Microarrays for bacterial detection and microbial community analysis

Jizhong Zhou

Several types of microarrays have recently been developed and evaluated for bacterial detection and microbial community analysis. These studies demonstrated that specific, sensitive and quantitative detection could be obtained with both functional gene arrays and community genome arrays. Although single-base mismatch can be differentiated with phylogenetic oligonucleotide arrays, reliable specific detection at the single-base level is still problematic. Microarray-based hybridization approaches are also useful for defining genome diversity and bacterial relatedness. However, more rigorous and systematic assessment and development are needed to realize the full potential of microarrays for microbial detection and community analysis.

Addresses
Environmental Sciences Division, Oak Ridge National Laboratory, PO Box 2008, Oak Ridge, TN 37831-6038, USA
e-mail: zhouj@ornl.gov

Current Opinion in Microbiology 2003, **6**:288–294

This review comes from a themed issue on
Techniques
Edited by Jo Handelsman and Kornelia Smalla

1369-5274/03/$ – see front matter
Published by Elsevier Science Ltd.

DOI 10.1016/S1369-5274(03)00052-3

Abbreviations

CGA	community genome array
FGA	functional gene array
POA	phylogenetic oligonucleotide array
rRNA	ribosomal RNA
SSU	small subunit

Introduction

Microarrays (or microchips) are a recently developed, powerful genomic technology that are widely used to monitor gene expression under different cell growth conditions, detect specific mutations in DNA sequences and characterize microorganisms in environmental samples. Similar to the situation in which microprocessors have increased the speed of computation, microarray-based genomic technologies have revolutionized genetic analyses of biological systems. The widespread, routine use of such genomic technologies will shed light on a wide range of important research questions from how cells grow, differentiate and evolve, to the medical challenges of pathogenesis, antibiotic resistance and cancer; from agricultural issues such as seed breeding and pesticide resistance, to the biotechnological challenges of drug discovery and the remediation of environmental contamination.

Although microarray technology has been used successfully to analyze global gene expression in pure cultures [1–8], adapting microarray hybridization for use in environmental studies presents great challenges in terms of specificity, sensitivity and quantitation [9••]. Recently, various formats of microarrays have been developed and evaluated for bacterial detection and microbial community analyses in complex environments. These studies indicated that microarray-based genomic technologies have great potential as specific, sensitive, quantitative, parallel high-throughput tools for microbial detection, identification and characterization in natural environments. The characteristics of different types of microarrays are listed in Table 1.

In this review, I discuss recent work on the development of microarray-based technology for bacterial detection and microbial community analysis, with an emphasis on environmental application.

Functional gene arrays

The genes encoding functional enzymes involved in biogeochemical cycling processes (e.g. of carbon, nitrogen, sulfate and metals) are very useful signatures for monitoring the physiological status and functional activities of microbial populations and communities in natural environments. Microarrays containing functional gene sequence information are referred to as functional gene arrays (FGAs), as they are primarily used for the functional analysis of microbial community activity in the environment [10••]. Similar to the microarrays used for monitoring gene expression, both oligonucleotides and DNA fragments derived from functional genes can be used for fabricating FGAs (Table 1).

The performance of PCR-product-based FGAs for detecting microorganisms in environmental samples was systematically evaluated in terms of specificity, sensitivity and quantitation [10••]. A prototype FGA containing ~120 genes (~400–800 base pairs) involved in nitrogen cycling from pure cultures and those cloned from marine sediments were constructed and evaluated. Microarray hybridization results indicated that genes possessing <80–85% sequence identity could be differentiated under high-stringency hybridization conditions (65°C). The detection limit for *nirS* (nitrite reductase genes) was approximately 1 ng of pure genomic DNA and 25 ng of soil community DNA using the optimized

Table 1

Major differences of various types of microarrays for environmental studies.

	CGAs	PCR-product-based FGAs	Oligonucleotides-based FGAs	POAs
Probe size	Entire genomic DNAs	Individual functional genes (typically 200–1000 bp)	Individual functional genes (typically 50–70 bp)	Ribosomal rRNA genes, (typically 18–25 bp)
Types of Information provided	Phylogenetic	Functional	Functional	Phylogenetic
Construction of comprehensive arrays	More difficult	More difficult	Easier	Intermediate
Reagent handling and tracking	Intermediate	More difficult	Easier	Easier
Targeted microorganisms	Culturable	Culturable and non-culturable	Culturable and non-culturable	Culturable and non-culturable
Specificity	Species	< 80–85% sequence homology	< 86–90% sequence homology	Single nucleotide difference
Sensitivity (ng of pure genomic DNA)	~ 0.2	~ one	~ eight	Undetermined
Quantitation	Yes	Yes	Yes	Unknown
Taxonomic resolution	Genus–species	Genus–species	Species–strains	Species–strains

protocol. This level of sensitivity should be sufficient for many studies in microbial ecology. In addition, there was a strong linear quantitative relationship ($r^2 = 0.89–0.94$) between signal intensity and target DNA concentration in the range of 1 to 100 ng for genomic DNA from pure cultures and mixed communities. DNA microarrays have also been used to measure differences in DNA copy number in breast tumours [11,12]. Single-copy deletions or additions can be detected [12], suggesting that microarray-based detection is potentially quantitative. A recent study in which λDNA was co-spotted with DNA from reference bacterial strains also indicated that microarrays could accurately quantify genes in DNA samples [13].

As sequence divergence strongly affects hybridization signal intensity, however, the quantitative capacity of microarrays for measuring the relative abundance of targeted genes in complex environmental samples is less clear. The most obvious problem is how to interpret appropriately hybridization results from environmental samples of unknown diversity. In addition, when constructing microarrays containing large DNA fragments, the probes used for microarray fabrication are generally amplified by the polymerase chain reaction (PCR) from environmental clones or from pure genomic DNAs. Obtaining this diversity of environmental clones and bacterial strains from various sources as templates for amplification is a big challenge. As a result, construction of comprehensive FGAs representing diverse environmental sequences is difficult (Table 1).

To circumvent this problem, another type of FGA containing 50-base-pair probes has been developed and evaluated in my laboratory. The main advantage of the 50-mer FGA is that the construction is much easier than that of PCR-product-based FGAs because the probes can be directly designed and synthesized on the basis of sequence information from public databases. Therefore, a comprehensive array representing diverse environmental sequences can be easily constructed. To determine the potential performance of this type of microarray for environmental studies, a 50-mer FGA was constructed and evaluated using 1033 genes involved in nitrogen cycling (*nir*S, *nir*K, *nif*H, *amo*A and *pmo*A) and sulfite reduction (*dsr*A and *dsr*B) from public databases and our own sequence collections (S Tiquia *et al.*, unpublished data). Genes with less than 86–90% sequence identity were clearly differentiated using hybridization conditions of 50°C and 50% formamide. As expected, the hybridization specificity of the 50-mer FGA is higher than that of the PCR-product-based FGA. By comparing the probe sequences from pure cultures involved in nitrification, denitrification, nitrogen fixation, methane oxidation and sulfate reduction, the average similarity of these functional genes at the species level ranged from 74 to 84%. These results suggest that the 50-mer FGA could provide species-level resolution for analyzing microorganisms involved in these biogeochemical processes. The detection limit was approximately 8 ng of pure genomic DNA. As expected, the sensitivity of the 50-mer FGA is ten times lower than the PCR-product-based FGA and 100 times lower than community genome arrays (CGAs) [9••,10••]. In addition, similar to the PCR-product-based FGAs, a strong linear relationship was observed between signal intensity and target DNA concentrations in the 8–1000 ng range for all six functional gene groups ($r^2 = 0.96–0.98$).

Finally, using 2 μg of bulk community DNA from marine sediments, very strong hybridizations were obtained with the 50-mer FGA. These results suggest that the amount of DNA sample should not be a major limiting factor when using this type of microarray for environmental studies, as the average DNA yields from many surface soil and sediment samples usually ranged between 10 and 400 μg of DNA per gram of dry weight [14,15]. The 50-mer FGA therefore has potential for use as a specific, sensitive and potentially quantitative parallel tool for characterizing the composition, structure, activity and dynamics of microbial communities in natural environments.

On the basis of these results, a much more comprehensive 50-mer FGA of more than 6000 probes is being designed and constructed in my laboratory. The probes represent very diverse groups of functional genes involved in nitrogen and carbon cycling, phosphorus utilization, organic contaminant degradation and metal resistance.

Community genome arrays

Many microorganisms have been isolated from a variety of natural habitats. However, little or nothing is known about the genomic sequences for the majority of these microorganisms. Such a large collection of pure cultures would be very useful for monitoring microbial community composition, structure and dynamics in natural environments if microarrays could be developed that did not require prior knowledge of gene sequences. Thus, a novel prototype microarray containing whole genomic DNA, CGA, was developed and evaluated in my laboratory.

To examine hybridization specificity under varying experimental conditions and to determine threshold levels of genomic differentiation, a prototype microarray was fabricated that contained genomic DNA isolated from 67 representative microorganisms from the environment, classified as α-, β- and γ-proteobacteria and Gram-positive bacteria. Many of the selected species are closely related to each other on the basis of small subunit ribosomal RNA (SSU rRNA) and *gyr*B gene phylogenies and primarily belong to three major bacterial genera, *Pseudomonas*, *Shewanella* and *Azoarcus*. Microarray hybridizations at 55°C with 50% formamide revealed species-level differences, and strain-level differentiation was obtained for hybridizations performed at 65°C and 75°C (L Wu *et al.*, unpublished data). The detection limit, with labeled pure genomic DNA, was estimated to be approximately 0.2 ng (Table 1). Strong linear relationships were observed between hybridization signal intensity and target DNA concentrations for pure cultures, a mixture of DNA templates and a population of mixed cells ($r^2 = 0.95$–0.98). This suggests that CGA hybridization is potentially a quantitative tool for the detection and identification of microorganisms in environmental samples. However, when using CGAs for detecting bacterial populations in a mixed microbial community, stringent hybridization conditions (e.g. 55–65°C and 50% formamide) should be used to minimize any potential cross-hybridization among closely related species and strains.

CGA is conceptually analogous to membrane-based reverse sample genome probing (RSGP) [16], but CGA hybridization is distinctly different from RSGP in terms of the arraying substrate and signal detection strategies. In contrast to RSGP, CGA uses a non-porous surface for fabrication and fluorescence-based detection. The capability of accurate and precise miniaturization with robots on non-porous substrates is one of the two key advances of microarray-based genomic technologies. The miniaturized microarray format, coupled with fluorescent detection, represents a fundamental revolution in biological analysis. Like RSGP, the main disadvantage of the CGA is that only the cultured components of a community can be monitored because the construction requires the availability of individual pure isolates, although CGA-based hybridization itself does not require culturing [16]. With the recent advances in environmental genomics, high-molecular-weight DNA from uncultivated microorganisms could be accessed through bacterial artificial chromosomes (BACs). BAC clones could also be used to fabricate CGAs, thus allowing the investigation of uncultivated components of a complex microbial community.

CGAs are likely to be very useful for bacterial identification at the species/strain level. Significant linear relationships were observed between CGA hybridization ratios and sequence similarity values derived from SSU rRNA and *gyr*B genes, DNA–DNA reassociation, or REP- and BOX-PCR fingerprinting profiles ($r^2 = 0.80$–0.95) (L Wu *et al.*, unpublished data). Because of its high capacity, it is possible to construct CGAs containing bacterial type strains and also appropriately related strains. By hybridizing genomic DNA from unknown strains with this type of microarray, one should be able to identify quickly and reliably unknown strains, provided a suitably related probe is on the array. When using CGAs for strain identification, less stringent hybridization conditions (e.g. 45°C and 50% formamide) should be used first to ensure that good hybridization signals can be obtained for distantly related target species. If multiple probes have significant hybridization with the unknown target strains, highly stringent hybridization conditions should then be used.

Compared with the traditional DNA–DNA reassociation approach, CGAs have several advantages for determining species relatedness. As many bacterial genomes can be deposited on microarray slides, the tedious and laborious pair-wise hybridizations associated with the traditional DNA–DNA reassociation approach among different species is not needed with CGAs. In contrast to the traditional DNA–DNA reassociation approach, which generally requires about 100 μg DNA, CGA-based hybridization requires only about 2 μg of genomic DNA. This is important for determining the relationships between bacterial species that are recalcitrant to cultivation or grow very slowly.

Phylogenetic oligonucleotide arrays

rRNA genes are powerful molecules for studying phylogenetic relationships between different organisms, for analyzing microbial community structure in natural environments, and as useful targets for developing microarray-based detection approaches. Oligonucleotide microarrays containing information from rRNA genes are referred to as phylogenetic oligonucleotide microarrays (POAs),

because such microarrays are used primarily for phylogenetic analysis of microbial communities (Table 1). The oligonucleotide probes can be designed in a phylogenetic framework to survey different levels of sequence conservation, from highly conserved sequences giving broad taxonomic resolution to hypervariable sequences giving genus- and species-level groupings. Because highly conserved universal primers for amplifying rRNA genes are available, the SSU rRNA genes can be amplified before array hybridization, and hence POA-based hybridization can be easily coupled with PCR amplification. Therefore, highly sensitive assays can be implemented. However, POA-based hybridization presents some unique technical challenges [9••].

As rRNA genes are highly conserved and present in all microorganisms, specific detection with POA-based hybridization can be difficult [9••]. Although several previous studies showed that specific detection can be achieved on the basis of rRNA genes using both gel-pad oligonucleotide microarrays [17–20] and glass-based two-dimensional microarrays [21,22], detection specificity is still a central critical issue. Recently, Urakawa *et al.* [23•] evaluated the effects of the single-base-pair near-terminal and terminal mismatches on the melting temperature and microarray hybridization signal intensity using gel-pad oligonucleotide microarrays. Their results demonstrated that the position of mismatch and type of mismatch have significant effects on melting temperature and signal intensity. We have also evaluated whether single-base mismatch of SSU rRNA genes can be achieved by glass-slide-based microarrays. The signal intensity of the probes with single-base mismatch was three- to tenfold lower than those of the perfect match probes ([9••]; X Zhou and J Zhou, unpublished data). Maximum discrimination and signal intensity was achieved with 19-base probes. Similarly, our results indicated that single-base discrimination for SSU rRNA genes can be achieved with glass-slide-based array hybridization, but complete discrimination appears to be problematic with SSU rRNA genes [19,21,23•,24••]. Single-base differentiation was also obtained with the heat-shock protein Hsp70 in *Cryptosporidium parvum* [25].

SSU rRNA gene-based oligonucleotide arrays are still in the early stages of development, and therefore, only a few studies have applied POAs to the analysis of microbial community structure of environmental samples. Using photolithography-based Affymetrix technology, Wilson *et al.* [26••] designed a microarray containing 31,179 20-mer oligonucleotide probes specific for SSU rRNA genes. All of the probes were derived from a small SSU rRNA gene region (i.e. *Escherichia coli* positions 1409–1491), which were bound on both ends by universally conserved segments. The microarray also contained control sequences, which were paired with the probe sequences. A control sequence was identical to the paired probe sequence except that there was a mismatch nucleotide at the 11th position. Thus, the microarray contained a total of 62,358 features. The number of probes for individual sequences contained in the Ribosomal Database Project (RDP version 5.0, containing approximately 3200 sequences) ranged from 0 to 70. A total of 17 pure bacterial cultures were used to assess the performance of this microarray and 15 bacterial species were identified correctly. However, it failed to resolve the individual sequences comprising complex mixed samples [26••].

Loy *et al.* [24••] developed a glass-based microarray containing 132 SSU rRNA-targeted oligonucleotide probes, representing all recognized groups of sulfate-reducing prokaryotes. Microarray hybridizations with 41 reference strains showed that, under the hybridization conditions used, clear discrimination between perfectly matched and mismatched probes were obtained for most, but not all of the 132 probes. They used this microarray to determine the diversity of sulfate-reducing prokaryotes in periodontal tooth pockets and a hypersaline cyanobacterial mat. The microarray hybridization results were consistent with those obtained using well-established conventional molecular methods. These results suggest that microarray hybridization is a powerful tool for analysing community structure, but great caution is needed in data interpretation because of the potential for cross-hybridization.

Whole-genome open reading frame arrays for revealing genome differences and relatedness

Many microorganisms that are closely related on the basis of SSU rRNA gene sequences show dramatic differences in phenotypic characteristics. A way of understanding the genetic basis for such phenotypic differences is to obtain whole-genome sequence information for all closely related species of interest. Patterns of sequence similarity and variability will provide insights on the conservation of gene functions, physiological plasticity and evolutionary processes. However, sequencing the whole genomes of all closely related species is expensive and time-consuming. In addition, it might not be necessary to sequence all closely related genomes once the complete genome sequence for one representative microorganism is available, because substantial portions of the genomic sequence will be common among closely related species. A way of circumventing the need for sequencing multiple closely related genomes is to use DNA microarrays containing individual open reading frames (ORFs) of a sequenced microorganism to view genome diversity and relatedness of other closely related microorganisms.

The whole-genome ORF-array-based hybridization approach has been used to reveal genome diversity and relatedness among closely related organisms in several studies. Murray *et al.* [27••] used this approach to evaluate genome diversity and relatedness of several related

metal-reducing bacteria within the *Shewanella* genus using partial ORF microarrays for the sequenced metal-reducing bacterium *Shewanella oneidensis* MR-1. Both conserved and poorly conserved genes were identified among the nine species tested. Under the conditions used in this study, the hybridization results were most informative for the closely related organisms with SSU rRNA sequence similarities greater than 93% and *gyrB* sequence similarities greater than 80%. Above this level of homology, the similarities of microarray hybridization profiles were strongly correlated with *gyrB* sequence divergence. In addition, most genes in operons had high levels of DNA relatedness, suggesting that this approach can be used to identify genes or operons that were horizontally transferred [27••].

Using the ORF arrays for *E. coli* K-12, Dong *et al.* [28•] identified the genes in a common endophyte of maize, *Klebsiella pneumoniae* strain 342, which is closely related to *E. coli*. About 3000 (70%) *E. coli* genes were found in *K. pneumoniae* strain 342 with greater than 55% identity, whereas about 24% of the *E. coli* genes were absent in strain 342. Genes with high sequence identity were those involved in cell division, DNA replication, transcription, translation, transport, regulatory proteins, energy, amino acid and fatty acid metabolism, and cofactor synthesis, whereas the genes that are less conserved were those involved in carbon-compound metabolism, membrane proteins, structural proteins, central intermediary metabolism and proteins involved in adaptation and protection. Genes that were not identified in *K. pneumoniae* strain 342 included putative regulatory proteins, putative chaperones, surface structure proteins, mobility proteins, putative enzymes and hypothetical proteins. These genomic diversity results are consistent with the physiological properties of these two strains, suggesting that the microarray-based whole genome comparison is a powerful approach for revealing the genomic diversity and relatedness of closely related organisms.

The whole-genome ORF-array approach has also been used successfully to identify genome differences among 15 *Helicobacter pylori* strains with differences in virulence [29] and to detect deletions existing in other strains such as *Mycobacterium tuberculosis* and *M. bovis* [30]. All of these studies suggest that whole-genome ORF arrays will be useful for revealing genome difference and relatedness. Whole-genome ORF arrays are available for many microorganisms and they will be valuable for studying genome diversity and relatedness of closely related microorganisms. For example, whole genome arrays for six environmentally important microorganisms, including *Shewanella oneidensis* MR-1, *Deinococcus radiodurans* R1, *Rhodopseudomonas palustris*, *Nitrosomonas europaea*, *Desulfovibrio vulgaris* and *Geobacter metallireducens*, are available at Oak Ridge National Laboratory, TN, USA; and we are currently using these whole-genome ORF arrays to understand the genome diversity and relatedness of other important environmental isolates.

Other types of microarrays for bacterial detection and characterization

DNA microarrays containing random genomic fragments have been used to determine species relatedness where the genome sequence is not currently known. In this approach, 60–96 genomic fragments of about 1 kb were randomly selected from four fluorescent *Pseudomonas* species as reference genomes for microarray fabrication [31•]. Cluster analysis of hybridization profiles from 12 well-characterized fluorescent *Pseudomonas* species indicated that such types of microarray hybridization could provide species-to-strain level resolution. This approach is likely to have higher resolution than CGA because extensive component information is obtained rather than an average for the whole genome. However, this approach is more time-consuming and costly to develop than CGA and such an array would be more limited in scope as many of the array positions would be used for each reference microorganism (L Wu, personal communication).

Recently, instead of using a gel electrophoresis method, a random nanomer oligonucleotide microarray was developed and evaluated for obtaining fingerprinting profiles among closely related strains [32••]. A prototype array containing 47 randomly selected nanomer oligonucleotides was constructed and used to differentiate 14 closely related *Xanthomonas* strains. REP-PCR was first carried out to obtain the fingerprints from different strains, then the amplified REP-PCR products were hybridized with the nanomer array and fingerprinting profiles for each strain were obtained on the basis of microarray hybridization. The results showed that the microarray-based fingerprinting methods provided clear resolution among all strains examined, including *X. oryzae* 43836 and *X. oryzae* 49072, which could not be resolved using traditional gel electrophoresis of REP-PCR amplification methods. This suggests that the microarray hybridization approach could provide higher resolution in strain differentiation than the conventional gel electrophoresis fingerprinting approach. This approach is attractive because a universal nanomer array can be developed to generate fingerprints from any microorganisms.

Conclusions

The development and application of microarray-based genomic technology for bacterial detection and microbial community analysis has received a great deal of attention. Because of its high density and high throughput capacity, it is expected that microarray-based genomic technologies will revolutionize the analyses of microbial community structure, function and dynamics. Various microarray formats have been developed and evaluated for bacterial detection and microbial community analyses of environmental samples. The results

demonstrate that microarray-based genomic technologies have great potential as specific, sensitive and quantitative parallel high-throughput tools for microbial detection, identification and characterization in natural environments. However, more rigorous and systematic assessment and development are needed to realize the full potential of microarrays for microbial ecology studies. Several key issues will need to be addressed, including novel experimental designs and strategies for minimizing inherently high variations in hybridization to improve microarray-based quantitative accuracy, novel approaches for increasing hybridization sensitivity to detect extremely low biomass in natural environments, novel computational tools for microarray data extraction and interpretation and broad integration and application of microarray technologies with environmental studies to address ecological and environmental questions and hypotheses.

Acknowledgements

The author thanks Lynn Kszos for editorial assistance. The author's work described in this review was supported by the US Department of Energy Office of Science as part of its Biological and Environmental Research Programs in Natural and Accelerated Bioremediation Research, Biotechnology Investigations Ocean Margins Program, Microbial Genome Program 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. The submitted manuscript has been authored by a contractor of the US. Government under contract DE-AC05-00OR22725.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Lockhart DJ, Dong H, Byrne MC, Follettie MT, Gallo MV, Chee MS, Mittmann M, Wang C, Kobayashi M, Horton H *et al.*: **Expression monitoring by hybridization to high-density oligonucleotide arrays**. *Nat Biotechnol* 1996, **14**:1675-1680.

2. DeRisi JL, Iyer VR, Brown PO: **Exploring the metabolic and genetic control of gene expression on a genomic scale**. *Science* 1997, **278**:680-686.

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

4. Richmond CS, Glasner JD, Mau R, Jin H, Blattner FR: **Genome-wide expression profiling in *Escherichia coli* K-12**. *Nucleic Acids Res* 1999, **27**:3821-3835.

5. Ye RW, Tao W, Beddzyk L, Young T, Chen M, Li L: **Global gene expression profiles of *Bacillus subtilis* grown under anaerobic conditions**. *J Bacteriol* 2000, **182**:4458-4465.

6. Thompson DK, Beliaev AS, Giometti CS, Tollaksen SL, Khare T, Lies DP, Nealson KH, Lim H, Yates J III, Brandt CC *et al.*: **Transcriptional and proteomic analysis of a ferric uptake regulator (Fur) mutant of *Shewanella oneidensis*: possible involvement of Fur in energy metabolism, transcriptional regulation, and oxidative stress**. *Appl Environ Microbiol* 2002, **68**:881-892.

7. Liu Y, Zhou J, Omelchenko M, Beliaev A, Venkateswaran A, Stair J, Wu L, Thompson DK, Xu D, Rogozin IB *et al.*: **Transcriptome dynamics of *Deinococcus radiodurans* recovering from ionizing radiation**. *Proc Natl Acad Sci USA* 2003, **100**:4191-4196.

8. Wodicka L, Dong H, Mittmann M, Ho MH, Lockhart DJ: **Genome-wide expression monitoring in *Saccharomyces cerevisiae***. *Nat Biotechnol* 1997, **15**:1359-1367.

9. •• Zhou J, Thompson DK: **Challenges in applying microarrays to environmental studies**. *Curr Opin Biotechnol* 2002, **13**:204-207.
This commentary overviewed the application of microarray-based genomic technology for environmental studies, with an emphasis on the key microarray hybridization issues related to specificity, sensitivity and quantitation. This review also highlighted the challenges and future research directions of microarray technology for environmental applications.

10. •• Wu L, Thompson D, Li G, Hurt RA, Tiedje JM, Zhou J: **Development and evaluation of functional gene arrays for detection of selected genes in the environment**. *Appl Environ Microbiol* 2001, **67**:5780-5790.
This is the first report that systematically evaluated the performance of functional gene arrays for environmental applications in terms of specificity, sensitivity and quantitation.

11. Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y *et al.*: **High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays**. *Nat Genet* 1998, **20**:207-211.

12. Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO: **Genome-wide analysis of DNA copy-number changes using cDNA microarrays**. *Nat Genet* 1999, **23**:41-46.

13. Cho JC, Tiedje TM: **Quantitative detection of microbial genes by using DNA microarrays**. *Appl Environ Microbiol* 2002, **68**:1425-1430.

14. Zhou J, Bruns MA, Tiedje JM: **DNA recovery from soils of diverse composition**. *Appl Environ Microbiol* 1996, **62**:461-468.

15. Hurt RA, Qiu X, Wu L, Roh Y, Palumbo AV, Teidje JM, Zhou J: **Simultaneous extraction of RNA and DNA from environmental samples**. *Appl Environ Microbiol* 2001, **67**:4495-4503.

16. Voordouw G: **Reverse sample genome probing of microbial community dynamics**. *ASM News* 1998, **64**:627-633.

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

18. Guschin DG, Yershov A, Zaslavsky A, Gemmell V, Shick D, Proudnikov D, Arenkov P, Mirzabekov A: **Manual manufacturing of oligonucleotide, DNA and protein microchips**. *Anal Biochem* 1997, **250**:203-211.

19. Bavykin SG, Akowski JP, Zakhariev VM, Barsky VE, Perov AN, Mirzabekov AD: **Portable system for microbial sample preparation and oligonucleotide microarray analysis**. *Appl Environ Microbiol* 2001, **67**:922-928.

20. Koizumi Y, Kelly JJ, Nakagawa T, Urakawa H, El-Fantroussi S, Al-Muzaini S, Fukui M, Urushigawa Y, Stahl DA: **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* 2002, **68**:3215-3225.

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

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

23. • Urakawa H, Noble PA, El Fantroussi S, Kelly JJ, Stahl DA: **Single-base-pair discrimination of terminal mismatches by using oligonucleotide microarrays and neural network analyses**. *Appl Environ Microbiol* 2002, **68**:235-244.
This study evaluated the effects of the single-base-pair near-terminal and terminal mismatches on the melting temperature and microarray hybridization signal intensity using gel-pad oligonucleotide microarrays.

24. Loy A, Lahner A, Lee N, Adamczyk J, Meier H, Ernst J, Schleifer KH,
•• Wagner M: **Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment**. *Appl Environ Microbiol* 2002, **68**:5064-5081.
A glass-based microarray containing 132 SSU rRNA-targeted oligonucleotide probes, which represents all recognized groups of sulfate-reducing prokaryotes was developed and used to determine the diversity of sulfate-reducing prokaryotes in periodontal tooth pockets and a hypersaline cyanobacterial mat.

25. Straub TM, Daly DS, Wunshel S, Rochelle PA, DeLeon R, Chandler DP: **Genotyping *Cryptosporidium parvum* with an *hsp70* single-nucleotide polymorphism microarray**. *Appl Environ Microbiol* 2002, **68**:1817-1826.

26. Wilson KH, Wilson WJ, Radosevich JL, DeSantis TZ,
•• Viswanathan VS, Kuczmarski TA, Andersen GL: **High-density microarray of small-subunit ribosomal DNA probes**. *Appl Environ Microbiol* 2002, **68**:2535-2541.
This paper evaluated a microarray containing 31,179 20-mer oligonucleotide probes specific for SSU rRNA genes for bacterial detection in environmental samples. These are the most comprehensive phylogenetic oligonucleotide arrays that have been designed and evaluated to date.

27. Murray AE, Lies D, Li G, Nealson K, Zhou J, Tiedje JM:
•• **DNA–DNA hybridization to microarrays reveals gene-specific differences between closely related microbial genomes**. *Proc Natl Acad Sci USA* 2001, **98**:9853-9858.
This paper used whole-genome open reading frame (ORF) arrays to evaluate genome diversity and relatedness of several related metal-reducing bacteria within the *Shewanella* genus. The results also showed that this approach can be used to identify genes or operons that were horizontally transferred.

28. Dong Y, Glasner JD, Blattner FR, Triplett EW: **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* 2001, **67**:1911-1921.
Open reading frame (ORF) arrays for *Escherichia coli* K-12 were used to identify genes in the closely related bacterium *Klebsiella pneumoniae* strain 342. About 70% of *E. coli* genes were found in *K. pneumoniae* strain 342 with greater than 55% identity and 24% of the *E. coli* genes were absent in this strain. The study demonstrated that the ORF array is useful for understanding genome diversity of closely related bacteria.

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

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

31. Cho JC, Tiedje JM: **Bacterial species determination from
• DNA–DNA hybridization by using genome fragments and DNA microarrays**. *Appl Environ Microbiol* 2001, **67**:3677-3682.
A microarray containing randomly selected DNA fragments was evaluated for assessing bacterial species relatedness.

32. Kingsley MT, Straub TM, Call DR, Daly DS, Wunschel SC,
•• Chandler DP: **Fingerprinting closely related *Xanthomonas* pathovars with random nanomer oligonucleotide microarrays**. *Appl Environ Microbiol* 2002, **68**:6361-6370.
This paper evaluated glass-based microarrays containing randomly selected nanomer oligonucleotides for fingerprinting closely related bacterial strains.