17
Microarray-Based Microbial Identification and Characterization

TERRY J. GENTRY AND JIZHONG ZHOU

Introduction

The development of molecular-based methodologies over the past two decades has dramatically improved our ability to detect microorganisms in clinical and environmental samples—enabling detection and identification within hours in many cases. However, most of these methods are only capable of monitoring individual or small groups of organisms at a time. Due to the extreme microbial diversity in many environments, such as the human intestine (Eckburg et al., 2005), it is necessary to monitor hundreds to thousands of different microbial populations simultaneously in order to detect all of the organisms of interest as a whole and understand these communities more comprehensively. Microarrays have the unprecedented potential to achieve this objective as specific, sensitive, quantitative, and high-throughput tools for microbial detection, identification, and characterization. Advances in printing technology have enabled the production of microarrays containing thousands to hundreds of thousands of probes. Although microarrays have been primarily developed and used for gene expression profiling of pure cultures of individual organisms, major advances have recently been made in their application to complex environmental samples. This chapter discusses the basis of different microarray formats and their application to issues of clinical interest. Several reviews on microarray technology have recently been published and may provide additional information of interest (Ye et al., 2001; Zhou and Thompson, 2002; Cook and Sayler, 2003; Zhou, 2003; Bodrossy and Sessitsch, 2004; Schadt and Zhou, 2005; Schadt et al., 2005).

Principles and Types of Arrays

Conceptually, microarrays are an extension of traditional membrane-based Northern and Southern blots where a labeled probe molecule is hybridized to target DNA or RNA attached to a membrane. However, this process is reversed in microarray analysis with the probe attached to the support substrate, usually a nonporous solid surface, and the labeled DNA or RNA then hybridized to the probe. The major

TABLE 17.1. Selected properties of microarrays for microbial detection and characterization.

	Type of array			
Property	POA	FGA	CGA	MGA
Probe template	Ribosomal rRNA genes	Functional genes	Whole genome	Environmental DNA
Probe length	~18–25 nt	~ 50–70 nt oligos or ~200–1000 nt PCR products	Whole genome	~1000+ nt
Targeted microorganisms	Cultured & uncultured	Cultured & uncultured	Cultured	Cultured & uncultured
Information provided	Phylogenetic	Functional	Phylogenetic	Functional
Specificity	Species level or single nucleotide difference	< 80–90% sequence homology	Species – strain	≥Strain
Sensitivity (ng of pure genomic DNA)	~ 500[a]	~1–8	~0.2	Undetermined
Quantitative	Depends on array design[a]	Yes	Yes	Undetermined

POA, phylogenetic oligonucleotide array; FGA, functional gene array; CGA, community genome array; MGA, metagenomic array.
[a]Undetermined for POAs based on perfectly matched and mismatched probe pairs.
Adapted from Zhou (2003).

advantage of this approach is that the sample can be screened with thousands of probes simultaneously.

Arrays with potential application to diagnostic clinical research can be divided into at least four major categories based on what genes are represented on the array: (1) phylogenetic oligonucleotide arrays (POAs), which are designed based on a conserved marker such as the 16S rRNA gene and are used to detect specific organisms and compare the relatedness of microbial communities; (2) functional gene arrays (FGAs), which are designed for key functional genes involved in various physiological processes, such as antibiotic resistance, and provide information on the genes and microbial populations involved with these processes; (3) community genome arrays (CGAs), which contain the whole genomic DNA of cultured microorganisms and can describe an isolate or community based on its relationship to these cultivated organisms; and (4) metagenomic arrays (MGA), which contain probes produced directly from environmental DNA itself and can be a potentially powerful technique because, unlike the other arrays, they can applied with no prior sequence knowledge of the community (Table 17.1).

Phylogenetic Oligonucleotide Arrays.

Most POAs contain short oligonucleotide (oligo) probes representing genes with diagnostic value, usually the small-subunit ribosomal RNA (16S rRNA) gene.

These arrays are commonly used to detect the presence of specific bacteria in complex samples based on unique sequence regions. Several factors make 16S rRNA ideal for microbial identification and differentiation including (1) 16S rRNAs genes are found in all bacteria; (2) there is no evidence of horizontal transfer of these genes between organisms; and (3) the genes contain both conserved and variable regions, either of which can be used for probe selection depending on the objective of the study (Olsen et al., 1986). Additionally, there is a vast amount (>100,000 sequences) of rRNA sequence data available via the Ribosomal Database Project (RDP) (Cole et al., 2005).

Because some regions of rRNA genes are highly conserved, it is often necessary to use short oligos (~20-mers) for POAs in order for the probes to be specific to individual organisms. Using shorter probes, it is possible to discriminate a single mismatch in a probe-target hybridization (Zhou et al., 2004). A commonly used POA design strategy consists of arraying several probes that perfectly match a given target along with corresponding probes containing a single mismatch (usually at the central position) relative to the target (Wilson et al., 2002a; El Fantroussi et al., 2003; Peplies et al., 2003; Urakawa et al., 2003). Detection of the target sequence is indicated by greater signal intensity for the perfectly matched probes compared with the mismatched probes. Although this strategy enables very specific detection of target sequences, it does have some potential disadvantages, which we discuss in a later section on specificity.

One of the challenges for 16S rRNA-based analysis is the innate propensity of these molecules to form stable secondary structures that may interfere with hybridization and lead to false-negative results. For example, one study (Peplies et al., 2003) reported that 17 out of 41 expected hybridization events were not detected. This possibility can be reduced by incorporating one of numerous available software programs, such as Mfold (Zuker, 2003), which can identify self-complementarity in oligo probes into the probe design process. This difficulty can also be addressed by either fragmenting the target prior to hybridization or by the inclusion of helper-probes in the hybridization mixture. The helper-probes are designed to disrupt the local secondary structure by binding to the target molecule adjacent to the actual probe binding site. However, there is a risk that the helper-probe could cause nonspecific binding if it binds too closely to the actual probe binding site (Chandler et al., 2003). Furthermore, disruption of secondary structure in one region may result in the formation of secondary structures in other regions that could possibly affect the binding sites of other probes.

Functional Gene Arrays

In contrast with POAs, which are primarily used for the detection of specific microorganisms, FGAs target genes involved in some process of interest. FGAs can also be used to determine the expression of these genes by measuring mRNA, but only a limited number of studies have used FGAs for mRNA analysis (Dennis et al., 2003; Rhee et al., 2004) due to the technical challenges of mRNA isolation (SalehLakha et al., 2005). Thus, FGAs not only provide a degree of phylogenetic

classification but they also can give information on the genetic capacity for, or activity of, a given process in the specific environment being studied.

The initial and one of the most critical steps for FGA analysis is the selection of the genes to be targeted by the array. This will depend on the specific research objectives and characteristics of the sample to be analyzed. In contrast with POAs, there is a limited sequence data available for many functional genes. This is an important consideration when selecting a gene(s) for inclusion on a FGA. This may also necessitate the generation of clone libraries for the gene and environment under study in order to obtain the requisite sequence information for probe design. Characteristics of an ideal target gene for an FGA include (1) it encodes a critical enzyme or protein in the process of interest; (2) its sequence is evolutionarily conserved but with sufficient divergence in different microorganisms to allow the design of species-specific probes; and (3) it has a wide spectrum of published sequence data from isolated organisms and environmental samples.

Once the target genes are selected, it is necessary to decide what type of probes will be used on the array. The molecules most commonly used as probes are PCR amplification products (∼200–1000 bp) and shorter, synthesized oligos (∼20–70 nt). The PCR amplicon probes have the advantage of being amplifiable from their source organisms via conserved primers without the need for specific sequence knowledge. These probes tend to be more sensitive than the shorter oligo probes (He et al., 2005), but they also have higher potential for cross-hybridization. Furthermore, depending on the number of organisms or genes to be represented on the array, it can be virtually impossible to acquire all of the necessary isolates and clones from their various sources in order to produce a comprehensive PCR amplicon-based array. Perhaps the greatest advantage of the oligo probes is that they can be designed and synthesized directly from available sequence data. This also enables greater control and flexibility in the design process, such as the ability to avoid highly conserved regions of genes.

Several factors that can affect probe specificity and thus should be considered during probe design include (1) nucleotide similarity of probe with nontarget sequences; (2) long stretches of a probe that are complementary to a nontarget sequence, which can lead to substantial nonspecific hybridization (Kane et al., 2000; Hughes et al., 2001); (3) the position of mismatches—more specific binding occurs when the mismatches are those distributed across the probe instead of concentrated in a single location (Letowski et al., 2004); and (4) the amount of free energy of probe-target duplexes (Li and Stormo, 2001; Held et al., 2003; Taroncher-Oldenburg et al., 2003). Specificity is also affected by the hybridization conditions (temperature, formamide concentration, salt concentration, etc.).

Recent research has shown that specific probes for FGAs could be produced using more relaxed design criteria when multiple probe-target characteristics were simultaneously considered during the design process (Liebich et al., unpublished). This indicated that specific hybridization at 50°C with 50% formamide could be achieved using 50-mer probes with a free energy release of $\leq$–35 kcal/mol and $\leq$90% similarity and $\leq$20 bp continuous stretches to nontarget sequences. Relaxing

the design criteria, even slightly, should increase the percentage of target genes for which probes can be designed.

Several software programs are currently available for the design of FGA oligo probes including: ArrayOligoSelector (Bozdech et al., 2003); OligoArray (Rouillard et al., 2002); OligoArray 2.0 (Rouillard et al., 2003); Oligopicker (Wang and Seed, 2003); OligoWiz (Nielsen et al., 2003); PRIMEGENS (Xu et al., 2002); PROBEmer (Emrich et al., 2003); ProbeSelect (Li and Stormo, 2001); and ROSO (Reymond et al., 2004). Although these programs work well for designing probes from whole-genome sequences, recent research in J. Zhou's laboratory (Li et al., unpublished) has found that a large portion of the probes designed by some of these programs from orthologous gene sequences, such as those produced by clone libraries, were not specific to the target sequence. Therefore, a new probe design software program, called CommOligo, was designed to correct this problem (Li et al., unpublished). This program incorporates a new global alignment algorithm to identify unique probes for each gene using multiple, simultaneously evaluated criteria that can be defined by the user. One major advantage of CommOligo is that it can also design group-specific probes for sets of sequences that are too similar for the design of unique probes, thus increasing the number of sequences covered by probes. Until improved design software is available, researchers should exercise caution when using software that was originally designed for use with whole-genome data for the design of probes from environmental sequences.

In addition to the actual experimental probes, it may be beneficial to design and include control probes on the array that have varying similarity to a control sequence. The control DNA can then be spiked into the hybridization solution to ensure that the correct hybridization stringency is achieved.

Community Genome Arrays

Unlike the other types of arrays, entire genomes of isolated organisms are used as probes for CGAs. These arrays are conceptually equivalent to membrane-based reverse sample genome probing (RSGP) (Greene and Voordouw, 2003), but they use a nonporous hybridization surface and fluorescence-based detection, which allows for high-throughput analyses but reduces sensitivity (Wu et al., 2004). CGAs can achieve strain-level differentiation of isolates and thus can be used to ascertain the genomic similarity of isolated bacteria or microbial communities in relation to the organisms represented on the array. The primary disadvantage of CGAs is that only cultured organisms (probes) are included on the arrays.

Metagenomic Arrays

Instead of genomic DNA from cultured organisms, MGAs use DNAs directly cloned from an environment of interest as the probes. This approach has only been used on a limited scale (Sebat et al., 2003), but it has great potential for many applications because the vast amount of unknown sequences in many samples

is one of the primary limitations for microarray analysis. This approach could also be used to produce a site-specific FGA for measuring microbial activity if sufficient mRNA could be obtained and reverse-transcribed to cDNA for the array probes.

Applications

The high-throughput capacity of microarray technology has numerous applications in diagnostic microbiology, specifically in relation to rapid pathogen detection, identification, and characterization. Many of these methods also have utility for clinical research. The following are specific examples of recent and emerging microarray applications, but the technology can be used to investigate virtually any microorganism or microbial process.

Microbial Detection and Identification

One of the potentially most powerful clinical uses of microarrays is for the rapid, simultaneous assay and detection of thousands of microorganisms. Arrays have been designed and used to detect many pathogenic microorganisms including bacteria, viruses, fungi, and protozoa (Straub et al., 2002; Wilson et al., 2002b; Diaz and Fell, 2005; Korimbocus et al., 2005). Although PCR-amplicons or longer oligos (50–70-mers) have been successfully used for arrays of limited scope, it may be necessary to use shorter oligo probes (~20-mers) for more comprehensive arrays due to reasons previously discussed. One of the most comprehensive arrays published to date for microbial detection was a POA containing 31,179 hierarchical probes perfectly matching their targets (and a corresponding number of mismatched probes) representing 1945 prokaryotic and 431 eukaryotic sequences (Wilson et al., 2002a). The array could successfully identify 15 of the 17 tested pure bacterial cultures. The diagnostic ability of the POA was then tested using microorganisms collected from a 1.4-million-liter air sample. Although the results generally agreed with those from an rDNA clone library, the array could only resolve differences to the third level of phylogenetic rank, as defined by RDP, and could not identify individual species. This illustrates the challenge for developing comprehensive, yet highly specific arrays for use with complex samples. The approach above used universal primers to amplify a single region of the target DNA for hybridization. An alternative approach, which may improve detection specificity, is to use species-specific primers to amplify multiple diagnostic regions (e.g., pathogenicity and virulence genes) for each organism of interest and then hybridize the pooled products to an array containing tiled probes covering the entirety of each of the diagnostic regions (Wilson et al., 2002b). In addition to non–culture-based detection, microarrays can also be used to specifically genotype isolated organisms (or enriched sequences) to the strain level or even differentiate point mutations depending on the array format (Vinje and Koopmans, 2000; Straub et al., 2002; Willse et al., 2004; Wu et al., 2004; Lin et al., 2005).

Community Dynamics and Activity

Microarrays now provide the researcher with the unprecedented ability to detect and monitor most, if not all, of the populations even in complex communities such as the human intestine. Microorganisms inhabiting the intestines play instrumental roles in the maintenance of health and development of disease, yet only recently have they begun to be investigated on a large scale (Eckburg et al., 2005). Researchers have used microarrays (primarily POAs) to monitor dominant bacteria and those with certain phenotypes (e.g., production of carcinogens) in intestinal and fecal samples (Wang et al., 2002, 2004a, 2004b), but to our knowledge, the communities in these environments have yet to be analyzed with comprehensive arrays representing thousands of organisms. Similarly, microarrays have also been used to measure gene expression in the intestine, but these studies have largely used human gene–based arrays to determine the response to beneficial or pathogenic bacteria (Caro et al., 2005; Galindo et al., 2005). Microarrays, specifically FGAs, could likewise be used to determine the activity of specific microbial populations (or even communities) in the intestines of different hosts or in response to different diets or stimuli (Stintzi et al., 2005).

Antibiotic Resistance

FGAs can be used to detect virtually any gene of interest including those encoding pathogenicity and virulence factors such as antibiotic resistance (Korczak et al., 2005). In contrast with traditional antibiotic resistance assays, which require isolation and growth of the organisms of interest, microarrays have the potential to rapidly screen a sample for the presence of multiple antibiotic resistance genes without the need for culturing. However, direct detection without amplification may be more applicable to antibiotic efflux- or modification-based resistance mechanisms than to mutation-based resistance given the difficulty in differentiating single (or a few) nucleotide differences with the longer probes, which provide better sensitivity. Very short oligo ($\sim$20-mers) are more appropriate for detecting resistance due to mutations but will likely require amplification of these specific genes. Once the antibiotic resistant organisms are isolated, FGAs can be used to genotype the antibiotic resistance genes, potentially revealing information on the gene's development and/or acquisition (Troesch et al., 1999; Call et al., 2003; Grimm et al., 2004; Perreten et al., 2005). This has application not only to clinical diagnostics but also to any research or regulatory program concerned with the spread of antibiotic resistance genes and the development of multidrug-resistant bacteria.

Challenges for Microarray Analysis

Clinical and environmental samples present several challenges for microarray analysis that are not encountered during the analysis of pure cultures. Although several recent studies have used microarrays to interrogate these types of samples, many

analytical challenges remain with respect to sensitivity, specificity, quantitation, and data analysis.

Specificity

One of the major challenges for microarray analysis is the design of probes specific to a given target. This is largely due to the conserved nature of many genes and the large amount of unknown sequence data present in many samples. Although longer probes may increase sensitivity, they also increase the potential for cross-hybridization with nontarget sequences. By using oligo probes, it is possible to avoid conserved regions of genes or areas containing stable secondary structure during the probe design process. The shorter oligo probes (~20-mers) can differentiate a single mismatch in a probe-target hybridization, making them ideal for use with highly conserved genes such as 16S rRNA in POAs (Wilson et al., 2002a; Urakawa et al., 2003; Zhou et al., 2004). A common format for these arrays includes sets of probes that perfectly match a target sequence and corresponding sets of probes containing one mismatched nucleotide, usually at a central position. Greater signal intensity for the perfect probes versus the mismatched probes indicates detection of the target sequence. Even though the mismatched probes typically have greatly decreased ability to bind the target of interest (Zhou et al., 2004), spurious results are sometimes obtained. This is likely due to the presence of similar yet unknown sequences and can make it difficult to achieve complete discrimination. One way to address this problem is to design and use multiple perfectly matched and mismatched probe combinations for each organism or gene of interest. The results from the probe pairs are then compared statistically, and those with abnormal results (higher signal intensity for the mismatched probe) are discarded during data analysis. It may also be possible to improve the differentiation of perfectly matched and mismatched probes by determining the thermal dissociation curve for each probe–target hybridization on a three-dimensional array platform (Liu et al., 2001; El Fantroussi et al., 2003; Urakawa et al., 2003), but this may be difficult for high-density planar arrays given the current technology.

Most functional genes are more variable than rRNA genes thus enabling the use of longer oligo probes (~40- to 70-mers), which have greater sensitivity, while still achieving species-level specificity. These longer oligo-based probes have been reported to discriminate sequences less than 80–90% similar to the probes (Taroncher-Oldenburg et al., 2003; Rhee et al., 2004). Specificity can also be increased or decreased, to an extent, by adjusting the stringency of the hybridization conditions (temperature, formamide concentration, salt concentration, etc.) (Wu et al., 2004). However, caution should be exercised when using an array under more or less stringent conditions than that for which it was designed, as this could cause overestimated or underestimated results and ultimately inaccurate conclusions.

Sensitivity

The different array types have not been directly compared with regard to sensitivity, but limits of 0.2 ng of target genomic DNA for a CGA (Wu et al., 2004), 1 ng for a

PCR-based FGA (Wu et al., 2001), and 5–8 ng for 50-mer oligo FGAs (Rhee et al., 2004; Tiquia et al., 2004) in the absence of background DNA have been reported. However, when background DNA is added to simulate environmental samples, these sensitivities are decreased around 10-fold (Rhee et al., 2004; Tiquia et al., 2004; Wu et al., 2004). The relative sensitivity of the array is correlated with probe length with shorter oligo probes typically being ~10- to 100-fold less sensitive than longer PCR-based or CGA probes (Wu et al., 2001, 2004; Denef et al., 2003; Rhee et al., 2004; He et al., 2005). Depending on the specific research objective, it may be desirable to use probes that are as long as reasonably possible without compromising specificity.

The most common approach used to increase sensitivity and detect less dominant populations is to PCR-amplify these specific organisms or groups. However, this potentially introduces other well-documented biases and limitations, making it preferable to avoid amplification if possible (Reysenbach et al., 1992; Farrelly et al., 1995; Crosby and Criddle, 2003). Magnetic beads or other capture techniques may also be useful for selecting specific populations (Tsai et al., 2003).

Although some populations or genes can be amplified with specific PCR primers as mentioned above, this option is not available for all genes and may not be the optimal approach if hundreds to thousands of genes are being simultaneously considered. In these situations, a nonspecific amplification approach is needed to amplify the community DNA or RNA. A whole community genome amplification (WCGA) procedure based on rolling circle amplification has recently been developed for use with microarray analysis (Wu et al., 2005). The method representatively detected individual genes or genomes starting from 1 to 100 ng DNA of individual or mixed genomes of equal or unequal abundance, and 1 to 500 ng of environmental DNAs. It could detect initial target DNA concentrations as low as 10 fg, but the representativeness of amplification was affected by the lower template concentrations. Hybridization of amplification products to several types of arrays indicated significantly linear relationships between initial DNA concentration and signal intensity across a range of DNA concentrations from pure cultures ($r^2 = 0.65$ to 0.99) and environmental samples ($r^2 = 0.96$ to 0.98). Other researchers are developing methods for amplification of prokaryotic mRNA including one approach that adds a poly(A) tail to the mRNA for subsequent amplification (Botero et al., 2005).

Researchers have used different nucleic acid labeling methods to increase sensitivity (Denef et al., 2003; Steward et al., 2004; Zhou and Zhou, 2004). For example, one study (Denef et al., 2003) used tyramide signal amplification labeling to increase the signal intensity of a 70-mer FGA ~10-fold, compared with the commonly used Cy dye-labeling techniques, which ultimately lowered the detection limit to ~1% of cells in the total community. Planar glass slides are commonly used for microarray analysis because they enable higher printing densities than the three-dimensional arrays. However, hybridization on these nonporous surfaces are several orders of magnitude less sensitive than membrane-based hybridizations due to the limited amount of probe material that can be attached to the nonorous surfaces (Cho and Tiedje, 2002). The development and use of new slide chemistries,

including ultrathin three-dimensional platforms, which enable increased binding capacities with high-density arrays, may also help to increase sensitivity (Guschin et al., 1997; Urakawa et al., 2003; Zhou et al., 2004).

Quantitation and Data Analysis

Due to the potential variability in steps including DNA extraction, labeling, hybridization, and analysis, there has been some debate whether microarray analysis is quantitative. Recent research has indicated that some array formats, including FGAs and CGAs, can be quantitative over a range of concentrations (Wu et al., 2001; Rhee et al., 2004; Wu et al., 2004). However, it is not currently known if arrays based on perfect match–mismatch probes sets (e.g., POAs) are quantitative.

Due to the same variations mentioned above, it can be difficult to compare data between, and even within, microarray experiments. Techniques such as the two-color dye-swap method works well for measuring relative levels of gene expression in pure cultures, but these may not be directly applicable to the types of analyses needed for many clinical and environmental samples. Different methods are needed to standardize data between slides and experiments. Alternative approaches have been developed where known amounts of labeled DNA fragments or oligos are spiked into the hybridization solution as a control (Cho and Tiedje, 2002; Dudley et al., 2002; Bodrossy et al., 2003). The array results are then normalized based on the hybridization signal intensity of this control DNA and corresponding control probes on the array.

However, it could be difficult to quantitatively correlate differences in hybridization signals with changes in specific microorganisms or genes due to the large amount of unknown nucleic acid sequences in many clinical and environmental samples—even if the microarray probes and experiments have been carefully designed and performed. Although it is typically assumed that the abundance of the target organism is directly proportional to the observed microarray hybridization signal intensity, nonspecific hybridization to uncharacterized microorganisms in the samples could occur and complicate interpretation. It may be beneficial to analyze key genes in selected samples with other methods, such as real-time PCR, to validate the conclusions drawn from microarray data (Rhee et al., 2004).

Conclusion

Microarrays have the unprecedented potential to simultaneously detect and monitor thousands of genes or organisms. Although microarrays were primarily developed for measuring gene expression in pure cultures, considerable effort has been expended over the past few years to adapt the technology for other applications. Several recent experiments have applied microarray technology to issues relevant to clinical microbiology including the rapid detection and identification of microorganisms and the genotyping of antibiotic resistance genes. The integration of microarray technology into clinical diagnostics is still in the early stages.

Although methodological and technological improvements are needed to broaden the applicability of the technology, future advances will undoubtedly expand the use of microarray technology in the clinical laboratory.

Acknowledgments. The authors' efforts in preparing this review were supported by the U.S. Department of Energy Office of Science as part of its Environmental Research Programs in Natural and Accelerated Bioremediation Research, Biotechnology Investigations Ocean Margins Program, Microbial Genome Program, Genomics:GTL program through the Virtual Institute of Microbial Stress and Survival (VIMSS; http://vimss.lbl.gov), and Carbon Sequestration (as part of the consortium on research to enhance Carbon Sequestration in Terrestrial Ecosystems [CSiTE]). Oak Ridge National Laboratory is managed by UT-Battelle LLC for the Department of Energy under contract DE-AC05-00OR22725.

References

Bodrossy, L., & Sessitsch, A. (2004). Oligonucleotide microarrays in microbial diagnostics. *Curr Opin Microbiol, 7*, 245–254.

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

Botero, L. M., D'Imperio, S., Burr, M., McDermott, T. R., Young, M., & Hassett, D. J. (2005). Poly(A) polymerase modification and reverse transcriptase PCR amplification of environmental RNA. *Appl Environ Microbiol, 71*, 1267–1275.

Bozdech, Z., Zhu, J., Joachimiak, M. P., Cohen, F. E., Pulliam, B., & DeRisi, J. L. (2003). Expression profiling of the schizont and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray. *Genome Biol, 4*, R9.

Call, D. R., Bakko, M. K., Krug, M. J., & Roberts, M. C. (2003). Identifying antimicrobial resistance genes with DNA microarrays. *Antimicrob Agents Chemother, 47*, 3290–3295.

Caro, S. D., Tao, H., Grillo, A., Franceschi, F., Elia, C., Zocco, M. A., Gasbarrini, G., Sepulveda, A. R., & Gasbarrini, A. (2005). *Bacillus clausii* effect on gene expression pattern in small bowel mucosa using DNA microarray analysis. *Eur J Gastroenterol Hepatol, 17*, 951–960.

Chandler, D. P., Newton, G. J., Small, J. A., & Daly, D. S. (2003). Sequence versus structure for the direct detection of 16S rRNA on planar oligonucleotide microarrays. *Appl Environ Microbiol, 69*, 2950–2958.

Cho, J. C., & Tiedje, J. M. (2002). Quantitative detection of microbial genes by using DNA microarrays. *Appl Environ Microbiol, 68*, 1425–1430.

Cole, J. R., Chai, B., Farris, R. J., Wang, Q., Kulam, S. A., McGarrell, D. M., Garrity, G. M., & Tiedje, J. M. (2005). The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. *Nucleic Acids Res, 33*(Database issue), D294–296.

Cook, K. L., & Sayler, G. S. (2003). Environmental application of array technology: promise, problems and practicalities. *Curr Opin Biotechnol, 14*, 311–318.

Crosby, L. D., & Criddle, C. S. (2003). Understanding bias in microbial community analysis techniques due to *rrn* operon copy number heterogeneity. *Biotechniques, 34*, 790–798.

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

Diaz, M. R., & Fell, J. W. (2005). Use of a suspension array for rapid identification of the varieties and genotypes of the *Cryptococcus neoformans* species complex. *J Clin Microbiol, 43*, 3662–3672.

Dudley, A. M., Aach, J., Steffen, M. A., & Church, G. M. (2002). Measuring absolute expression with microarrays with a calibrated reference sample and an extended signal intensity range. *Proc Natl Acad Sci U S A, 99*, 7554–7559.

Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S. R., Nelson, K. E., & Relman, D. A. (2005). Diversity of the human intestinal microbial flora. *Science, 308*, 1635–1638.

El Fantroussi, S., Urakawa, H., Bernhard, A. E., Kelly, J. J., Noble, P. A., Smidt, H., Yershov, G. M., & Stahl, D. A. (2003). Direct profiling of environmental microbial populations by thermal dissociation analysis of native rRNAs hybridized to oligonucleotide microarrays. *Appl Environ Microbiol, 69*, 2377–2382.

Emrich, S. J., Lowe, M., & Delcher, A. L. (2003). PROBEmer: A web-based software tool for selecting optimal DNA oligos. *Nucleic Acids Res, 31*, 3746–3750.

Farrelly, V., Rainey, F. A., & Stackebrandt, E. (1995). Effect of genome size and *rrn* gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. *Appl Environ Microbiol, 61*, 2798–2801.

Galindo, C. L., Fadl, A. A., Sha, J., Pillai, L., Gutierrez, C., Jr., & Chopra, A. K. (2005). Microarray and proteomics analyses of human intestinal epithelial cells treated with the *Aeromonas hydrophila* cytotoxic enterotoxin. *Infect Immun, 73*, 2628–2643.

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

Grimm, V., Ezaki, S., Susa, M., Knabbe, C., Schmid, R. D., & Bachmann, T. T. (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.

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

Held, G. A., Grinstein, G., & Tu, Y. (2003). Modeling of DNA microarray data by using physical properties of hybridization. *Proc Natl Acad Sci U S A, 100*, 7575–7580.

Hughes, T. R., Mao, M., Jones, A. R., Burchard, J., Marton, M. J., Shannon, K. W., Lefkowitz, S. M., Ziman, M., Schelter, J. M., Meyer, M. R., Kobayashi, S., Davis, C., Dai, H., He, Y. D., Stephaniants, S. B., Cavet, G., Walker, W. L., West, A., Coffey, E., Shoemaker, D. D., Stoughton, R., Blanchard, A. P., Friend, S. H., & Linsley, P. S. (2001). Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat Biotechnol, 19*, 342–347.

Kane, M. D., Jatkoe, T. A., Stumpf, C. R., Lu, J., Thomas, J. D., & Madore, S. J. (2000). Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Res, 28*, 4552–4557.

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.

Korimbocus, J., Scaramozzino, N., Lacroix, B., Crance, J. M., Garin, D., & Vernet, G. (2005). DNA probe array for the simultaneous identification of herpesviruses, enteroviruses, and flaviviruses. *J Clin Microbiol, 43*, 3779–3787.

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 Methods, 57*, 269–278.

Li, F., & Stormo, G. D. (2001). Selection of optimal DNA oligos for gene expression arrays. *Bioinformatics, 17*, 1067–1076.

Lin, Z. H., Shen, X. H., Jin, Z., Kim, Y., Lee, E., Kim, H., & Kim, I. (2005). Human papillomavirus genotyping by oligonucleotide microarray and p16 expression in uterine cervical intraepithelial neoplasm and in invasive carcinoma in Korean women. *Pathol Int, 55*, 491–496.

Liu, W. T., Mirzabekov, A. D., & Stahl, D. A. (2001). Optimization of an oligonucleotide microchip for microbial identification studies: a non-equilibrium dissociation approach. *Environ Microbiol, 3*, 619–629.

Nielsen, H. B., Wernersson, R., & Knudsen, S. (2003). Design of oligonucleotides for microarrays and perspectives for design of multi-transcriptome arrays. *Nucleic Acids Res, 31*, 3491–3496.

Olsen, G. J., Lane, D. J., Giovannoni, S. J., Pace, N. R., & Stahl, D. A. (1986). Microbial ecology and evolution: a ribosomal RNA approach. *Annu Rev Microbiol, 40*, 337–365.

Peplies, J., Glockner, F. O., & 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.

Reymond, N., Charles, H., Duret, L., Calevro, F., Beslon, G., & Fayard, J. M. (2004). ROSO: optimizing oligonucleotide probes for microarrays. *Bioinformatics, 20*, 271–273.

Reysenbach, A. L., Giver, L. J., Wickham, G. S., & Pace, N. R. (1992). Differential amplification of rRNA genes by polymerase chain reaction. *Appl Environ Microbiol, 58*, 3417–3418.

Rhee, S. K., Liu, X., Wu, L., Chong, S. C., Wan, X., & Zhou, J. (2004). Detection of genes involved in biodegradation and biotransformation in microbial communities by using 50-mer oligonucleotide microarrays. *Appl Environ Microbiology, 70*, 4303–4317.

Rouillard, J. M., Herbert, C. J., & Zuker, M. (2002). OligoArray: genome-scale oligonucleotide design for microarrays. *Bioinformatics, 18*, 486–487.

Rouillard, J. M., Zuker, M., & Gulari, E. (2003). OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. *Nucleic Acids Res, 31*, 3057–3062.

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

Schadt, C. W., & Zhou, J. (2005). Advances in Microarrays for Soil Microbial Community Analyses. In: Nannipieri, P., & Smalla, K., eds. *Soil Biology: Nucleic Acids and Proteins in Soil*. Springer-Verlag, New York.

Schadt, C. W., Liebich, J., Chong, S. C., Gentry, T. J., He, Z., Pan, H., & Zhou, J. (2005). Design and use of functional gene microarrays (FGAs) for the characterization of microbial communities. In: Savidge, T., & Pothulakis, H., eds. *Microbial Imaging*, Vol. 34. Academic Press, London, pp. 331–368.

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

Steward, G. F., Jenkins, B. D., Ward, B. B., & Zehr, J. P. (2004). Development and testing of a DNA macroarray to assess nitrogenase (*nifH*) gene diversity. *Appl Environ Microbiol, 70*, 1455–1465.

Stintzi, A., Marlow, D., Palyada, K., Naikare, H., Panciera, R., Whitworth, L., & Clarke, C. (2005). Use of genome-wide expression profiling and mutagenesis to study the intestinal lifestyle of *Campylobacter jejuni*. *Infect Immun, 73*, 1797–1810.

Straub, T. M., Daly, D. S., Wunshel, S., Rochelle, P. A., DeLeon, R., & Chandler, D. P. (2002). Genotyping *Cryptosporidium parvum* with an *hsp70* single-nucleotide polymorphism microarray. *Appl Environ Microbiol, 68*, 1817–1826.

Taroncher-Oldenburg, G., Griner, E. M., Francis, C. A., & 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.

Tiquia, S. M., Wu, L., Chong, S. C., 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.

Troesch, A., Nguyen, H., Miyada, C. G., Desvarenne, S., Gingeras, T. R., Kaplan, P. M., Cros, P., & Mabilat, C. (1999). *Mycobacterium* species identification and rifampin resistance testing with high-density DNA probe arrays. *J Clin Microbiol, 37*, 49–55.

Tsai, Y. L., Le, J. Y., & Olson, B. H. (2003). Magnetic bead hybridization to detect enterotoxigenic *Escherichia coli* strains associated with cattle in environmental water sources. *Can J Microbiol, 49*, 391–398.

Urakawa, H., El Fantroussi, S., Smidt, H., Smoot, J. C., Tribou, E. H., Kelly, J. J., Noble, P. A., & Stahl, D. A. (2003). Optimization of single-base-pair mismatch discrimination in oligonucleotide microarrays. *Appl Environ Microbiol, 69*, 2848–2856.

Vinje, J., & Koopmans, M. P. (2000). Simultaneous detection and genotyping of "Norwalk-like viruses" by oligonucleotide array in a reverse line blot hybridization format. *J Clin Microbiol, 38*, 2595–2601.

Wang, X., & Seed, B. (2003). Selection of oligonucleotide probes for protein coding sequences. *Bioinformatics, 19*, 796–802.

Wang, R. F., Beggs, M. L., Robertson, L. H., & Cerniglia, C. E. (2002). Design and evaluation of oligonucleotide-microarray method for the detection of human intestinal bacteria in fecal samples. *FEMS Microbiol Lett, 213*, 175–182.

Wang, R. F., Beggs, M. L., Erickson, B. D., & Cerniglia, C. E. (2004a). DNA microarray analysis of predominant human intestinal bacteria in fecal samples. *Mol Cell Probes, 18*, 223–234.

Wang, R. F., Chen, H., Paine, D. D., & Cerniglia, C. E. (2004b). Microarray method to monitor 40 intestinal bacterial species in the study of azo dye reduction. *Biosensors Bioelectron, 20*, 699–705.

Willse, A., Straub, T. M., Wunschel, S. C., Small, J. A., Call, D. R., Daly, D. S., & Chandler, D. P. (2004). Quantitative oligonucleotide microarray fingerprinting of *Salmonella enterica* isolates. *Nucleic Acids Res, 32*, 1848–1856.

Wilson, K. H., Wilson, W. J., Radosevich, J. L., DeSantis, T. Z., Viswanathan, V. S., Kuczmarski, T. A., & Andersen, G. L. (2002a). High-density microarray of small-subunit ribosomal DNA probes. *Appl Environ Microbiol, 68*, 2535–2541.

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

Wu, L. Y., Thompson, D. K., Li, G. S., Hurt, R. A., Tiedje, J. M., & Zhou, J. Z. (2001). Development and evaluation of functional gene arrays for detection of selected genes in the environment. *Appl Environ Microbiol, 67*, 5780–5790.

Wu, L. Y., Thompson, D. K., Liu, X. D., Fields, M. W., Bagwell, C. E., Tiedje, J. M., & Zhou, J. Z. (2004). Development and evaluation of microarray-based whole-genome hybridization for detection of microorganisms within the context of environmental applications. *Environ Sci Technol, 38*, 6775–6782.

Wu, L., Liu, X., Schadt, C. W., & Zhou, J. (2006). Microarray-based analysis of subnanogram quantities of microbial community DNA using Whole Community Genome Amplification (WCGA). *Proc Natl Acad Sci U S A*, (in press).

Xu, D., Li, G., Wu, L., Zhou, J., & Xu, Y. (2002). PRIMEGENS: robust and efficient design of gene-specific probes for microarray analysis. *Bioinformatics, 18*, 1432–1437.

Ye, R. W., Wang, T., Bedzyk, L., & Croker, K. M. (2001). Applications of DNA microarrays in microbial systems. *J Microbiol Methods, 47*, 257–272.

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

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

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

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.

Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res, 31*, 3406–3415.