

Chapter 11

Hybridization of Environmental Microbial Community Nucleic Acids by GeoChip

Joy D. Van Nostrand, Huaqin Yin, Liyou Wu, Tong Yuan, and Jizhong Zhou

Abstract

Functional gene arrays, like the GeoChip, allow for the study of tens of thousands of genes in a single assay. The GeoChip array (5.0) contains probes for genes involved in geochemical cycling (N, C, S, and P), metal homeostasis, stress response, organic contaminant degradation, antibiotic resistance, secondary metabolism, and virulence factors as well as genes specific for fungi, protists, and viruses. Here, we briefly describe GeoChip design strategies (gene selection and probe design) and discuss minimum quantity and quality requirements for nucleic acids. We then provide detailed protocols for amplification, labeling, and hybridization of samples to the GeoChip.

Key words GeoChip, Functional gene array, Microbial communities, Microbial ecology, Hybridization, Fluorescent labeling, Whole community genome amplification

1 Introduction

Microorganisms are critical for global biogeochemical cycling of N, C, S, and metals; however, a large percentage of microorganisms (>99 %) remain uncultured [1–3] making it difficult to fully examine microbial community activity. To overcome this limitation, culture-independent approaches are needed. Examination of functional genes, those genes that are involved in processes of interest (e.g., *nifH* for N fixation, *dsrA* for sulfite reduction, etc.), can shed light on the functional abilities of microbial communities. While there are a wide variety of culture-independent approaches that could be used to examine functional genes, most can provide information on only a small number of functional genes. Microarrays, on the other hand, allow the examination of tens of thousands of genes at one time, thus allowing for a comprehensive examination of a wide range of genes.

Functional gene microarrays (FGAs) probe for multiple functional genes involved in microbial functional processes of interest at one time [4–6]. Because FGAs can shed light on the

functional potential of microbial communities, they are ideal linking microbial functional structure with environmental processes. The most comprehensive FGA to date is the GeoChip microarray [5–7]. The GeoChip was designed to address two major challenges in using FGAs for microbial community analysis: (1) the need for adequate oligonucleotide probe specificity for genes that have high homologies and (2) the lack of a truly comprehensive probe set [5].

The following chapter briefly describes the process of gene selection and probe design for the GeoChip microarrays and describes the sample preparation and hybridization protocols for the GeoChip arrays. The GeoChip is currently manufactured by Agilent Technologies (Santa Clara, CA, USA), but there are other companies that manufacture arrays and they can also be printed in-house. These hybridization methods included are specific for Agilent arrays, but other methods can be used for other array types and there are methods for using Agilent arrays with automated hybridization stations. A general overview of the GeoChip analysis protocol is shown in Fig. 1.

The first step for GeoChip design is the selection of genes representing processes of interest. GeoChip was designed to

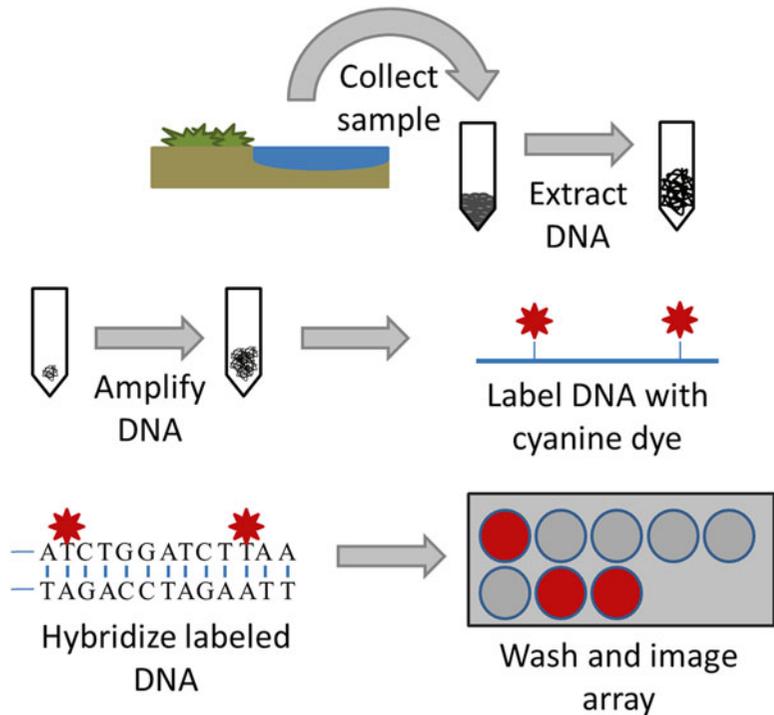


Fig. 1 Overview of the GeoChip protocol. Samples from the environment of interest are collected and DNA is extracted. If the yield of DNA is insufficient, whole community genome amplification can be performed to increase the quantity of DNA. The DNA is then labeled with a cyanine dye and hybridized to the GeoChip. Any unhybridized DNA is then washed off and the array is imaged

examine microbial community functional potential, so only genes for those proteins actively involved in a process of interest, i.e., those containing catalytic subunits or active sites, are included. Next, public databases are searched using keywords that are chosen to select a wide range of sequences that may be the gene of interest. These sequences are then confirmed by HMMER alignment (<http://hmmmer.wustl.edu/>) with seed sequences that have been experimentally confirmed. These confirmed sequences are then used for probe design (50mer) using experimentally determined criteria based on sequence homology (≤ 90 % identity for gene-specific probes and ≥ 96 % for group-specific probes), continuous stretch length (≤ 20 bases for gene-specific probes and ≥ 35 for group-specific probes), and free energy (≥ -35 kJ/mol for gene-specific probes and ≤ -60 kJ/mol for group-specific probes) [8, 9] with new versions of the CommOligo software [10]. The final step is to confirm specificity by BLASTing the probe sequences against the GenBank database.

There are a number of DNA extraction and purification kits and methods available for microbial community samples. As long as the resultant nucleic acid quantity and quality is sufficient and the fragment length is fairly large, any method can be used. The absorbance ratio (A260:A280) should be ~ 1.8 and > 1.9 for DNA and RNA, respectively, and an A260:A320 > 1.7 . The A260:A230 ratio is most important for microarray success [11]. A high A230 value may indicate contamination with EDTA, carbohydrates, phenol, or guanidine HCl [12]. For samples with high humics, a gel purification strategy followed by a phenol-chloroform-butanol extraction [13, 14] may be needed. If a low yield of nucleic acid is expected, large nucleic acid fragments are needed for efficient amplification.

GeoChip requires 0.5–1 μg of DNA or 10 μg of RNA. If this amount cannot be obtained, an amplification step will be needed. Whole community genome amplification (WCGA) using the Templphi 500 amplification kit (phi 29 DNA polymerase, GE Healthcare, Piscataway, NJ) and a modified reaction buffer can be used. This method provides a sensitive (10 fg detection limit) and representative amplification (< 0.5 % of amplified genes showed > 2 -fold different from unamplified) from 1 to 100 ng template DNA [15]. Whole community RNA amplification (WCRA) provides a representative amplification with 50–100 ng of starting material [16].

2 Materials

2.1 Reagents and Kits

1. Templphi 500 amplification kit (GE Healthcare, Piscataway, NJ, USA).
2. Single-stranded binding protein (SSB) (Affymetrix, Santa Clara, CA, USA).

3. Quant-iT™ PicoGreen® dsDNA kit (Life Technologies, Carlsbad, CA, USA).
4. dNTP mix: 5 mM (2.5 mM dTTP). Combine 5 μ L each of 100 mM dATP, dGTP, and dCTP and 2.5 μ L 100 mM dTTP. Bring to 100 μ L with 82.5 μ L nuclease-free water.
5. 40 U/ μ L Klenow (IMER, Inc., San Diego, CA, USA).
6. 25 nM Cy3-dUTP (GE Healthcare).
7. QIAquick PCR Purification kit (Qiagen, Valencia, C, USA).
8. Cot-1 Human DNA (catalog number 5190-3393) (Agilent).

2.2 Buffers and Solutions

1. 2.4 mM spermidine. Weight out 1.66 g spermidine and make up to 100 mL with water.
2. Hybridization buffer: Part of the Oligo aCGH hybridization kit, large, catalog number 5188-5380 (Agilent).
3. Blocking agent: Add 1350 μ L water to the 10 \times aCGH Blocking Agent (Part of the Oligo aCGH hybridization kit, large, catalog number 5188-5380) (Agilent). Incubate at room temperature for at least 6 h before using.
4. Wash Buffers: Oligo aCGH/ChIP-on-Chip Wash Buffer kit (catalog number 5188-5226, contains Wash Buffers 1 and 2) (Agilent).

2.3 Probes and Primers

1. Random primers (mostly hexamers, 3 μ g/ μ L, catalog number 48190-011) (Life Technologies).
2. Universal standard. The universal standard or common oligo reference standard (CORS) target is complimentary to a CORS probe printed onto the GeoChip microarray. The CORS target can be synthesized by an array of manufacturers. We use Eurofins MWG Operon for target synthesis. The nucleotide sequence contains Cy-5 label on the 5' with purification by desalting. [Cy5]GCCAGCACAGCTACACGTCCTCAAACG ATTGTGTGCGGTCCGAGGTGCGG. See Liang et al. [13], for full details of the CORS development.

2.4 Equipment

1. NanoDrop spectrometer (Thermo Scientific, Wilmington, DE, USA) or similar instrument capable of measuring cy-dye incorporation (568 and 647 nm).
2. General molecular lab equipment: thermocyclers, fluorescent plate reader (for PicoGreen measurement), vacuum concentrator for drying samples.
3. Microarrays: GeoChip 5 uses Agilent arrays (SurePrint G3 custom array) in two sizes: 8 \times 60K, 8 arrays of 60,000 probes per slide or 4 \times 180K, 4 arrays of 180,000 probes per slide.

4. Gasket slides (catalog number G2534-60013 for four arrays per slide or G2534-60016 for eight arrays per slide) (Agilent).
5. Hybridization equipment: hybridization oven, hybridization oven rotator rack, hybridization oven conversion rod, hybridization chamber (Agilent), magnetic stir plate with heating element.
6. Microarray scanner: The SureScan Microarray Scanner (Agilent Technologies, Santa Clara, CA) is ideal for Agilent microarrays, but other microarray scanners having a 2–3 μm resolution and the ability to image Cyanine 3 and 5 can be used.

3 Methods

3.1 Amplification

The GeoChip requires 1 μg DNA for hybridization. While many samples, such as soil, can easily meet this criterion, sites with low microbial abundance or that have restrictions on the amount of sample that can be taken, may not yield enough DNA. In these cases, WCGA can be used to increase the amount of DNA available. WCGA uses a modification of the Templiphi 500 amplification kit (phi 29 DNA polymerase) [15]. The amplification buffer is supplemented with single-stranded DNA binding protein (SSB) and spermidine and a larger amount of enzyme to increase sensitivity and representative amplification. The SSB and spermidine likely assist with DNA replication and bind inhibitors [17, 18]. Using 1–100 ng DNA provides a sensitive (10 fg detection limit) and representative amplification (<0.5 % of amplified genes showed >2-fold different from unamplified) [15].

This amplification reaction is very sensitive and will amplify any contaminating DNA. As such, all steps should be performed in a PCR hood. Zhang et al. [19] have outlined additional steps that should be followed to minimize contamination. These include UV irradiation of the hood and all items to be used in the protocol (i.e., tips, tubes, pipettors, tube racks, ice and ice bucket, etc.). Negative controls should always be run alongside the samples.

1. Add 10 μL sample buffer to a 0.2 mL PCR tube (*see Note 1*).
2. Add DNA (preferably 100 ng) (*see Notes 2 and 3*).
3. Mix DNA and buffer thoroughly and incubate at room temperature (RT) for 10 min.
4. While DNA and buffer are incubating, prepare Templiphi premix in a 1.7 mL tube (*see Note 1*). [For each reaction: 10 μL of reaction buffer, 0.6 μL enzyme mixture (both supplied in the Templiphi kit), 1.25 μL of 5 $\mu\text{g}/\mu\text{L}$ SSB (USB; Cleveland, OH), and 1 μL of 2.4 mM spermidine stock.]

5. Transfer 12.85 μL of the premix to the DNA/buffer mixture, mix well and spin twice.
6. Incubate the reaction at 30 $^{\circ}\text{C}$ for 6 h then heat-inactivate the enzyme at 65 $^{\circ}\text{C}$ for 10 min.
7. Run an aliquot ($\sim 2 \mu\text{L}$) of the product on a gel. The product should produce a smear rather than a single band (*see Note 4*).
8. Quantify the amplified DNA using a dye-binding assay, such as PicoGreen (Quant-iT™ PicoGreen® dsDNA kit). Using the 260/280 ratio will result in an erroneous quantity due to the primers and dNTPs remaining in the sample.

3.2 Labeling

DNA for microarray hybridization is generally labeled using fluorescent dyes such as cyanine dyes. The DNA can be labeled directly (dyes are directly integrated into the target DNA) or indirectly (targets are labeled after hybridization). A direct labeling approach is detailed here. Either amplified or unamplified DNA can be used.

1. Combine 5.5 μL random primers (3 $\mu\text{L}/\mu\text{L}$) and 1 μg target DNA in a 0.2 mL PCR tube and bring to 35 μL with nuclease-free water (*see Note 5*).
2. Mix well and incubate at 99.9 $^{\circ}\text{C}$ for 5 min, then immediately chill on ice.
3. In a separate 1.7 mL microcentrifuge tube, prepare a master mix. [For each sample: 6 μL nuclease-free water, 5 μL 10 \times reaction buffer (included with enzyme), 2.5 μL 5 mM dNTP mix (2.5 mM dTTP), 1 μL klenow (40 U/ μL), and 0.5 μL CyDye (25 nM, Cy3-dUTP)] (*see Notes 5 and 6*).
4. Add 15 μL of the master mix to primer/target DNA tube and mix well.
5. Incubate the reaction at 37 $^{\circ}\text{C}$ for 6 h, heat-inactivate the enzyme at 95 $^{\circ}\text{C}$ for 3 min and then cool to 4 $^{\circ}\text{C}$.
6. Purify the labeled DNA with a Qiagen QIAquick Kit as described by the manufacturer.
7. Elute the DNA using 100 μL H₂O and check the CyDye incorporation using NanoDrop (*see Note 7*).
8. Dry the labeled DNA using a Speed Vac at Vacuum Level 5.1 for 2 h at 45 $^{\circ}\text{C}$.

3.3 Hybridization

There are currently two versions of GeoChip 5.0. The smaller array has $\sim 60,000$ probes (60K) and is more for general microbial ecology studies ($\sim 60\text{K}$ probes) and was designed to cover the core biogeochemical cycles (C, N, S, and P) as well as degradation genes for the more common contaminants (such as BTEX) and metals and antibiotic resistance genes as long as they changed the metal or antibiotic (oxidation, reduction, degradation) and not just acted as

a pump. The larger version has ~180,000 probes (180K) and covers all the genes on the smaller array plus more contaminant degradation genes and includes genes for non-transformative metal resistance mechanisms (e.g., pumps) as well as general metal homeostasis genes, and several other categories (viral and protist genes, stress response, etc.). A complete list of genes covered by both GeoChip 5.0 versions can be found by clicking the GeoChip summary links at <http://ieg.ou.edu/entrance.html>.

GeoChip 5.0 uses an Agilent format and the following protocol is based on one of Agilent's hybridization protocols [20] but has some modifications. The 60K array has eight arrays per slide and there are four 180K arrays per slide; because of this, the 60K and 180K arrays use different volumes of hybridization buffer. All quantities below are for the 60K array with the volumes for the 180K in parentheses.

3.3.1 Sample Preparation

1. Prepare hybridization buffer (*see Note 8*). (Per sample: 27.5 μL 2 \times HI-RPM hybridization buffer, 5.5 μL blocking agent, 2.4 μL Cot-1 DNA, 1.1 μL universal standard [13], and 5.5 μL formamide.)
2. Add 13 μL nuclease-free water to the labeled DNA.
3. Add 42 μL of the hybridization buffer to the DNA and mix well by pipetting up and down and then spin to collect liquid in the bottom of the tube.
4. Heat samples at 95 $^{\circ}\text{C}$ for 3 min, then immediately transfer samples to 37 $^{\circ}\text{C}$ and incubate for another 30 min (*see Note 9*).
5. Centrifuge samples for 1 min at 6000 $\times g$ to collect the sample in the bottom of the tube.

3.3.2 Array Assembly

1. Place a new gasket slide (gasket side up) into the Agilent SureHyb chamber. Make sure the slide is aligned properly and is flush with the chamber base.
2. Slowly pipette 48 μL of the sample into a gasket well, avoiding touching the slide and making sure the liquid does not touch the gasket (*see Note 10*). Repeat with next sample until all gasket wells have been filled.
3. Place the microarray onto the gasket slide, making sure the active side (the text Agilent is printed on the active side) is down and that the array slide is properly aligned with the gasket slide.
4. Place the chamber cover on the slides, slide the clamp into place, and then firmly tighten the clamp (*see Note 11*).
5. Lift the assembled chamber and rotate to wet the slides and confirm that the air bubble moves freely and that there are no small bubbles that may inhibit mixing (*see Note 12*).

3.3.3 Hybridization Protocol

1. Put the slide chambers in the hybridization oven's rotator rack (*see Note 13*). Use empty slide chambers to keep the rotator balanced, if necessary.
2. Hybridize arrays at 67 °C for 24 h with a rotation speed of 20 rpm.

3.3.4 Washing

1. Prepare the wash buffers in three separate wash dishes (*see Notes 14 and 15*).
 - (a) Wash 1—The Wash Buffer 1 (WB1) dish should be placed on the benchtop and maintained at room temperature.
 - (b) Wash 2—A second room temperature WB1 dish should be placed on a magnetic stir plate and contain a slide rack and a stirbar. There should be sufficient WB1 to completely cover the slide rack.
 - (c) Wash 3—The third dish contains the prewarmed Wash Buffer 2 (WB2) (*see Note 16*). This dish should be placed on a magnetic stir plate with heating element to maintain the buffer at 37 °C. There should be sufficient WB2 to completely cover the slide rack. Insert a stir bar.
2. Remove one hybridization chamber from the incubator (*see Note 17*).
3. Check to determine if any bubbles formed, if there was any leakage, and if the sample is still able to rotate freely (*see Note 18*).
4. Place the chamber on a flat surface, loosen the screw and remove the clamp and chamber cover. Carefully lift one end of the array and gasket slide and then hold the sides of the array, maintaining its current orientation (array on top, gasket slide on bottom) quickly place it into Wash 1.
5. Make sure the slide is completely submerged and then using forceps to make a gentle twisting motion, pry the array slide off the gasket slide. Leave the gasket slide and quickly place it in the slide rack in Wash 2.
6. Repeat **steps 2–4** until all slides have been transferred or there are five slides in Wash 2, whichever comes first (*see Note 19*).
7. Turn on the magnetic stirrer in Wash 2, such that there is a depression on the surface of the buffer without creating a vortex. We were able to obtain this with a speed of 250–300 rpm on a VWR (Radnor, PA) model magnetic stirplate with a maximum speed of 1600 rpm. Incubate for 5 min.
8. Transfer the slide rack to Wash 3 and then turn on stirrer as described in **step 6**. Incubate for 1.5 min.
9. Slowly remove the slide rack from Wash 3. It should take about 10 s to remove. No additional spinning or drying is needed. The slides are hydrophobic and should shed the buffer.
10. Discard used buffers and replenish if additional slides need to be washed (*see Note 14*).

3.3.5 Scanning

1. For the Agilent SureScan microarray scanner, the array slides must be placed into slide holders. The array should be placed array-side up. Close the holders, making sure you hear a click.
2. Place the slide holders containing microarray slides into the scanner cassette.
3. Select the appropriate scanning protocol and check the settings (*see Note 20*). For GeoChip arrays, which use the Agilent platform, using the SureScan Microarray Scanner, scanning is done in red and green channels (lasers with excitation wavelengths at 640 and 532 nm, respectively), 3 μm resolution, 20 bit Tiff dynamic range ($>10^5$), and 100 % photomultiplier tube sensitivity for both channels.
4. Scan the slides and then extract the data using Agilent's Feature Extraction program. Select the appropriate grid template (each array design should have its own specific template file generated by the array manufacturer) and the Feature Extraction program will automatically place and optimize the grid placement. Array features are automatically selected and mean pixel intensity is scored using the program's default settings.
5. Evaluate hybridization quality of the array. The slide image can be displayed in the Feature Extraction program once scanning is complete. Examine the images to make sure positive control spots are present. GeoChip 5 contains a series of 16S rRNA gene probes across the array in an easily observable pattern (green channel) and the CORS probes should also be visible across the array (red channel) (Fig. 2). Also make sure the arrays have even hybridization and no obvious problems [areas with no positive probes, very bright or dim areas, or "flares" where spots are obscured by dust or other fluorescent contaminants (Fig. 3)]. Make sure the background is even and that none of the arrays have a obviously higher background by switching to log scale.

3.4 Data Processing and Analysis

There are a large variety of microarray analysis software available to use. Select the software that best meets your particular needs. The GeoChip microarrays use an in-house developed data analysis pipeline. The pipeline allows the user to select normalization protocols, the method to determine signal cutoff and what controls to use.

GeoChip data normalization and quality filtering is performed with multiple steps [13, 21]. As a general rule, the following protocol is followed although other settings may be used depending on the samples (*see Note 21*). First, the average signal intensity of common oligo reference standard is calculated for each array, and the maximum average value is applied to normalize the signal intensity of samples in each array. Second, the sum of the signal intensity of samples is calculated for each array, and the maximum

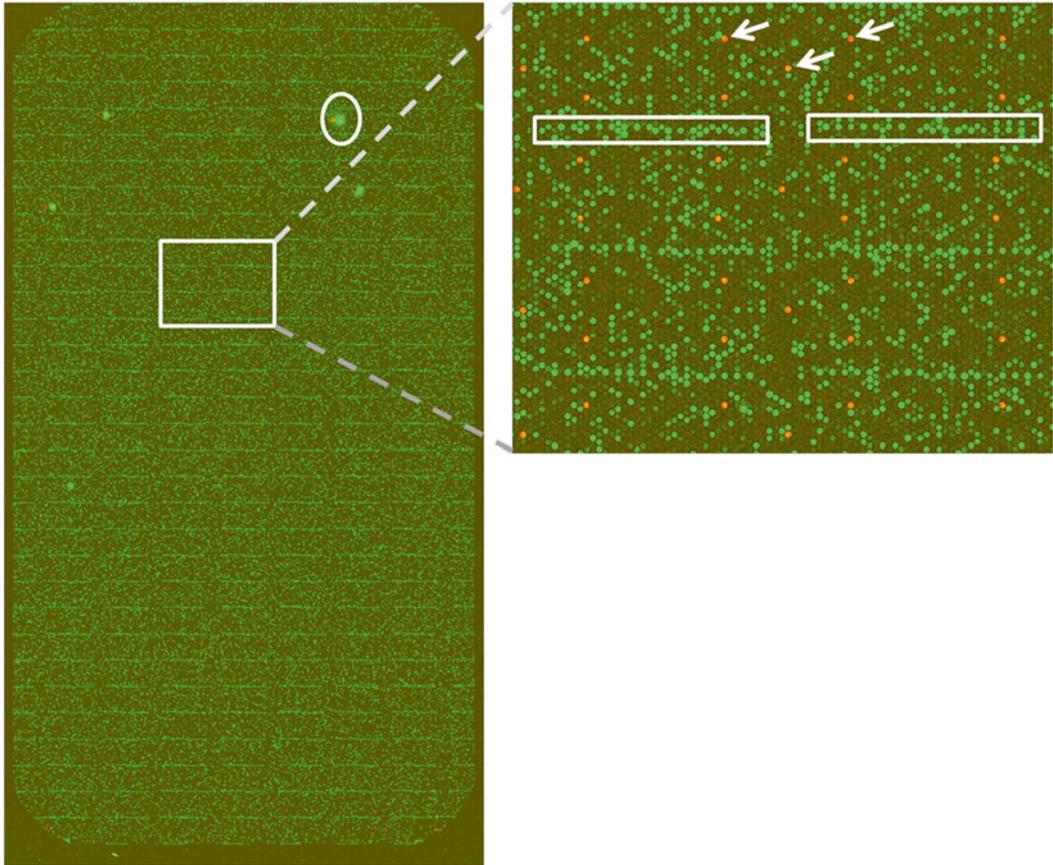


Fig. 2 GeoChip 5 microarray image. Image to the *left* is a hybridized GeoChip array. The *circle spot* is fluorescence from a piece of dust or other debris. The *boxed area* is enlarged to the *right*. The *arrows* point to a few of the CORS control probes (shown as *red spots*). The *boxed area* in the *right image* shows the 16S rRNA control probes

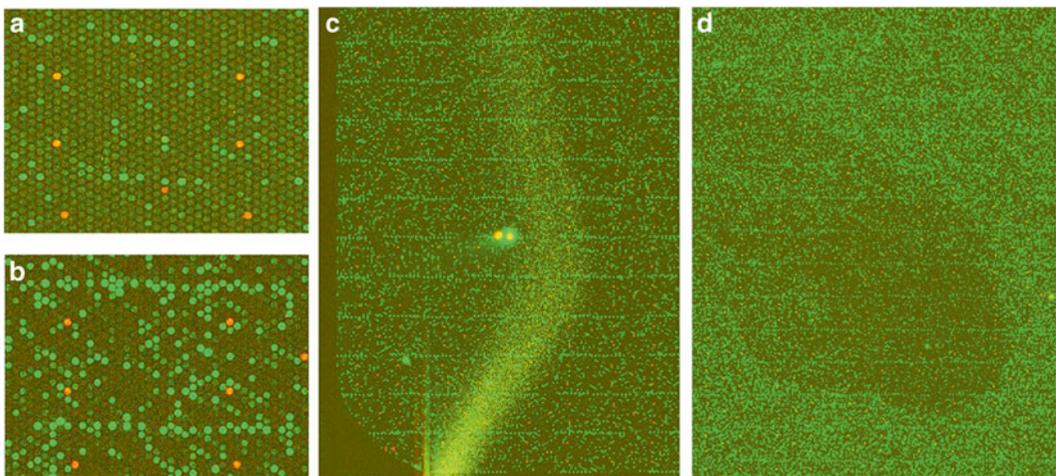


Fig. 3 Examples of poor quality hybridization or artifacts. **(a)** Low or no 16S rRNA gene probe signal and poor sample hybridization (compare with **b**). **(c)** A “flare” likely from incomplete washing, and **(d)** an area of no/poor hybridization

sum value is applied to normalize the signal intensity of all spots in an array, producing a normalized value for each spot in each array. Spots are then scored as positive and retained if the signal-to-noise ratio [$SNR = (\text{signal mean} - \text{background mean}) / \text{background standard deviation}$] is ≥ 2.0 , the coefficient of variation (CV) of the background is < 0.8 , and the signal intensity is at least 1.3 times the background. In addition, spots with signal intensities less than ~ 200 are discarded. The minimum signal value chosen should be at least twice the average background signal for a given set of samples. Spots that were detected in less than two samples (either within a replicate group or across all samples) are also removed. Before statistical analysis, logarithmic transformation is carried out for the remaining spots, and the signals of all spots are transferred into relative abundances.

Microarray data analysis can be challenging due to the large amount of data generated and its multivariate structure. The GeoChip data analysis pipeline has a variety of analysis tools for microarray data. These include relative abundance of genes or gene categories or subcategories, richness and α and β diversity of functional genes, and gene overlap between individual samples or sample groups. To look at differences between conditions, hierarchical cluster analysis, *T*-tests, analysis of variance (ANOVA), and dissimilarity tests could be used. Response ratios can be used to compare gene levels or signal intensity between conditions (e.g., treatment vs. control, contaminated vs. uncontaminated) [22]. Unconstrained ordination methods, such as principal component analysis (PCA) and correspondence analysis (CA) maximize the visible variability of data sets by reducing the dimensionality of variables. Due to the differences in the assumed data structures of ecological studies, CA may be preferred over PCA. Non-metric multidimensional scaling (NMDS), which represents the relative inter-relatedness of samples on a priori dimensions, could also be used.

Interrelationships among functional genes detected and other abiotic and biotic factors can be examined by constrained ordination, such as canonical correspondence analysis (CCA) [23], distance-based redundancy analysis (db-RDA) [24], variation partitioning analysis (VPA) [25, 26], and Mantel test. CCA is commonly used for GeoChip-based studies to better understand how environmental factors impact and drive community structure, while canonical VPA, based on the results of CCA and partial CCA, provides information on the relative influence of individual parameters on the microbial community structure. The Mantel test can be used to compare environmental factors with functional genes detected by GeoChip.

4 Notes

1. Before use, mix reagents well, then briefly spin them in a microcentrifuge or microfuge to collect all the liquid in the bottom of the tube. Repeat to make sure all reagents are well mixed.
2. Make sure the same amount of DNA is used for all samples within an experimental group.
3. Efficient amplification requires high molecular weight DNA of the highest quality obtainable.
4. No or minimal amplification is likely due to inhibitors in the sample. To overcome this, decrease the amount of sample volume used or run the sample through a serial dilution (two to three dilution steps) to reduce the concentration of inhibitor or “wash” the DNA. Inhibitors may be present even if the DNA quality is high. Additionally, the DNA can be re-precipitated (using ethanol or isopropanol protocols [27]) to reduce inhibitors. If there are still problems with the amplification, the incubation time can be increased or multiple amplification products can be combined to obtain sufficient DNA.
5. For best results, high quality, fresh reagents should be used [4].
6. CyDyes are light sensitive so the samples should be protected from the light as much as possible after the dye is added.
7. Minimum dye incorporation should be >50 pmol (pmol/ μ L \times total μ L).
8. Prepare blocking agent ahead of time by adding 1350 μ L nuclease-free water to the 10 \times aCGH Blocking Agent and keep at room temperature for at least 6 h.
9. Prior to starting, preheat two heat blocks or thermocyclers to 95 and 37 °C.
10. The Agilent arrays use bubble mixing, so the hybridization buffer will not completely fill the array chambers.
11. Make sure to tighten the clamp as firmly as possible using your hand. The clamp is designed to prevent damage to the slide so over tightening will not break the slide.
12. If small bubbles are present, firmly tap the assembly on a hard surface to dislodge the bubbles.
13. Preheat hybridization oven to 67 °C and allow it to equilibrate for at least 6 h.
14. Prior to use and after use, thoroughly rinse all glassware and stirbars used in the washing steps with copious amounts of Milli-Q water or equivalent and allow to air dry.

15. Washing should be performed in a low ozone environment such as an ozone-free hood as ozone can degrade cy-dye signal [28, 29]. The most critical time for cy-dye signal loss is during the transition from wet to dry [30], so it is important to limit ozone exposure during washing.
16. Preheat Wash Buffer 2 to 37 °C in an incubator. The Wash Buffer should be preheated in the wash dish and remain at 37 °C overnight.
17. Maintain rotation and temperature for the remaining chambers.
18. Any of these issues may result in inefficient hybridization. Make note if any of these issues occurred and continue with washing. Once the slide is scanned, hybridization quality can be assessed.
19. This setup is sufficient for five slides. If more slides need to be washed, fresh buffers should be used.
20. Other microarray scanners can be used for scanning the GeoChip as long as the scanner has at least a 3 µm resolution, can scan both the red and green channels and has at least a 16 bit dynamic range.
21. Based on results from hundreds to thousands of samples analyzed by various GeoChip versions, we anticipate about 30 % of probes to be positive for most soil samples. Other sample types, such as groundwater or bioreactors, may have fewer positive probes. For samples with more positive probes than expected, settings can be adjusted by increasing the SNR cutoff used or increasing the minimum signal required. Removing singletons (those probes detected in only one sample) can also reduce noise in the data.

Acknowledgement

Efforts for writing this Chapter were supported by the Department of Energy's Carbon Cycling program (DE-SC0004601 and DE-SC0010715).

References

1. Amann RI, Ludwig W, Schleifer KH (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev* 59:143–169
2. Furrman JA, Campbell L (1998) Marine ecology: microbial microdiversity. *Nature* 393:410–411
3. Whitman WB, Coleman DC, Wiebe WJ (1998) Prokaryotes: the unseen majority. *Proc Natl Acad Sci U S A* 95:6578–6583
4. Wu L, Thompson DK, Li G et al (2001) Development and evaluation of functional gene arrays for detection of selected genes in the environment. *Appl Environ Microbiol* 67:5780–5790

5. He Z, Gentry TJ, Schadt CW et al (2007) GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. *ISME J* 1:67–77
6. He Z, Deng Y, Van Nostrand JD et al (2010) GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity. *ISME J* 4:1167–1179
7. Tu Q, Yu H, He Z et al (2014) GeoChip 4: a functional gene arrays-based high throughput environmental technology for microbial community analysis. *Mol Ecol Resour* 14:914–928
8. He Z, Wu LY, Li XY et al (2005) Empirical establishment of oligonucleotide probe design criteria. *Appl Environ Microbiol* 71:3753–3760
9. Liebich J, Schadt CW, Chong SC et al (2006) Improvement of oligonucleotide probe design criteria for functional gene microarrays in environmental applications. *Appl Environ Microbiol* 72:1688–1691
10. Li X, He Z, Zhou J (2005) Selection of optimal oligonucleotide probes for microarrays using multiple criteria, global alignment and parameter estimation. *Nucleic Acids Res* 33: 6114–6123
11. Ning J, Liebich J, Kästner M et al (2009) Different influences of DNA purity indices and quantity on PCR-based DGGE and functional gene microarray in soil microbial community study. *Appl Microbiol Biotechnol* 82: 983–993
12. NanoDrop (2007) 260/280 and 260/230 Ratios NanoDrop® ND-1000 and ND-8000 8-Sample Spectrophotometers. Technical Support Bulletin T009
13. Liang Y, He Z, Wu L et al (2010) Development of a common oligonucleotide reference standard (CORS) for microarray data normalization and comparison across different microbial communities. *Appl Environ Microbiol* 76:1088–1094
14. Xie J, Wu L, Van Nostrand JD et al (2012) Improvements on environmental DNA extraction and purification procedures for metagenomic analysis. *J Cent South Univ* 19:3055–3063
15. Wu L, Liu X, Schadt CW et al (2006) Microarray-based analysis of submicrogram quantities of microbial community DNAs by using whole-community genome amplification. *Appl Environ Microbiol* 72: 4931–4941
16. Gao H, Yang ZK, Gentry TJ et al (2007) Microarray-based analysis of microbial community RNAs by whole-community RNA amplification. *Appl Environ Microbiol* 73:563–571
17. Zhang K, Martiny AC, Reppas NB et al (2006) Sequencing genomes from single cells by polymerase cloning. *Nat Biotechnol* 24:680–686
18. Khan AU, Mei YH, Wilson T (1992) A proposed function for spermine and spermidine: protection of replicating DNA against damage by singlet oxygen. *Proc Natl Acad Sci U S A* 89:11426–11427
19. Marceau AH (2012) Functions of single-strand DNA-binding proteins in DNA replication, recombination, and repair. *Methods Mol Biol* 922:1–21
20. Agilent (2012) Agilent Oligonucleotide Array-Based CGH for Genomic DNA Analysis. Version 3.4, July 2012. Agilent Technologies
21. Deng Y, He Z (2014) Microarray data analysis. In: He Z (ed) *Microarrays: current technology, innovations and applications*. Caister Academic Press, Norwich, UK, <http://www.horizonpress.com/microarrays2>
22. Luo Y, Hui D, Zhang D (2006) Elevated CO₂ stimulates net accumulations of carbon and nitrogen in land ecosystems: a meta-analysis. *Ecology* 87:53–63
23. ter Braak CJF (1986) Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67:1167–1179
24. Legendre P, Anderson MJ (1999) Distance-based redundancy analysis: testing multi-species responses in multi-factorial ecological experiments. *Ecol Monogr* 69:1–24
25. Økland RH, Eilertsen O (1994) Canonical correspondence analysis with variation partitioning: some comments and an application. *J Veg Sci* 5:117–126
26. Ramette A, Tiedje JM (2007) Multiscale responses of microbial life in spatial distance and environmental heterogeneity in a patchy ecosystem. *Proc Natl Acad Sci U S A* 104:2761–2766
27. Sambrook J, Russell DW (2001) *Molecular cloning a laboratory manual*, vol 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
28. Fare TL, Coffey EM, Dai H (2003) Effects of atmospheric ozone on microarray data quality. *Anal Chem* 75:4672–4675
29. Branham WS, Melvin CD, Han T et al (2007) Elimination of laboratory ozone leads to a dramatic improvement in the reproducibility of microarray gene expression measurements. *BMC Biotechnol* 7:8
30. Byerly S, Sundin K, Raja R et al (2009) Effects of ozone exposure during microarray posthybridization washes and scanning. *J Mol Diagn* 11:590–597