Tiquia SM, Wu LY, Chong SC, Passovets S, Xu D, Xu Y et al.. Evaluation of 50-mer oligonucleotide arrays for detecting microbial populations in environmental samples. Biotechniques...

Article *in* BioTechniques · March 2004

Source: PubMed

Evaluation of 50-mer oligonucleotide arrays for detecting microbial populations in environmental samples

Sonia M. Tiquia, Liyou Wu, Song C. Chong, Sergei Passovets, Dong Xu, Ying Xu, and Jizhong Zhou

BioTechniques 36:664-675 (April 2004)

*Microarrays fabricated with oligonucleotides longer than 40 bp have been introduced for monitoring whole genome expression but have not been evaluated with environmental samples. To determine the potential of this type of microarray for environmental studies, a 50-mer oligonucleotide microarray was constructed using 763 genes involved in nitrogen cycling: nitrite reductase (*nirS *and* nirK*), ammonia monooxygenase (*amoA*), nitrogenase (*nifH*), methane monooxygenase (*pmoA*), and sulfite reductase (*dsrAB*) from public databases and our own sequence collections. The comparison of the sequences from pure cultures indicated that the developed microarrays could provide species-level resolution for analyzing microorganisms involved in nitrification, denitrification, nitrogen fixation, methane oxidation, and sulfite reduction. Sensitivity tests suggested that the 50-mer oligonucleotide arrays could detect dominant populations in the environments, although sensitivity still needs to be improved. A significant quantitative relationship was also obtained with a mixture of DNAs from eight different bacteria. These results suggest that the 50-mer oligonucleotide array can be used as a specific and quantitative parallel tool for the detection of microbial populations in environmental samples.*

INTRODUCTION

Our understanding of microbial community structure and dynamics in the environment continues to advance rapidly due to the development and application of molecular methods. Many novel uncultured microorganisms have been identified using conventional molecular methods (1,2). Microarray-based technology has the potential to overcome the limitations of traditional molecular methods for studying microbial community structure as high-throughput, cost-effective assessment tools for monitoring microbial communities (3).

In the past several years, microarray-based genomic technology has attracted tremendous interest among biologists because it permits the simultaneous analysis of thousands of genes in a single assay. DNA microarray technology has already been proven to be a particularly powerful tool to study whole genome expression (4–6). It is only recently that microarray-based genomic technology has been extended to the studies of microbial communities in the environment (3).

On the basis of the types of probes arrayed, microarrays used in environmental studies can be divided into three major classes (7); namely, functional gene arrays, community genome arrays, and phylogenetic oligonucleotide arrays. Functional gene arrays contain genes encoding key enzymes involved in various ecological processes such as carbon fixation, nitrification, denitrification, and sulfate reduction. Both DNA fragments and oligonucleotides derived from functional genes can be used for fabricating functional gene arrays. To avoid confusion, the former is referred to as DNA-based functional gene arrays, whereas the later is referred to as oligonucleotide-based functional gene arrays. These types of arrays are useful in studying physiological status and functional activities of microbial communities in natural environments (1,3). Community genome arrays are constructed using whole genomic DNA isolated from pure culture microorganisms and can be used to describe a microbial community in terms of its cultivable component (L. Wu, D.K. Thompson, X. Liu, C.E. Bagwell, M.W. Fields, J.M. Tiedje, and J. Zhou, manuscript submitted). Phylogenetic oligonucleotide arrays are constructed with short synthetic oligonucleotides from rRNA genes and can be used for phylogenetic analyses of microbial community composition and structure in environmental samples. Recently, oligonucleotide microarrays containing probes longer than 40 bp have been extensively evaluated and used for whole genome expression studies (8,9). These microarrays have high specificity and are easy to construct. Thus, they provide an important alternative array-based approach for monitoring gene expression. However, this type of microarray has not been tested and validated within the context of environmental samples. We therefore evaluated 50-mer oligonucleotide-based functional gene arrays for potential environmental application by focusing on the genes involved in nitrification, denitrification, nitrogen fixation, methane oxidation, and sulfate reduction. Our results demonstrated that the 50-mer func-

Oak Ridge National Laboratory, Oak Ridge, TN, USA

tional gene array is potentially specific and quantitative for detecting microbial populations in environmental samples, and it can provide species-level resolution for analyzing microorganisms involved in nitrification, denitrification, nitrogen fixation, methane oxidation, and sulfate reduction.

MATERIALS AND METHODS

Bacterial Strains, Samples, DNA Extraction, and Purification

The majority of the pure bacteria used in this study were obtained from the ATCC (Manassas, VA, USA) and our own isolate collection: *Desulfovibrio vulgaris* subsp. vulgaris (ATCC 29579) containing the *dsrAB* gene, *Pseudomonas stutzeri* containing the *nirS* gene, *Pseudomonas* sp. G-179 containing the *nirK* gene, *Methanococcus maripaludis* (ATCC 43000) containing the *nifH* gene, *Nitrosomonas europaea* (ATCC 19718) containing the *amoA* gene, and *Methylomicrobium album* (ATCC 33003) containing the *pmoA* gene. The genomic DNAs from these strains were extracted as previously described (10), and the plasmid clones containing the targeted functional genes have also been described (3,11–13).

The performance of microarray hybridization was evaluated using a sample of marine sediment collected from the top 1 cm in the ocean off the Gulf of Mexico. This sample was provided by Allan Devol (University of Washington, Seattle, WA, USA). Community DNAs from the sediment sample were extracted and purified as previously described (10). Two grams of the marine sediment sample from frozen stocks were used for DNA extraction. The extracted DNAs were purified using the Qiagen RNA/DNA Purification System® with Qiagen-tip 20® (Qiagen, Valencia, CA, USA) according to the manufacturer's instructions.

Microarray Design and Construction

The microarray was designed using the information developed by our cloning and sequencing efforts and sequence information from GenBank®. The oligonucleotide microarray was constructed using a diverse set of functional genes involved in nitrogen (*amoA*, *nirS*, *nirK*, and *nifH*), methane (*pmoA*), and sulfur (*dsrAB*) cycling. Almost half of the clone sequences were obtained from our laboratory and isolated from diverse environments such as soil, marine sediment, and ground water (3,11–13). A custom database was established containing the majority of the published *amoA*, *nirS*, *nirK*, *pmoA*, *dsrAB*, and *nifH* sequences and many unpublished sequences obtained from our laboratory. Sequences of interest in GenBank were identified by a keyword search through the National Center for Biotechnology Information (NCBI) web site (http://www.ncbi.nlm.nih.gov/).

The 50-mer oligonucleotide probes were designed using a modified version of the software, PRIMEGENS (http://compbio.ornl.gov/structure/primegens/) (14). Each individual gene sequence was compared against the entire sequence database using the Basic Local Alignment Search Tool (BLAST®; NCBI) and aligned with the other sequences showing more than 85% similarity using dynamic programming. Based on these global optimal alignments, segments of 50-mer with less than 85% nucleotide identity to the corresponding aligned regions of any of BLAST hit sequences were selected as potential probes. Among those identified as potential probes, one probe was selected by considering GC content, melting temperature, and self-complementarity. Outputs of the designed probes were imported into Microsoft® Excel®, and a pivot table was constructed containing the sequence information of each probe. In summary, a total of 763 oligonucleotide probes consisting of 35 *amoA*, 3 *pmoA*, 93 *nirK*, 108 *nirS*, 218 *nifH*, 170 *dsrA*, and 136 *dsrB* genes were designed. Human genes (9) were also used as negative or quantitative controls, and 5 conserved 16S rRNA gene fragments were used as positive controls.

The 50-mer oligonucleotides were synthesized without any modification by MWG-Biotech (High Point, NC, USA) in a 96-well plate format. The oligonucleotides were diluted to a final concentration of 50 pmol/µL in 50% dimethyl sulfoxide (DMSO; Sigma, St. Louis, MO, USA). Microarray printing was carried out as previously described (3). The slides were then cross-linked using an ultraviolet (UV) Stratalinker® 1800 (Stratagene, La Jolla, CA, USA) by exposing the slides to 80 mJ doses of UV irradiation and washing at room temperature with 0.1% sodium dodecyl sulfate (SDS) for 4 min, followed by water for 2 min. Finally, the slides were dried by centrifugation at 50× *g* for 5 min, transferred to a clean slide box, and stored overnight prior to use. The 763 probes and negative and positive control probes were printed on a slide in 2–4 replicates.

Labeling of Target DNA and Microarray Hybridization

Two methods were used to fluorescently label the DNA. For genomic DNA labeling, from 1 ng to 1 µg of genomic DNA was mixed with 1 µg random octamers, denatured by boiling for 5 min, and immediately chilled on ice. The labeling reaction mixture contained denatured DNA, 2.5 µL of dNTP mixture [5 mM each dATP, dTTP, and dGTP, 2.5 mM dCTP (New England Biolabs, Beverly, MA, USA)], 1 µL of Cy3 (25 nmol Cy3-dUTP) or Cy5-dUTP (Amersham Biosciences, Piscataway, NJ, USA), and 40 U of Klenow fragment (Invitrogen, Carlsbad, CA, USA) in a total volume of 50 µL. The reaction mixture was incubated at 37°C for 6 h. Labeled target was purified using a QIAquick™ PCR purification column (Qiagen), concentrated in a SpeedVac® (Thermo Savant, Holbrook, NY, USA) at 40°C for 50 min, and then resuspended in an appropriate volume of distilled water.

DNA targets from plasmid clones were also indirectly labeled by PCR amplification using gene-specific primers. Each PCR used 10 pg of plasmid containing the desired target gene, 20 pmol PCR primers (TA-F, 5′-GCCGCCAGTGTGCTGGAATT-3′ and TA-R, 5′-TAGATGCATGCTCGAGCGGC-3′); 25 mM of dATP, dCTP, dGTP; 15 mM dTTP; 10 mM amino-allyl-dUTP (Sigma); and 0.5 U *Taq* DNA polymerase in a total volume of 30 µL. The amino-allyl-derivitized PCR products were purified using QIAquick PCR purification kit (Qiagen), washed with 75% ethanol, and eluted with water. The amino-allyl-labeled cDNA was dried in a SpeedVac for 30

min and resuspended in 4.5 µL 0.1 M sodium carbonate buffer, pH 9.0. The solution was then reacted with N-hydroxysuccinimide esters Cy3 (NHS-Cy; Amersham Biosciences) and incubated in the dark for 1 h. After incubation, 35 µL of 100 mM sodium acetate, pH 5.2, were added. The labeled products were then purified using QIAquick columns.

Table 1. Sequence Similarities of Selected Genes from Pure Cultures Collected from GenBank

Phylogenetic Hierarchies	Sequence Similarity[a] (%)					
	dsrAB	*nirS*	*nirK*	*nifH*	*amoA*	*pmoA*
Strain	0.93 ± 0.03	0.93 ± 0.04	0.91 ± 0.05	0.93 ± 0.06	0.99 ± 0.01	0.95 ± 0.03
Species	0.73 ± 0.13	0.70 ± 0.18	0.76 ± 0.00	0.82 ± 0.06	0.75 ± 0.11	0.79 ± 0.55
Genus	0.70 ± 0.09	0.67 ± 0.16	0.71 ± 0.07	0.70 ± 0.07	0.71 ± 0.07	0.75 ± 0.15
Family or higher	0.66 ± 0.13	0.57 ± 0.14	0.62 ± 0.14	0.66 ± 0.10	0.65 ± 0.13	N.A.

[a]Data show the mean (± SD) of similarity values of the six gene groups at different phylogenetic levels. N.A., not applicable.

All hybridizations were carried out in triplicate. The hybridization solution contained 3× standard saline citrate (SSC) (1× SSC contained 150 mM NaCl and 15 mM trisodium citrate), 1 µg of unlabeled herring sperm DNA (Promega, Madison, WI, USA), 0.30% SDS, and 50% formamide in a total volume of 15 µL. To avoid bubbles, the hybridization solution was deposited directly onto the microarrays. The microarray (probe-side down) was placed on the coverslip (22 × 22 mm) and then into a hybridization chamber (Corning® hybridization chamber; Corning, Acton, MA, USA). Fifteen microliters of 3× SSC were dispensed into the hydration wells on each side of the microarray slide, and hybridization was carried out for 12–15 h. To determine the effect of temperature on signal intensity, hybridization was carried out at 45°, 50°, 55°, 60°, 65°, 70°, and 75°C. Following hybridization, the arrays were washed with 1× SSC plus 0.2% SDS at hybridization temperature and 0.1× SSC plus 0.2% SDS for 5 min, and then with 0.1× SSC for 30 s each at ambient temperature prior to drying via centrifugation at 500× *g* for 5 min.

Various small microarrays were constructed to determine the optimum oligonucleotide probe concentration, effects of probe sequence similarity on hybridization, and optimal hybridization conditions (i.e., sensitivity and quantitation experiments). To determine the optimal probe concentration for microarray fabrication, a Cy3-labeled oligonucleotide was directly synthesized, diluted to different concentrations (5–200 pmol/µL) in 50% DMSO, and spotted on SuperAmine® slides (TeleChem International, Sunnyvale, CA, USA) in 5 replicates. Immediately after printing, the slides were scanned, and the intensity of the spots was quantified. The intensity of the spots was also quantified after preprocessing (i.e., UV linking and washing as described above) and/or mock hybridization at 55°C for 15 h.

The effects of probe similarity on hybridization signals was determined using small microarrays containing 10 *nirS* oligonucleotide probes with varying levels of similarity (42%–100%) to the *nirS* gene sequence from *P. stutzeri* D7-6. Target DNA was indirectly labeled by PCR amplification as described above and hybridized to the microarrays at 45°, 50°, 55°, 60°, and 65°C for 15 h. Finally, sensitivity and quantitation were evaluated using small microarrays containing six gene probes from different functional gene groups of *dsrB*, *nirS*, *nirK*, *nifH*, *amoA*, and *pmoA*. Each gene was spotted on the arrays in five replicates. Different concentrations of genomic DNA (4–1000 ng) from pure cultures containing *dsrB*, *nirS*, *nirK*, *nifH*, *amoA*, and *pmoA* genes were directly labeled with Cy3 using a random priming method and then hybridized on the small microarrays as described above.

Image Processing and Data Analyses

The microarrays were scanned using an ArrayWoRx™ Station (Applied Precision, Issaquah, WA, USA). The images were saved as 16-bit TIFF files, and each spot was quantified using the ImaGene™ software version 5.0 (Biodiscovery, Los Angeles, CA, USA). Local background measurements were subtracted for each spot. The poor quality spots were flagged and removed from the data set for further analysis. The signal-to-noise ratio (SNR) was also computed for each spot to discriminate true signals from noise. The SNR ratio was calculated as follows:

$$\text{SNR} = \frac{\text{Signal mean - Background mean}}{\text{Background standard deviation}} \quad \text{[Eq. 1]}$$

A commonly accepted criterion for the minimum signal (threshold) that can be accurately quantified is SNR ≥ 3 (15). Spots that appeared to be lower than the threshold value were removed from the data set. Regression analyses of hybridization signal intensity and target or probe concentration were performed using SigmaPlot™ 5.0 (Jandel Scientific, San Rafael, CA, USA).

To determine the level of taxonomic resolution that the designed 50-mer functional gene arrays can provide, the sequence similarities of *dsrAB*, *nirS*, *nirK*, *nifH*, *amoA*, and *pmoA* from pure cultures were compared and evaluated in terms of taxonomic classification (Table 1). Partial sequences of pure cultures containing *dsrAB*, *nirS*, *nirK*, *nifH*, *amoA*, and *pmoA* genes were downloaded from GenBank and grouped according to their taxonomic classification. The obtained sequences were aligned using the PILEUP program from the Genetics Computer Group (GCG) (16). The nucleotide sequence similarities were then determined using the OLDDISTANCE command in GCG (16).

RESULTS

Determination of Optimum Concentration of Oligonucleotide Probes for Microarray Fabrication

The retention of probes on the microarrays was used to determine the optimal probe concentration for fabrication. After printing, the signal intensity increased as the probe concentration in-

creased from 5–100 pmol/µL (data not shown). Within these concentrations, a 5-fold difference in signal intensity was observed. However, between 100–200 pmol/µL, the signal intensity was very similar. These results indicated that the saturation point of probe concentration was 100 pmol/µL after printing.

To assess the effect of preprocessing and hybridization on probe retention, the intensity of the spots was also quantified following preprocessing and mock hybridization at 55°C for 15 h. After preprocessing, no significant difference in signal intensity was observed between 50 and 100–200 pmol/µL (data not shown), indicating that the saturation point of signal intensity after preprocessing was about 50 pmol/µL. Following mock hybridization, the saturation point further declined to between 30–40 pmol/µL (data not shown). The signal intensity at these concentrations was similar to that of 50, 100, and 200 pmol/µL. Our data suggest that the optimum oligonucleotide concentrations were between 30–40 pmol/µL for microarray fabrication, and excess probes will be washed away after hybridization. Because synthesizing oligonucleotides for array fabrication is expensive, an oligonucleotide probe concentration of 40 pmol/µL was therefore used for microarray construction in later experiments.

Effect of Temperature on the Oligonucleotide Microarray Hybridization

To determine the effect of temperature on oligonucleotide hybridization, the PCR-amplified DNA fragments of the *dsrB*, *nirK*, *nirS*, *nifH*, *pmoA*, and *amoA* genes were hybridized on the arrays containing 763 probes at different temperatures (data not shown). Prior to hybridization, slides from the beginning, middle, and end of the printing run were stained with POPO™ iodide (Molecular Probes, Eugene, OR, USA) to ensure that the quality of spots was uniform across the array of slides (40 slides per run) by visual observation. The quality of the spots was similar for slides from the beginning, middle, and end of the print (data not shown). At 45°C, the signal intensity was the highest but significant cross-hybridization was observed (data not shown). As hybridization temperature increased, the signal intensity decreased dramatically. At 70° and 75°C, little or no hybridization was observed for all of the genes examined. Oligonucleotide probes hybridized strongly to their complementary sequences at 50°C, and the hybridization signal intensity was much stronger compared to that at 55°, 60°, or 65°C. No cross-hybridization was observed among different gene groups

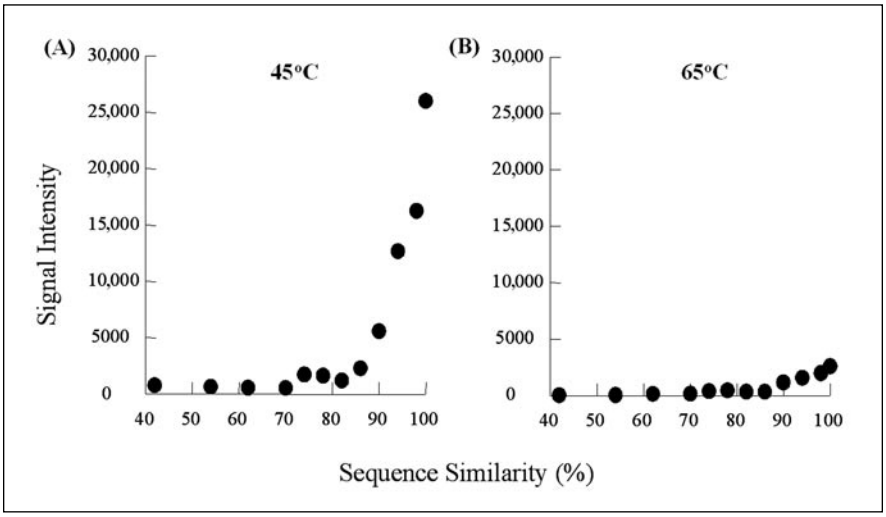

Figure 1. Effects of probe sequence identities on hybridization signal intensity. The experiments were carried out using small arrays fabricated with the probe sequences derived from DNA from *nirS*-containing *Pseudomonas stutzeri.* These probe sequences have varying degrees of similarities (40%, 50%, 60%, 70%, 80%, 90%, 94%, 98%, and 100%) to the target sequence. Target DNA was labeled by PCR amplification and hybridized to the microarrays at (A) 45°C and (B) 65°C for 15 h. One human gene probe was also spotted on the arrays as a negative control.

at 50°–65°C (data not shown). In addition, no hybridization was observed with any of the 10 human gene probes. Based on these results, a hybridization temperature at 50°C was used in most of the later experiments.

Effect of Probe Sequence Similarities on Hybridization Signal Intensity

Previous work demonstrated that target genes having less than 80%–85% sequence identity could be differentiated at high stringency using DNA fragments (400–800 bp) as probes (3). Based on this finding, all of the oligonucleotide probes that were selected for microarray fabrication had sequence identities less than 85%. No probes were designed for the genes having more than 85% similarity.

To experimentally determine the resolution power of the hybridization with the 50-mer functional gene arrays, the effects of probe sequence similarity on hybridization signal intensity of the target *nirS* gene from *P. stutzeri* were evaluated. The signal intensity of the *nirS* gene to the oligonucleotide probes was the highest at low stringency (45°C) and the lowest at high stringency (65°C) (Figure 1). The highest signal intensity was observed with the oligonucleotide probe having 100% similarity to the *nirS* gene from *P. stutzeri.* As the similarity decreased, the signal intensity also decreased. At all temperatures, little hybridization was observed for probes showing 40%–86% similarity to the target DNA, whereas the signal intensity increased substantially for probes showing more than 86% similarity to the target DNA. In addition, the signal intensities of probes having 86%, 90%, 94%, 98%, and 100% identity to the target DNA were also substantially different from each other (Figure 1).

The signal intensities for the probes that were less than 86% similar were comparable to background signal intensity at hybridization temperatures greater than 50°C. To determine whether the hybridization signal intensities were true signals, the SNR was calculated (Table 2). SNR decreased with increasing hybridization temperatures (Table 2). For instance, at 45°C, the perfectly matched target DNA gave an SNR of 35.6 (10 times higher than the threshold). However, this ratio decreased to below threshold (SNR = 2.2) at 65°C. SNR also decreased with the decrease of probe sequence similarities. At 45°, 50°, and 55°C, the lowest detectable signals were observed at 86%, 90%, and 94% similarities, respectively. At 60° and 65°C, the signal intensities of all targeted probes were below the detection limit of SNR = 3 (Table 2). These results suggest that at the hybridization conditions of 50°C, the oligonucleotide microarray hybridization can differentiate sequences having less than 86% similarity, whereas at 55°C,

Table 2. Signal-to-Noise Ratio of Microarray Hybridizations at Different Hybridization Temperatures

Probe Sequence Identity (%)	Signal-to-Noise Ratio (°C)				
	45	50	55	60	65
100	35.6 ± 7.1	18.4 ± 2.9	8.2 ± 4.3	2.9 ± 0.9	2.2 ± 0.7
98	23.5 ± 10.4	11.0 ± 4.7	7.1 ± 3.8	2.9 ± 1.0	2.3 ± 0.2
94	19.4 ± 10.9	6.6 ± 1.5	4.4 ± 1.9	2.3 ± 0.3	1.9 ± 0.4
90	7.9 ± 5.9	5.1 ± 1.0	2.7 ± 1.4	2.3 ± 0.2	1.6 ± 0.4
86	3.8 ± 0.8	1.8 ± 0.5	1.2 ± 0.5	1.1 ± 0.4	0.5 ± 0.2
82	2.3 ± 1.0	1.8 ± 0.7	1.2 ± 0.7	1.4 ± 0.7	0.6 ± 0.3
78	3.4 ± 1.8	1.4 ± 1.0	0.8 ± 0.6	1.1 ± 0.6	0.9 ± 0.4
74	3.4 ± 1.6	1.5 ± 0.5	1.1 ± 0.9	1.2 ± 0.7	0.7 ± 0.3
70	1.4 ± 1.2	1.1 ± 0.7	1.2 ± 1.3	1.0 ± 0.5	0.3 ± 0.1
62	1.8 ± 1.8	1.3 ± 0.9	1.1 ± 1.1	1.2 ± 0.9	0.3 ± 0.2
54	1.4 ± 0.8	0.8 ± 0.5	0.5 ± 0.3	0.4 ± 0.3	0.0 ± 0.1
42	1.1 ± 0.2	1.0 ± 0.5	0.6 ± 0.4	0.4 ± 0.3	0.0 ± 0.1

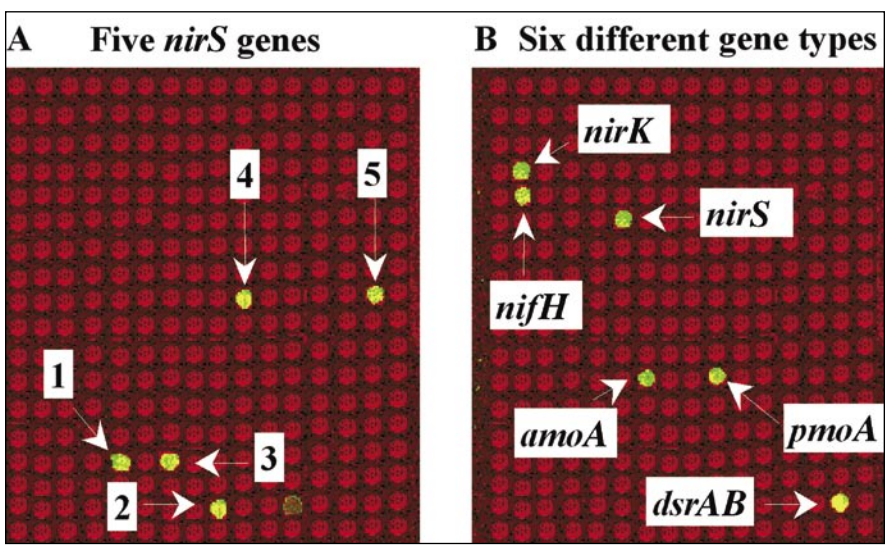

Figure 2. Images showing hybridization specificity. (A) Five different *nirS* genes (clone M305027, clone D10-300-17, clone M318A96, clone M312A24, and clone M312A29). All of these genes have less than 85% nucleotide identity. The range of similarities between the probes was 44%–60%. (B) The combination of six different gene groups (*dsrA*, clone FW003001A; *nirK*, clone NKTT14; *nirS*, clone M304B04; *nifH*, clone NIS3-3; *pmoA*, clone C01A15143; *amoA*, clone H06A16104). The target DNAs were labeled with Cy3 using PCR amplification and hybridized separately with the microarrays containing 763 probes at 50°C. The background shows the layout of the array spots stained with POPO dye (red). Specific hybridizations are shown in green.

sequences having less than 90% similarity can be differentiated.

Specificity of the 50-mer Oligonucleotide Array Hybridization

The specificity of hybridization with the 50-mer oligonucleotide array across 763 probes was determined using environmental clones from different functional gene groups (*dsrA, nirK*, *nirS*, *nifH*, *pmoA*, and *amoA*) as targets. These cloned genes were labeled with Cy3 by PCR amplification. Each labeled DNA was hybridized separately to the oligonucleotide microarrays at 50°C to determine the specificity of individual probes. Strong hybridizations were observed with each target gene at 50°C, whereas no hybridization was observed with the nontarget genes (data not shown). Specific hybridization was obtained when five different *nirS* clones showing less than 85% sequence identity (clones M305027, D10-300-17, M318A96, M312A24, and M312A29) were mixed together (Figure 2A). In addition, when the labeled target genes from six different functional gene groups were combined and hybridized to the oligonucleotide arrays, strong hybridization was observed only with their corresponding probes sequences (Figure 2B). Finally, five additional, different genes from each functional gene group (total of 30 genes) were selected and hybridized with the arrays. As expected, strong hybridization with the target genes and no cross-hybridization with the nontarget genes were observed (data not shown). These results indicated that the 50-mer functional gene arrays containing 763 gene probes appear to be specific to their corresponding target genes.

Detection Sensitivity of 50-mer Functional Gene Array Hybridization

To evaluate the detection sensitivity of hybridization with the oligonucleotide array, different concentrations of genomic DNAs from pure cultures containing *dsrB*, *nirS*, *nirK*, *nifH*, *amoA*, or *pmoA* genes were randomly labeled with fluorescent dyes and hybridized with an array containing a small number of probes. Very strong hybridization signals were observed with DNA concentrations between 32–1000 ng from all gene groups (data not shown), while hybridization signals were weaker with DNA concentrations between 4–8 ng (Figure 3). The positive controls (16S rRNA oligonucleotide probes) had similar hybridization signal intensities to the probes from the six different functional gene groups (Figure 3). Little or no hybridization was observed at 4 ng of genomic DNA for all six functional gene groups and for positive controls (data not shown). The detection limit with randomly labeled genomic DNA from pure cultures under the above hybridization conditions was therefore estimated to be approximately 8 ng.

In the above experiment, the detection limit was determined with genomic DNA from a single pure culture in the absence of heterogeneous DNA templates, and thus the detection limit could be lower. In addition, the detection limit determined with genomic DNA from a single pure culture may not be directly applicable to real environmental samples because of the complexity of microbial communities in environmental samples. The existence of other nontarget DNAs may affect the hybridization with target DNA and therefore decrease detection sensitivity. To evaluate the detection sensitivity within the context of environmental applications, genomic DNAs representing eight target bacteria from different genera and species were mixed at different concentrations, fluorescently labeled, and hybridized with the functional gene arrays (see Figure 4 legend for details). The signal intensity was significantly higher than background level when DNA concentrations were greater than 50 ng, whereas the signal intensity was comparable to the background level when the genomic DNA concentration ranged from 8–30 ng (Figure 4).

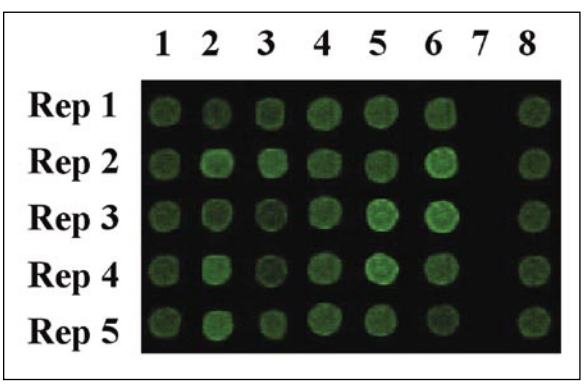

Figure 3. Detection sensitivity of the oligonucleotide microarrays. A small array was constructed with five replicates of each probe. These probes are *dsrB*-containing *Desulfovibrio vulgaris* (lane 1), *nirS*-containing *Pseudomonas stutzeri* (lane 2), *nirK*-containing *Pseudomonas* sp. G-179 (lane 3), *nifH*-containing *Methanococcus maripaludis* (lane 4), *amoA*-containing *Nitrosomonas europea* (lane 5), *pmoA*-containing *Methylomicrobium album* (lane 6), a human gene as a negative control (lane 7), and 16S rRNA (lane 8). Different amounts of target DNAs (4–1000 ng) were mixed together, labeled with Cy3 using the random primer labeling method, and hybridized with the smaller arrays. The hybridization image with 8 ng genomic DNA is presented.

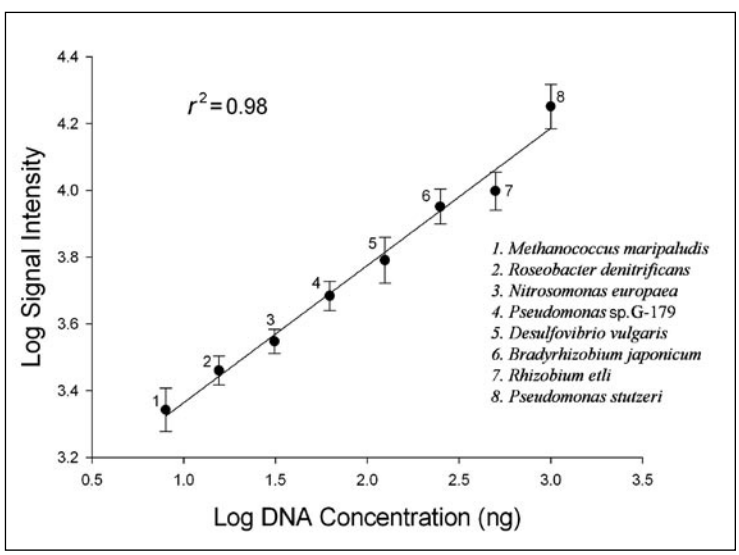

Figure 4. Quantitative analysis of the 50-mer functional gene arrays using a mixed genomic DNA population. Eight different genomic DNAs were mixed together in different concentrations and labeled with Cy5. The plot shows the log-transformed average hybridization intensity versus the log-transformed target DNA concentration for each strain.

The detection limit of the functional gene array-based hybridization in the presence of nontarget DNAs under the standard hybridization condition was therefore approximately 60 ng.

Quantitative Potential of Oligonucleotide Array Hybridization

The quantitative capability of microarrays is important for understanding the activities and dynamics of microbial communities. To evaluate whether oligonucleotide-based arrays can be used as a quantitative tool, the relationship between target DNA concentration and hybridization signal was determined. Genomic DNAs from pure cultures containing *dsrB*, *nirS*, *nirK*, *nifH*, *amoA*, or *pmoA* genes were diluted at different quantities, labeled with fluorescent dyes, and hybridized with the arrays. A good linear relationship (n = 40) was observed between signal intensity and target DNA for *dsrB* (r^2 = 0.98; slope = 0.72), *nirK* (r^2 = 0.97; slope = 0.73), *nirS* (r^2 = 0.97; slope = 0.67), *nifH* (r^2 = 0.96; slope = 0.81), *amoA* (r^2 = 0.96; slope = 0.74), and *pmoA* (r^2 = 0.98; slope = 0.78) genes from pure cultures (data not shown). These results suggest that the microarray hybridization for pure bacterial cultures appears to be quantitative within a wide range of DNA concentrations (8–1000 ng).

To determine whether the functional gene array hybridization is quantitative for targeted organisms in the presence of other nontarget DNAs, the quantitative relationships between signal intensity and DNA concentrations were further examined with the mixture of DNAs from eight different bacteria that represented different genera and species as mentioned above. A significant linear relationship (r^2 = 0.96; $P < 0.01$) was observed between signal intensity and target DNA concentration within a concentration range of 60–1000 ng (Figure 4). These results suggest that the 50-mer functional gene array-based hybridization could be quantitative for mixed DNA templates.

Environmental Application

To evaluate the potential feasibility of using 50-mer functional gene arrays to detect microbial populations in environmental samples, 5 µg of bulk community DNAs from a marine sediment sample were labeled with Cy3 using the random primer labeling method and hybridized directly with the oligonucleotide arrays. The hybridization image indicated that the 50-mer oligonucleotide arrays hybridized reasonably well with the DNAs from marine sediment (Figure 5). Strong signals were obtained with some nitrogenases (*nifH*), dissimilatory sulfite reductase (*dsrAB*), ammonia monooxygenase (*amoA*), methane monooxygenase (*pmoA*), and nitrite reductase (*nirSK*) (Figure 4, partial image shown). These results suggested that the oligonucleotide microarray technology developed in this study is potentially useful for detecting dominant microbial populations involved in these functional processes.

DISCUSSION

In this paper, we examined the specificity, sensitivity, and quantitative aspects of using the 50-mer functional gene array-based hybridization to analyze environmental samples. Other researchers have evaluated the use of different oligonucleotide lengths for microarrays (17–21). In general, oligonucleotides longer than 30 bp are used (17,20) because the sensitivity of microarrays containing longer oligonucleotides is much higher than that of shorter oligonucleotides or that of cDNA microarrays (17,20). For instance, the hybridization signals of the oligonucleotide arrays containing 60-mer are 10-fold higher than those with 25-mer and 2-fold higher than those with 30-mer (19). The arrays with 60-mer can also provide significantly better sensitivity than those with 25- to 30-mer in complex mixtures of RNAs (19). Therefore, longer oligonucleotides, such as those investigated in the present study, provide higher sensitivity compared to oligonucleotides containing shorter oligonucleotides. In contrast to the construction of arrays containing shorter oligonucleotides, no modification of the oligonucleotides is required for constructing microarrays containing longer oligonucleotides. However, the specificity of a short oligonucleotide (25-mer) is 20-fold better than that of a long oligonucleotide (60-mer) for differentiating single mismatch (19). In addition, it is more difficult to analyze oligonucleotides longer than 50 bp using mass spectrometry. Therefore, in this study, the probe size of 50 bp was selected.

In this study, we found that the 50-

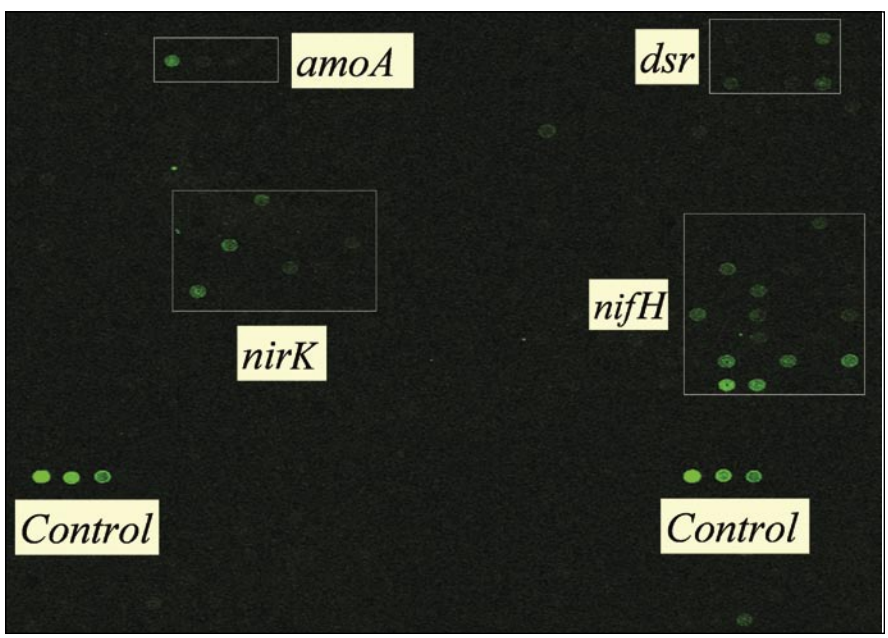

Figure 5. Microarray hybridization with a marine sediment sample. Five micrograms of DNA from the marine sediment sample were labeled with Cy3 using random primer labeling method and hybridized at 50°C overnight to the 50-mer oligonucleotide arrays. A small portion of the hybridization image is shown.

mer oligonucleotide arrays containing probes involved in nitrogen and sulfur cycling were able to discriminate gene sequences with less than 86%–90% similarity at 50°–55°C. As expected, higher hybridization specificity can be achieved with the 50-mer functional gene arrays than the DNA-based functional gene arrays.

With the 50-mer oligonucleotide arrays, *dsrB*, *nirS*, *nirK*, *nifH*, *amoA*, and *pmoA* genes could be detected with 8 ng of pure genomic DNA. However, as expected, the sensitivity is 10 times lower than the DNA-based functional gene arrays and 100 times lower than community gene arrays (3,13). Using the DNA-based functional gene arrays, *nirS* genes were detected with about 1 ng of labeled pure genomic DNA (3). For community gene arrays, the detection limit was estimated to be approximately 0.2 ng with pure labeled genomic DNA (Wu et al., manuscript submitted). One of the main reasons for the lower sensitivity of the 50-mer functional gene arrays is that the probes are much shorter than the probes used in DNA-based functional gene arrays and community gene arrays, which have more binding sites available for capturing the labeled target DNAs. In the presence of background DNA, the sensitivity of the 50-mer functional gene arrays was further decreased (approximately 50 ng pure genomic DNA). If these values can be directly applicable to real environmental samples, then the level of the 50-mer-based functional gene array detection sensitivity should be sufficient for detecting dominant microbial populations, but it may not be sensitive enough to detect less abundant microbial populations. The current detection limit of the technology is about 5% of the total population analyzed (22,23). This means that if cells are present at a lower ratio, they may be missed. Therefore, new technology will be needed to improve the performance of microarrays for environmental applications in the future.

Environmental studies require experimental tools that not only detect the presence or absence of a particular group of microorganisms but also provide quantitative information. As shown by our results, there were good quantitative correlations between hybridization signal and the amount of genomic DNAs. This result is consistent with those found by Relogio et al. (19) using oligonucleotide probes and by Wu et al. (3) using DNA-based functional gene arrays. However, there are a few obstacles to quantitation of microbial populations with microarrays. These include differences in gene copy numbers, differences in signal intensities due to minor sequence variation, and possible bias with Klenow labeling. Because environmental samples are very diverse, potential cross-hybridiza-

tion could occur among the sequences showing more than 86%–90% similarities. Thus, caution is needed in data interpretation when dealing with environmental samples of unknown diversity.

Because only the probes with less than 85% similarity were used to construct the 50-mer functional gene arrays, it is unclear what level of taxonomic resolution the designed oligonucleotide microarrays can provide. To address this question, the sequence similarities of *dsrAB*, *nirS*, *nirK*, *nifH*, *amoA*, and *pmoA* from pure cultures were compared and evaluated in terms of taxonomic classification (Table 1). At the strain level, the average sequence similarity for *amoA* was 99%, whereas it was lower (91%–95%) for the other five functional gene groups. At the species level, the average similarity decreased to 70%–82%, and it further dropped at the levels of genus (67%–75%) and family or higher (57%–66%) (Table 1). These results indicated that the 50-mer functional gene arrays appear to be able to provide species-level resolution of microorganisms involved in these biogeochemical processes. However, to obtain strain-level resolution, microarrays containing shorter oligonucleotides (e.g., 20 bp) are needed.

ACKNOWLEDGMENTS

S.M.T. and L.W. contributed equally to this work. We thank Lynn Kszos for editorial assistance. This research was supported by the US Department of Energy (DOE) Office of Science as part of its Biological and Environmental Research Programs in Biotechnology Investigations-Ocean Margins program, Natural and Accelerated Bioremediation Research, 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 US DOE under contract no. DE-AC05-00OR22725.

REFERENCES

1. **Amann, R.I., W. Ludwig, and K.H. Schleifer.** 1995. Phylogenetic identification of in situ detection of individual microbial cells without cultivation. Microbiol. Rev. *59*:143-169.
2. **Zhou, J.** 2003. Microarrays for bacterial detection and microbial community analysis. Curr. Opin. Microbiol. *6*:288-294.
3. **Wu, L.Y., D.K. Thompson, G. Li, R.A. Hurt, J.M. Tiedge, and J. Zhou.** 2001. Development and evaluation of functional gene arrays for detection of selected genes in the environment. Appl. Environ. Microbiol. *67*:5780-5790.
4. **Beliaev, A.S., D.K. Thompson, M.W. Fields, L. Wu, D.P. Lies, K.H. Nealson, and J. Zhou.** 2002. Microarray transcription profiling of a *Shewanella oneidensis etr*A mutant. J. Bacteriol. *184*:4612-4616.
5. **DeRisi, J.L., V.R. Iyer, and P.O. Brown.** 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. Science *278*:680-686.
6. **Liu, Y., J. Zhou, M. Omelchenko, A. Beliaev, A. Venkateswaran, J. Stair, L. Wu, D.K. Thompson, et al.** 2003. Transcriptome dynamics of *Deinococcus radiodurans* recovering from ionizing radiation. Proc. Natl. Acad. Sci. USA *100*:4191-4196.
7. **Zhou, J. and D.K. Thompson.** 2002. Microarrays: application in environmental microbiology, p. 1968-1979. *In* G. Bitton (Ed.), Encyclopedia of Environmental Microbiology, vol IV. John Wiley and Sons, New York.
8. **Hughes, T.R., M. Mao, A.R. Jones, J. Burchard, M.J. Marton, K.W. Shannon, S.M. Lefkowitz, M. Ziman, et al.** 2001. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. Nat. Biotechnol. *19*:342-347.
9. **Lessard, I.A.D., G.J. Domingo, A. Borges, and R.N. Perham.** 1998. Expression of genes encoding the E2 and E3 components of the *Bacillus stereothermophilus* pyruvate dehydrogenase complex and the stoichiometry of subunit interaction in assembly in vitro. Eur. J. Biochem. *258*:491-501.
10. **Zhou, J.Z., M.A. Bruns, and J.M. Tiedje.** 1996. DNA recovery from soils of diverse composition. Appl. Environ. Microbiol. *62*:461-468.
11. **Zhou, J.Z., B. Xia, H. Huang, D.S. Treves, L.J. Hauser, R.J. Mural, A.V. Palumbo, and J.M. Tiedje.** 2003. Phylogenetic diversity and a novel candidate division of two humid region sandy surface soils. Soil Biol. Biochem. *35*:915-924.
12. **Liu, X., S.M. Tiquia, G. Holguin, L. Wu, A. Devol, S.C. Nold, A.V. Palumbo, J.M. Tiedje, et al.** 2003. Molecular diversity of denitrifying genes in continental margin sediments within the oxygen deficient zone of the Pacific Coast of Mexico. Appl. Environ. Microbiol. *69*:3549-3560.
13. **Yan, T., M.W. Fields, L. Wu, Y. Zu, J.M. Tiedje, and J. Zhou.** 2003. Molecular diversity and characterization of nitrite reductase gene fragments (*nirK* and *nirS*) from nitrate- and uranium-contaminated groundwater. Environ. Microbiol. *5*:13-24.
14. **Xu, D., G. Li, L. Wu, J.-Z. Zhou, and Y. Xu.** 2001. PRIMEGENS: a computer program for robust and efficient design of gene-specific probes for microarray analysis. Bioinformatics *18*:1432-1437.
15. **Verdnik, D., S. Handran, and S. Pickett.** 2002. Key considerations for accurate microarray scanning and image analysis, p. 89-98. *In* G. Kamberova, (Ed.), DNA Image Analysis: Nuts & Bolts. DNA Press LLC, Salem, MA.
16. **Genetics Computer Group (GCG).** 1994. GCG program manual. Genetics Computer Group, Madison, WI.
17. **Hughes, T.R. and D.D. Shoemaker.** 2001. DNA microarrays for expression profiling. Curr. Opin. Chem. Biol. *5*:21-25.
18. **Proudnikov, D., E. Timofeev, and A. Mirzabekov.** 1998. Immobilization of DNA in polyacrylamide gel for the manufacture of DNA and DNA-oligonucleotide microchips. Anal. Biochem. *259*:34-41.
19. **Relogio, A., C. Schwager, A. Richter, W. Ansorge, and J. Valcarcel.** 2002. Optimization of oligonucleotide-based DNA microarrays. Nucleic Acids Res. *30*:e51.
20. **Selinger, D.W., K.J. Cheung, R. Mei, E.M. Johansson, C.S. Richmond, F.R. Blattner, D.J. Lockhart, and G.M. Church.** 2000. RNA expression analysis using a 30-base pair resolution *Escherichia coli* genome array. Nat. Biotechnol. *18*:1262-1268.
21. **Wang, D.L., L. Coscoy, M. Zylberberg, P.C. Avila, H.A. Boushey, D. Ganem, and J.L. DeRisi.** 2002. Microarray-based detection and genotyping of viral pathogens. Proc. Natl. Acad. Sci. USA *99*:15687-15692.
22. **Denef, V.J., J. Park, J.L. Rodrigues, T.V. Tsoi, S.A. Hashsham, and J.M. Tiedje.** 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.
23. **Bodrossy, L., N. Stralis-Pavese, J. Colin Murell, S. Radajewski, A. Weiharter, and A. Sessitsch.** 2003. Development and validation of a diagnostic microbial microarray for methanotrophs. Environ. Microbiol. *5*:566-582.

Received 12 November 2003; accepted 3 February 2004.

Address correspondence to:

Jizhong Zhou
Environmental Sciences Division
Oak Ridge National Laboratory
Oak Ridge, TN 37831, USA
e-mail: zhouj@ornl.gov