinas SG, Hunt D (2003) A(r)Ray of hope in analysis of the function and diversity of microbial communities. Biol Bull 204:196–199

Pozhitkov A, Chernov B, Yershov G, Noble PA (2005) Evaluation of gel-pad oligonucleotide microarray technology by using artificial neural networks. Appl Environ Microbiol 71:8663–8676

Radajewski S, Ineson P, Parekh NR, Murrell JC (2000) Stable-isotope probing as a tool in microbial ecology. Nature 403:646–649

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

Ro HS, Jung SO, Kho BH, Hong HP, Lee JS, Shin YB, Kim MG, Chung BH (2005) Surface plasmon resonance imaging-based protein array chip system for monitoring a hexahistidine-tagged protein during expression and purification. Appl Environ Microbiol 71:1089–1092

Rudi K, Skulberg OM, Skulberg R, Jakobsen KS (2000) Application of sequence-specific labeled 16S rRNA gene oligonucleotide probes for genetic profiling of cyanobacterial abundance and diversity by array hybridization. Appl Environ Microbiol 66:4004–4011

Salama N, Guillemin K, McDaniel TK, Sherlock G, Tompkins L, Falkow S (2000) A whole-genome microarray reveals genetic diversity among *Helicobacter pylori* strains. Proc Natl Acad Sci USA 97:14668–14673

Saleh-Lakha S, Miller M, Campbell RG, Schneider K, Elahimanesh P, Hart MM, Trevors JT (2005) Microbial gene expression in soil: methods, applications and challenges. J Microbiol Methods 63:1–19

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

Sessitsch A, Gyamfi S, Stralis-Pavese N, Weilharter A, Pfeifer U (2002) RNA isolation from soil for bacterial community and functional analysis: evaluation of different extraction and soil conservation protocols. J Microbiol Methods 51:171–179

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

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

Sunkara V, Hong BJ, Park JW (2007) Sensitivity enhancement of DNA microarray on nano-scale controlled surface by using a streptavidin-fluorophore conjugate. Biosens Bioelectron 22:1532–1537

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

Thompson M, Cheran LE, Zhang MQ, Chacko M, Huo H, Sadeghi S (2005) Label-free detection of nucleic acid and protein microarrays by scanning Kelvin nanoprobe. Biosens Bioelectron 20:1471–1481

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 36:664–675

Torsvik V, Ovreas L (2002) Microbial diversity and function in soil: from genes to ecosystems. Curr Opin Microbiol 5:240–245

Tringe SG, von Mering C, Kobayashi A, Salamov AA, Chen K, Chang HW, Podar M, Short JM, Mathur EJ, Detter JC, Bork P, Hugenholtz P, Rubin EM (2005) Comparative metagenomics of microbial communities. Science 308:554–557

Tsai YL, Le JY, Olson BH (2003) Magnetic bead hybridization to detect enterotoxigenic *Escherichia coli* strains associated with cattle in environmental water sources. Can J Microbiol 49:391–398

Tyson GW, Chapman J, Hugenholtz P, Allen EE, Ram RJ, Richardson PM, Solovyev VV, Rubin EM, Rokhsar DS, Banfield JF (2004) Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature 428:37–43

Urakawa H, El Fantroussi S, Smidt H, Smoot JC, Tribou EH, Kelly JJ, Noble PA, Stahl DA (2003) Optimization of single-base-pair mismatch discrimination in oligonucleotide microarrays. Appl Environ Microbiol 69:2848–2856

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

Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, Wu D, Paulsen I, Nelson KE, Nelson W, Fouts DE, Levy S, Knap AH, Lomas MW, Nealson K, White O, Peterson J, Hoffman J, Parsons R, Baden-Tillson H, Pfannkoch C, Rogers YH, Smith HO (2004) Environmental genome shotgun sequencing of the Sargasso Sea. Science 304:66–74

Volkov I, Banavar JR, Maritan A (2006) Comment on "Computational improvements reveal great bacterial diversity and high metal toxicity in soil". Science 313:918; author reply 918

Wagner M, Nielsen PH, Loy A, Nielsen JL, Daims H (2006) Linking microbial community structure with function: fluorescence in situ hybridization-microautoradiography and isotope arrays. Curr Opin Biotechnol 17:83–91

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

Wan Y, Broschat SL, Call DR (2007) Validation of mixed-genome microarrays as a method for genetic discrimination. Appl Environ Microbiol 73:1425–1432

Wilkes T, Laux H, Foy CA (2007) Microarray data quality – review of current developments. Omics 11:1–13

Wilson KH, Wilson WJ, Radosevich JL, DeSantis TZ, Viswanathan VS, Kuczmarski TA, Andersen GL (2002) High-density microarray of small-subunit ribosomal DNA probes. Appl Environ Microbiol 68:2535–2541

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

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

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

Yergeau E, Kang S, He Z, Zhou J, Kowalchuk GA (2007) Functional microarray analysis of nitrogen and carbon cycling genes across an Antarctic latitudinal transect. ISME J 1:163–179

Yilmaz LS, Noguera DR (2004) Mechanistic approach to the problem of hybridization efficiency in fluorescent in situ hybridization. Appl Environ Microbiol 70:7126–7139

Yilmaz LS, Okten HE, Noguera DR (2006) Making all parts of the 16S rRNA of Escherichia coli accessible in situ to single DNA oligonucleotides. Appl Environ Microbiol 72:733–744

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

Zhou X, Wu L, Zhou J (2004) Fabrication of DNA microarrays on nanoengineered polymeric ultrathin film prepared by self-assembly of polyelectrolyte multilayers. Langmuir 20:8877–8885

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