

1 **Oligonucleotide Fingerprinting of Ribosomal RNA Genes (OFRG)**

2

3 LEA VALINSKY<sup>1</sup>, ALEXANDRA J. SCUPHAM<sup>1</sup>, GIANLUCA DELLA VEDOVA<sup>4</sup>, ZHENG  
4 LIU<sup>2</sup>, ANDRES FIGUEROA<sup>2</sup>, KATECHAN JAMPACHAISRI<sup>3</sup>, BEI YIN<sup>1</sup>, ELIZABETH  
5 BENT<sup>1</sup>, ROBERT MANCINI-JONES<sup>1</sup>, JAMES PRESS<sup>3</sup>, TAO JIANG<sup>2</sup>, and JAMES  
6 BORNEMAN<sup>1</sup>

7

8 *Departments of Plant Pathology<sup>1</sup>, Computer Science<sup>2</sup>, and Statistics<sup>3</sup>, University of California, Riverside, CA 92521 USA, Dipartimento di*  
9 *Statistica<sup>4</sup>, Università degli Studi di Milano-Bicocca, via Bicocca degli Arcimboldi, 8, I-20126, Milano, Italy.*

10

11 **Introduction**

12

13 Microorganisms are crucial components of ecosystems and human civilization. They play vital  
14 roles in processes such as soil formation, nutrient cycling, detoxification of contaminated  
15 environments, and the transformation of waste materials into useful commodities such as  
16 compost [1, 16]. In the plant rhizosphere, they help decompose organic matter providing  
17 essential nutrients required for plant growth [1, 8]. For biotechnology, microorganisms contain a  
18 considerable resource of potentially useful compounds and enzymes, and provide vehicles for the  
19 production of genetically engineered products [5, 7]. Despite these gains in basic knowledge and  
20 biotechnology, the microbial contribution to ecosystem functioning and their potential for society  
21 has yet to be fully realized, as thorough assessments of microbial diversity have been hindered  
22 by technological limitations.

23

1           Oligonucleotide fingerprinting of ribosomal RNA genes (OFRG) is an array-based  
2 method that enables extensive analysis of microbial community composition by sorting rRNA  
3 gene clones into taxonomic clusters [17, 18]. Clone libraries are constructed using PCR primers  
4 designed to selectively amplify ribosomal RNA genes from a specific taxonomic group. The  
5 cloned rDNA fragments are arrayed on nylon membranes and then subjected to a series of  
6 hybridization experiments, each using a single DNA oligonucleotide probe. For every  
7 hybridization experiment, the signal intensities are transformed into three discrete values 0, 1,  
8 and N, where 0 and 1 respectively specify negative and positive hybridization events and N  
9 designates an uncertain assignment. This process creates a hybridization fingerprint for each  
10 clone, which is a vector of values resulting from its hybridizations with all probes. The clones are  
11 identified by clustering their hybridization fingerprints with those of known sequences and by  
12 nucleotide sequence analyses of representative clones within a cluster. Alternative methods for  
13 both transforming the hybridization data and clustering the hybridization fingerprints have also  
14 been recently developed (see Tools for Data Analysis section).

15

## 16 **Procedures**

17

18 The OFRG process can be divided into six parts: (1) DNA extraction, (2) PCR of rRNA  
19 genes, (3) Clone library construction, (4) Array construction, (5) Array hybridization, and  
20 (6) Data analysis.

21

### 22 *DNA extraction*

23

1 Extract DNA from the sample of interest using an appropriate method. For soil, the Fast  
2 DNA Spin Kit for Soil (Qbiogene, Carlsbad, CA, USA) or UltraClean Soil DNA Kit (Mo  
3 Bio Laboratories, Carlsbad, CA, USA) work well. The inclusion of  $\text{AlNH}_4(\text{SO}_4)_2$  in the  
4 extraction process may be useful to remove PCR inhibitors from the DNA [4]. Extract  
5 DNA from replicate samples, pool an aliquot from each sample, and reduce the volume  
6 of this pooled sample using a concentrator system such as the Savant Speed Vac  
7 System. Purify the DNA by subjecting it to electrophoresis in an agarose gel, excising  
8 the portion of the gel containing the larger fragments and isolating the DNA using a kit  
9 such as the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA, USA). PCR amplify  
10 rRNA genes from this purified DNA as follows.

11

12 *PCR amplification of rRNA genes*

13

14 *PCR optimization.* Several factors need to be considered when optimizing PCR  
15 conditions. (i) Template concentration may affect library composition [6]. (ii) Some DNA  
16 samples contain substances that inhibit amplification. Gel isolation is an effective way to  
17 purify DNA for PCR amplification. Diluting the template or adding BSA or other  
18 components to the amplification reaction can also alleviate the inhibition [11, 14]. (iii)  
19 The optimum number of PCR cycles is the fewest possible. We define this to be the  
20 fewest number of cycles that produces a DNA band that is barely visible when 5  $\mu\text{l}$  of an  
21 amplification reaction is resolved on a 1% agarose gel and stained with ethidium  
22 bromide. (iv) Zhou et al. present useful guidelines for reducing the number of chimeras  
23 and heteroduplexes [13].

1  
2 *PCR for library construction.* Carry out 200 µl amplification reactions for each sample  
3 using the optimized PCR conditions determined above. PCR reagents and cycling  
4 parameters can vary. We typically perform 200 µl reactions that contain 50 mM Tris-HCl  
5 (pH 8.3), 2.5 mM MgCl<sub>2</sub>, 500 µg/ml BSA, 250 µM of each dNTP, 400 nM each primer  
6 (HPLC purified), and 10 U AmpliTaq DNA Polymerase (Applied Biosystems, Foster City,  
7 CA, USA). The bacterial rRNA gene-selective primers are: 27fOFRG (forward),  
8 AGRRTTTGATYHTGGYTCAG and 1492rOFRG (reverse), GBTACCTTGTTACGACTT,  
9 which are modified versions of 27f and 1492r [12]. Given the high number of degenerate  
10 positions in the 27fOFRG primer, amplification efficiency may be enhanced by  
11 increasing its concentration to 4 µM. The fungal rDNA primers are: nu-SSU-0817-5'  
12 (forward), TTAGCATGGAATAARRAATAGGA and nu-SSU-1536-3' (reverse),  
13 ATTGCAATGCYCTATCCCCA [3]. Cycling parameters for bacterial rRNA genes are:  
14 94°C for 5 min; followed by x cycles at 94°C for 30 sec, 48°C for 30 sec, 72°C for 4 min;  
15 followed by a 2 min incubation at 72°C. Cycling parameters for fungal rDNA are: 94°C  
16 for 5 min; followed by x cycles at 94°C for 30 sec, 56°C for 30 sec, 72°C for 2 min;  
17 followed by a 2 min incubation at 72°C. The number of cycles (x) is determined by  
18 optimization experiments described in the previous section. (After these reactions have  
19 been completed, an optional step can be performed to reduce the number of  
20 heteroduplexes [15].) Use a concentrator system such as the Savant Speed Vac  
21 System to reduce the volume of the amplification reactions to 5-10 µl. Resolve DNA on  
22 a 1% agarose gel. Excise the bands containing the rRNA gene fragments and isolate  
23 the DNA using a kit such as the Mini Elute Gel Extraction Kit (Qiagen).

1

2 *rRNA gene clone library construction*

3

4 Clone rDNA fragments using a T-vector system such as pGEM-T Vector System  
5 (Promega, Madison, WI, USA) or other suitable vector following the instructions  
6 provided by the manufacturer. Alternatively, preliminary experiments in our lab suggest  
7 that the USER Friendly Cloning Kit (New England Biolabs, Beverly, MA) may be a  
8 superior cloning system. It provides directional cloning and it appears to enhance  
9 cloning efficiency. Note: if this system is used, additional nucleotides will need to be  
10 added to the 5' ends of the PCR primers (rDNA-selective) as described by the  
11 manufacturer; in addition, new primers will need to be developed that specifically anneal  
12 to the multiple cloning site of this vector for amplifying the rRNA genes for array printing.

13

14 *Array construction*

15

16 These instructions are for constructing arrays composed of 1536 rDNA clones. Note:  
17 maintaining the orientation of the 384 well culture and PCR plates is very important.  
18 Number the upper right corner of the plates (1-4). This helps keep the plates in a  
19 uniform orientation throughout the procedure.

20

21 *Colony selection and transfer.* Transfer white colonies (those containing rRNA gene  
22 inserts) into 384-well plates containing Luria-Bertani (LB) media (100 µg/ml ampicillin)  
23 using toothpicks or another suitable transfer strategy. Each well should contain 30 µl of

1 medium except for the perimeter wells, which should contain 60  $\mu$ l because the medium  
2 in these wells evaporates more quickly. After inoculating a well, leave the toothpick  
3 temporarily in the well. This directs the placement of the next inoculation. Do not leave  
4 toothpicks in the wells for very long, because they absorb the medium. Toothpicks can  
5 be reused after autoclaving. For two of the 384-well plates (plates 2 and 4 work best),  
6 inoculate the column on the right side of the plate with positive and negative array  
7 controls, which are clones that have defined nucleotide sequences. Note: for each  
8 probe, there must be at least one control clone that will hybridize (positive control) and  
9 one that does not hybridize (negative control). See the Data Analysis section for an  
10 explanation of how these controls are used.

11

12 *Clone library growth.* Place the plates in an unsealed plastic bag with the open edge  
13 folded under to prevent excessive drying. Grow the bacteria in this bag for 16-18 hours  
14 at 37°C with shaking at 300 rpm. After this growth period, these bacteria are used for  
15 the subsequent PCR step and are also stored for future use as described below.

16

17 *PCR.* PCR is performed by inoculating the reaction mix with the bacteria cultured in  
18 384-well plates (as described above), which contain rRNA gene templates. PCR can be  
19 performed using a thermocycler capable of holding 384-well plates or with water baths.

20 The water bath method has the advantage of allowing all PCR amplifications to be  
21 performed simultaneously, which can reduce variation and therefore minimize error.

22 Note: the water bath method also requires that the 384-well plates be sealed using a  
23 device such as the Combi Thermo-sealer (ABgene, Epsom, UK).

- 1
- 2 1. Clean a 384-pin, 0.5  $\mu$ l slot pin replicator (VP 384S.5) (V&P Scientific, San
- 3 Diego, CA) with V&P Scientific Pin Cleaning Solution as described by the
- 4 manufacturer: add diluted cleaning solution (1 part cleaning solution, 4 parts
- 5 distilled water) to a pipette tip box lid (or other suitable container) until 1 cm deep.
- 6 Prepare three additional wash stations in pipette tip box lids: fill two with distilled
- 7 water and one with alcohol (95-100% EtOH, Isopropanol or Methanol) until 1 cm
- 8 deep. (i) Submerge replicator pins in cleaning solution. (ii) Blot on
- 9 chromatography paper (0.35 mm thick) or Lint-Free Blotting Paper (V&P
- 10 Scientific). (iii) Submerge replicator pins in distilled water. (iv) Blot on
- 11 chromatography paper (0.35 mm thick) or Lint-Free Blotting Paper. (v) Repeat
- 12 steps iii and iv using clean water. (vi) Submerge replicator pins in alcohol. (vii)
- 13 Dry the pins by using a hair drier or by flaming the alcohol.
- 14 2. Make reaction mix for four 384-well plates and store on ice. Combine 41 ml H<sub>2</sub>O,
- 15 6 ml 0.5 M Tris pH 8.3, 6 ml 25 mM MgCl<sub>2</sub>, 3 ml 10  $\mu$ g/ml BSA, 3 ml dNTPs mix
- 16 containing 5mM each dNTP, 240  $\mu$ l 100  $\mu$ M primer T725,
- 17 GGCCCGACGTCGCATGCTC (HPLC purified), 240  $\mu$ l 100  $\mu$ M primer SP650,
- 18 TGGTCGACCTGCAGGCGGC (HPLC purified), and 3000 Units *Taq* DNA
- 19 Polymerase. Note: these primers anneal to the multiple cloning site of pGEM-T. If
- 20 another vector is used for cloning, different PCR primers will be needed.
- 21 3. Place four 384-well PCR plates on ice and add 35  $\mu$ l reaction mix to each well.
- 22 To avoid bubbles at the bottom of the wells, which can lead to the displacement
- 23 of reaction mix when inoculated with the bacteria (Step 5), fill the wells by placing

- 1 the pipette tips at the bottom of the wells and then pull the pipette tips out of the  
2 well while simultaneously dispensing the reaction mix.
- 3 4. If using the water bath method, turn on the Combi Thermo-sealer (ABgene,  
4 Epsom, UK), which needs about 10 min to warm up.
- 5 5. Samples from the 384-well culture plates are transferred to the PCR reaction  
6 plates using the 0.5  $\mu$ l slot pin replicator. Inoculate the PCR plates with the  
7 freshly grown bacteria described above. Pay attention to the plate numbers and  
8 orientation. Inoculate 0.5  $\mu$ l of each bacterial culture 2 to 4 times such that a total  
9 of 1 to 2  $\mu$ l is transferred to each well. Note: larger amounts (up to 10% of total  
10 PCR volume) of bacteria may generate more PCR product, which typically  
11 enhances the hybridization signal intensities and the overall performance of the  
12 OFRG process. V&P Scientific will construct pin replicators that deliver a wide  
13 range of volumes. It may be advantageous to use a 1 or 2  $\mu$ l slot pin replicator  
14 such that only one transfer is needed. If several inoculations are being performed  
15 with a 0.5  $\mu$ l replicator to transfer a total of 1-2  $\mu$ l, dry the pins between  
16 inoculations with alcohol and flaming to ensure accurate and consistent volume  
17 transfers.
- 18 6. To prevent carry over from one 384-well plate to another, the replicator needs to  
19 be cleaned between inoculating each different plate. Prepare three wash stations  
20 in pipette tip box lids. The first station should contain 10% bleach, the second  
21 distilled water, and the third alcohol (95-100% EtOH, Isopropanol or Methanol).  
22 Blot the pins on chromatography paper (0.35 mm thick) or Lint-Free Blotting  
23 Paper. Successively dip the pins in each wash station while blotting them on

1 chromatography paper (0.35 mm thick) or Lint-Free Blotting Paper between each  
2 wash. Each successive station should contain slightly more solution to ensure  
3 removal of the previous solution. After dipping the pins in alcohol, dry them using  
4 a hair dryer or by flaming the alcohol.

5 7. If using the water bath method, seal the PCR plates using a Combi Thermo-  
6 sealer (ABgene, Epsom, UK) and Thermo-Seal foil (Marsh Bio Products,  
7 Rochester, NY, USA) glossy side up. Press and hold the sealer on the plate for  
8 exactly 3 sec, turn the plate 180°, press and hold the sealer on the plate for an  
9 additional 1 sec. The plates will not seal properly if the top surface of the plate is  
10 wet, which could occur for several reasons. First, care must be taken when filling  
11 the wells with the PCR reaction mix. Second, if bubbles were introduced when  
12 filling the wells, this mixture may overflow from the wells when the mix is  
13 inoculated with the bacteria. If using a conventional thermocycler, use an  
14 appropriate sealing material.

15 8. Perform thermocycling: 10 min at 94°C; 35 cycles of 1 min at 94°C then 2 min at  
16 72°C; 5 min at 72°C. Store PCR products at 4°C.

17  
18 *Clone library storage.* After the amplification reactions have been inoculated, the  
19 bacteria are stored for future use. This can be done in three ways. (1) The original set of  
20 culture plates can also be stored at -70 or -80°C after adding an equal volume of LB  
21 medium containing 30% glycerol. (2) Replica plates can be made by filling four 384-well  
22 plates with LB medium plus ampicillin (100 µg/ml). Each well should contain 30 µl LB  
23 except for the perimeter wells, which should contain 60 µl because the medium in these

1 wells evaporates more quickly. Inoculate these plates with the bacteria using a 0.5  $\mu$ l  
2 slot pin replicator. The plates should be placed in an unsealed plastic bag with the open  
3 edge folded under to prevent excessive drying. Grow for 16-18 hours at 37°C with  
4 shaking at 300 rpm. Add 30  $\mu$ l LB medium containing 30% glycerol to each well, briefly  
5 mix by shaking at 300 rpm, and store at -70–80°C. (3) The original set of culture plates  
6 can also be stored at -70–80°C without adding LB-glycerol. The bacteria stored by any  
7 of these methods can be thawed and used as inoculum for subsequent PCR reactions,  
8 if needed. Additional information on the storage and growth conditions of bacteria used  
9 as inoculum for PCR has been described previously [9].

10

11 *Array printing.* The arrays are printed by applying the PCR amplified rRNA gene  
12 fragments to nylon membranes using a 0.5  $\mu$ l slot pin replicator (VP 384S.5) with a  
13 Multi-Print registration device (VP 382) (V&P Scientific, San Diego, CA) or equivalent  
14 printing device.

15

- 16 1. After the thermocycling has been completed, unseal the PCR plates. To prevent  
17 drying, open the plates only on the day the membranes will be printed.
- 18 2. To ensure the amplification reactions worked properly, resolve 1  $\mu$ l each from 8  
19 random wells from each plate on a 1% agarose gel, stain the gel with ethidium  
20 bromide and view the results. (After collecting the 1  $\mu$ l samples from the 384-well  
21 plates, cover the plates with a lid or PCR plate sealing tape to prevent drying  
22 during this gel electrophoresis step.) If the PCR products are the predicted size  
23 and are of sufficient concentration, proceed with printing. Although the exact

1 concentration of PCR product needed to produce consistent hybridization signals  
2 has not been precisely determined, low concentrations of DNA will result in  
3 undetectable hybridization signals while high concentrations could result in large  
4 and noncircular spots if the binding capacity of the membrane was exceeded.  
5 We currently attempt to print PCR products that have a concentration between  
6 10 and 100 ng/ $\mu$ l.

- 7 3. Clean the 0.5  $\mu$ l slot pin replicator as described in Step 1 from the PCR  
8 subsection of the Array Construction section.
- 9 4. Cut Hybond N+ nylon membrane (Amersham Biosciences) into 10 x 13.5 cm  
10 pieces using scissors.
- 11 5. Find a stable, flat and relatively low surface for array printing.
- 12 6. Arrange membranes on two layers of chromatography paper (0.35 mm thick) on  
13 this surface.
- 14 7. Arrange Multi-Print registration device (V&P Scientific) on membranes in the  
15 orientation that allows four offset printing positions. Use double-sided tape to  
16 adhere the registration devices to the membrane.
- 17 8. The arrays are made by printing the PCR products from four 384-well plates on  
18 one membrane using the Multi-Print registration device. We recommend  
19 performing this operation in the following manner: plate 1 (upper left position),  
20 plate 2 (upper right position), plate 3 (lower left position), and plate 4 (lower right  
21 position).
- 22 9. Using the 0.5  $\mu$ l slot pin replicator, print PCR products onto Hybond N+  
23 membranes. Firmly press replicator pins on the nylon membrane for 5 sec using

1 your body weight to assist with the process. Repeat this printing one more time  
2 such that 1  $\mu$ l of each PCR product is applied to the membrane. (V&P Scientific  
3 will construct pin replicators that deliver a wide range of volumes. It may be  
4 advantageous to use a 1  $\mu$ l slot pin replicator such that only one transfer is  
5 needed.) Print the PCR products from all four plates on as many membranes as  
6 possible, which is typically 25-30. Clean and dry the replicator between different  
7 PCR plates as described in Step 6 of the PCR subsection of the Array  
8 Construction section. Before the Multi-Print registration device is removed from  
9 the membrane, ensure that all four plates were printed and number each  
10 membrane with pencil in the bottom right corner for orientation and identification  
11 purposes. Trim membranes, leaving approximately a 0.5 to 1 cm border around  
12 the printed area. It may also be helpful to notch one corner to assist with  
13 orientation for subsequent steps. The printed membranes are now called arrays.  
14 (If the following hybridizations reveal inconsistent printing, try one the following to  
15 remedy the situation for subsequent experiments: (i) clean the pins more  
16 frequently throughout the process using the protocol described in Step 6 from the  
17 PCR subsection of the Array Construction section, (ii) use a bristle brush when  
18 washing the pins (VP 425, V&P Scientific), (iii) substitute the Lint-Free Blotting  
19 Paper (V&P Scientific) for the chromatography paper throughout this protocol.)  
20 10. UV crosslink the arrays (70 mJ/cm<sup>2</sup>).  
21 11. The arrays can be stored dry between sheets of paper for at least 2-3 months at  
22 room temperature.  
23

1 *Array hybridization*

2

3 *Array processing and hybridization.* The rRNA genes fragments on the arrays are  
4 denatured and then subjected to a series of hybridization reactions, each using a single  
5 <sup>33</sup>P-labeled DNA probe.

6

- 7 1. Preheat water bath or block heater to 37°C for kinase reactions (Step 9).
- 8 2. Place a piece of chromatography paper (0.35 mm thick), which is large enough  
9 to hold the arrays, on a layer of plastic wrap.
- 10 3. Place the arrays DNA side up on the chromatography paper (0.35 mm thick),  
11 which has been saturated with Denaturation Solution (0.5 M NaOH, 1.5 M NaCl).  
12 Incubate 5 min. Repeat.
- 13 4. Place the arrays on chromatography paper (0.35 mm thick) saturated with  
14 Neutralization Solution (50 mM NaPhosphate, pH 7.2). Incubate 3 min. Repeat  
15 two more times. For both the Denaturation and Neutralization steps, best results  
16 are obtained when air pockets between the membrane and chromatography  
17 paper are avoided and no solution puddles on the top surface of the membrane.  
18 Note: Steps 2-4 are only done on newly printed arrays that have not been  
19 hybridized.
- 20 5. Place arrays in pipette tip boxes (without the lattice that holds the tips) or other  
21 suitable containers and add 100 ml boiling 0.1% SDS. (This step increases  
22 hybridization signal intensities, presumably through further strand separation.) To

- 1           increase efficiency, two arrays can be placed in one box with their DNA sides  
2           facing opposite directions.
- 3       6. Allow the arrays in SDS solution to cool for 5-10 min.
- 4       7. Put arrays, one array per bottle, in labeled hybridization bottles (35 x 150 mm)  
5           containing 5 ml Hybridization Buffer (600 mM NaCl, 60 mM sodium citrate, 7.2%  
6           Sarkosyl).
- 7       8. Incubate for 30 min at 11°C with rotation.
- 8       9. 5'-end-label the oligonucleotide probes with <sup>33</sup>P using T4 Polynucleotide Kinase  
9           (T4 PNK) (New England Biolabs). Kinase reactions (10 µl) contain 2 µM DNA  
10          oligonucleotide, 1.5 µl gamma <sup>33</sup>P-ATP (370 MBq/ml, 111 TBq/mmol), 1 µl 10X  
11          T4 PNK buffer, 6.5 U T4 PNK and are incubated at 37°C for 30 min. Table 1  
12          shows the nucleotide sequences of each probe.
- 13      10. Add kinase reactions (labeled primers) to the hybridization bottles, one reaction  
14          per bottle, containing the Hybridization Buffer and arrays.
- 15      11. Incubate 16-18 hours at 11°C with rotation.
- 16      12. Store used hybridization solutions (Hybridization Buffer + kinase reaction) at 4°C  
17          for future use (or discard in an appropriate manner). Although most but not all  
18          hybridization solutions can be used twice, optimum results are obtained when  
19          freshly made hybridization solutions are used.
- 20      13. Wash the arrays by adding 30 ml of the appropriate SSC solution (pre-chilled to  
21          11°C) to hybridization bottles and rotate for X min at 11°C (see Table 1 for array  
22          washing conditions). Discard washing solutions appropriately. Repeat.

1 14. Remove arrays with a forceps and remove excess solution by briefly placing the  
2 arrays on chromatography paper (0.35 mm thick). This typically takes  
3 approximately 1-3 min. The arrays should remain moist, but not have any visible  
4 puddles of solution. Do not dry the arrays at this stage. If they dry, the probe will  
5 be difficult to remove, making the array unusable for further hybridization  
6 experiments.

7 15. Arrange the arrays neatly on plastic wrap. Place another layer of plastic wrap on  
8 top of the arrays to prevent them from drying. If possible, avoid wrinkles and  
9 folds in the portions of the plastic wrap that cover the membranes. Place the  
10 arrays on a phosphorimaging screen or autoradiography film in an  
11 autoradiography cassette. Expose for 5-6 hours. Note: different plastic wraps  
12 have various thicknesses, which can influence the hybridization signal intensities  
13 as thinner wraps produce higher signals.

14 16. Scan the screen with a phosphorimager at 100 dpi (or, if using film, develop the  
15 film). Print and save the electronic image in TIF format.

16 17. Expose the phosphorimaging screen (or film) to the arrays overnight.

17 18. Scan the screen (or develop the film). Print and save the electronic image in TIF  
18 format.

19 19. At this point, the arrays can be (i) stored in 1X SSC at 4°C, (ii) stripped and  
20 stored (see below) or (iii) stripped and rehybridized (see below).

21

22 *Array stripping.* Arrays can usually be used more than once. To reuse them, the probe  
23 must be removed or stripped from the array.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22

1. Place arrays in pipette tip boxes (without the lattice that holds the tips) or other suitable containers and add 100 ml boiling Stripping Buffer (0.1% SDS, 1X SSC, 200 mM Tris, pH 7.5). To increase efficiency, two arrays can be placed in one box with their DNA sides facing opposite directions. Allow solution and arrays to cool for 5-10 min.
2. Remove arrays with a forceps and place on chromatography paper (0.35 mm thick). Use a radioisotope monitor and measure counts per minute (CPM).
3. If greater than 250 CPM, repeat stripping procedure (Steps 1-2).
4. After successfully stripping the arrays, they can be rehybridized immediately or dried and stored. To rehybridize, place the arrays in hybridization bottles and either (i) perform Steps 7-19 from the Array Processing and Hybridization section or (ii) add used hybridization solution (collected from Step 12 from the Array Processing and Hybridization) and perform Steps 11-19 from the Array Processing and Hybridization section.

*Array rehybridization.* Arrays that have been stored dry should be subjected to the stripping process before they are rehybridized. This step increases hybridization signal intensities, presumably through further strand separation.

1. Place arrays in pipette tip boxes (without the lattice that holds the tips) or other suitable containers and add 100 ml boiling Stripping Buffer (0.1% SDS, 1X SSC,

- 1           200 mM Tris, pH 7.5). To increase efficiency, two arrays can be placed in one  
2           box with their DNA sides facing opposite directions.
- 3           2. Allow solution and arrays to cool for 5-10 min.
- 4           3. Remove the arrays with a forceps, place them in hybridization bottles, and either  
5           (i) perform Steps 7-19 from the Array Processing and Hybridization section or (ii)  
6           add used hybridization solution (collected from Step 12 from the Array  
7           Processing and Hybridization) and perform Steps 11-19 from the Array  
8           Processing and Hybridization section.

9

10 *Data analysis*

11

12 In the data analysis section, the signal intensity data from the hybridization experiments  
13 will be transformed into hybridization fingerprints, which are subsequently used to sort  
14 the rDNA clones into taxonomic groups. The data analysis can be performed in several  
15 ways and with numerous software packages. Below is a general description of the  
16 process followed by brief descriptions of software tools we use to facilitate these  
17 analyses.

18

19 *Hybridization signal generation and processing.* There are several ways in which the  
20 signal intensity data can be generated and processed before the hybridization  
21 fingerprints are produced. A) *No processing-* Hybridization signals are obtained by  
22 performing a single hybridization experiment with each of the probes. B) *Normalization-*  
23 (i) Hybridization signals are obtained by performing a single hybridization experiment

1 with each of the probes. (ii) These arrays are subsequently stripped and rehybridized  
2 with a reference probe, which is a probe that should hybridize to all clones. The signal  
3 intensity values from (i) are divided by those from (ii). This process should reduce error  
4 due to variations in DNA concentration, PCR amplification efficiency, etc. C) *Repetition-*  
5 *Hybridization* signals are obtained by performing multiple hybridization experiments with  
6 each of the probes. The signal intensities from each probe are then averaged. D) *Other-*  
7 *Combinations* of the above options may also be useful.

8  
9 *Hybridization fingerprints.* In OFRG, a series of hybridization experiments, each using a  
10 single DNA oligonucleotide, generates a hybridization fingerprint for each rRNA gene  
11 clone. These fingerprints specify whether the probes hybridized or did not hybridize to  
12 the clones. An example of a hybridization fingerprint created by 26 probes is  
13 000101N001000N110101111000. The fingerprints are generated by transforming the  
14 hybridization signal intensity data into three discrete values 0, 1, and N, where 0 and 1  
15 respectively specify negative and positive hybridization events and N designates an  
16 uncertain assignment. The signal intensity data from the unidentified rDNA clones are  
17 transformed into 0, 1, and N based on the signal intensities from control clones, which  
18 are clones with defined nucleotide sequences. Typical arrays contain 1504 unknown  
19 rDNA clones and 32 control clones. For most probes, the control clones expected not to  
20 hybridize with the probe (negative controls) have signal intensity values less than the  
21 control clones expected to hybridize with the probe (positive controls); conversely, the  
22 signal intensity values from the positive clones are higher than those from the negative  
23 controls. For probes that function in this manner, clones with intensity values less than

1 or equal to x are given a 0 classification, where x is the highest intensity value  
2 generated by a negative control. Clones with intensity values greater than or equal to y  
3 are given a 1 classification, where y is the lowest value generated by a positive control.  
4 All other clones are given an N classification. For some probes, not all of the control  
5 clones will perform in the predicted manner; for example, some positive control clones  
6 may have intensity values that are lower than some of the negative control values and  
7 vice versa. For probes that function in this manner, clones with intensity values less  
8 than x are given a 0 classification, where x is the lowest intensity value generated from  
9 a positive control. Clones with intensity values greater than y are given a 1  
10 classification, where y is the highest value generated by a negative control. All other  
11 clones are given an N classification. Performing this analysis with all probes for all  
12 clones creates a hybridization fingerprint for each clone.

13

14 *Taxonomic assignments.* Hybridization fingerprints from the unidentified rDNA clones  
15 are clustered with fingerprints from identified rRNA gene sequences using UPGMA  
16 (Unweighted Pair Group Method with Arithmetic Mean) from PAUP (Sinauer Associates,  
17 Sunderland, MA, USA) with the default parameters. The unidentified rRNA gene clones  
18 are taxonomically classified by their association with identified rRNA gene sequences in  
19 the UPGMA tree. That is, unidentified clones whose fingerprints cluster with those from  
20 identified rRNA gene sequences are likely to have high sequence identity with those  
21 identified sequences. Previous work has shown that the bacterial probe set described in  
22 Table 1 was capable of near species level resolution, as clones with average sequence  
23 identities of 97% were grouped into the same cluster [18]; where, a cluster was defined

1 as a group of clones with the same fingerprint. A prior study with the fungal probe set  
2 described in Table 1 showed that clones with average sequence identities of 99.2%  
3 were grouped into the same cluster [17]. If the unidentified clones do not cluster with  
4 identified rRNA gene sequences, these clones can be identified by a nucleotide  
5 sequence analysis of a representative clone within that cluster.

6  
7 *Probe design.* The OFRG probe sets (Table 1) were designed using a previously  
8 described simulated annealing algorithm [2]. Simulated annealing is a popular heuristic  
9 method for efficiently solving difficult optimization problems [10].

10  
11 *Tools for data analysis.* Software to facilitate data analysis are under development  
12 including tools for data processing, algorithms to design probes, and several strategies  
13 to generate and cluster the hybridization fingerprints. These software will be accessible  
14 by linking to the OFRG website from the homepages of James Borneman (see faculty  
15 listing at <http://www.plantpathology.ucr.edu/>), Tao Jiang (<http://www.cs.ucr.edu/~jiang/>),  
16 or James Press (<http://www.statistics.ucr.edu/press.htm>). As new data analysis tools  
17 and probe sets are developed, they will be posted on the OFRG website.

18

## 19 **Application of the method**

20

21 OFRG enables extensive analysis of microbial community composition by sorting rRNA  
22 gene clones into taxonomic groups. Thus far, it has been used to examine bacterial and  
23 fungal composition in soil [17, 18] and mouse intestines (Scupham et al., unpublished

1 results). These studies have shown that OFRG is an efficient method for identifying  
2 large numbers of rRNA gene clones. However, as with most PCR-based rRNA gene  
3 analysis methods, OFRG does not provide quantitative population data. Thus, if firm  
4 conclusions concerning population trends are required, independent and quantitative  
5 analysis should be performed to verify the results obtained by OFRG.

6

### 7 **Acknowledgements**

8

9 This work was funded in part by grants from the NSF BDI Program (J.B. and T.J.), the UC  
10 Biotechnology Research & Training program (J.B.) and the NSF ITR Program (G.D.V. and T.J).  
11 L.V. was supported by Vaddia-BARD Postdoctoral Award FI-306-00 from BARD, The United  
12 States-Israel Binational Agricultural Research and Development Fund. G.D.V. was supported by  
13 MURST grant “Bioinformatica e Ricerca Genomica.”

14

### 15 **References**

16

- 17 1. Atlas, RM, and Bartha, R (1998) Microbial Ecology Fundamental and Applications.  
18 Benjamin/Cummings.
- 19 2. Borneman, J, Chrobak, M, Della Vedova, G, Figueroa, A, and Jiang, T (2001) Probe  
20 selection algorithms with applications in the analysis of microbial communities.  
21 Bioinformatics 17(Suppl. 1):S39-48.
- 22 3. Borneman, J, and Hartin, RJ (2000) PCR primers that amplify fungal rRNA genes from  
23 environmental samples. Appl. Environ. Microbiol. 66:4356-4360.

- 1 4. Braid, MD, Daniels, LM, and Kitts, CL (2002) Removal of PCR inhibitors from soil  
2 DNA by chemical flocculation. *J. Microbiol. Methods* (in press).
- 3 5. Bull, AT, Ward, AC, and Goodfellow, M (2000) Search and discovery strategies for  
4 biotechnology: The paradigm shift. *Microbiol. Mol. Biol. Rev.* 64:573-606.
- 5 6. Chandler, DP, Fredrickson, JK, and Brockman, FJ (1997) Effect of PCR template  
6 concentration on the composition and distribution of total community 16S rDNA clone  
7 libraries. *Molecular Ecology* 6:475-482.
- 8 7. Chapela, IH (1997) Bioprospecting: Myths, Realities and Potential Impact on Sustainable  
9 Development, p. 238-256. *In* M. E. Palm and I. H. Chapela (ed.), *Mycology in*  
10 *sustainable development: Expanding concepts, vanishing borders.* Parkway Publishers,  
11 Boone, North Carolina, USA.
- 12 8. Christensen, M (1989) A view of fungal ecology. *Mycologia* 81:1-19.
- 13 9. Gaspard, R, Dharap, S, Malek, J, Qi, R, and Quackenbush, J (2001) Optimized Growth  
14 Conditions for Direct Amplification of cDNA Clone Inserts from Culture. *Biotechniques*  
15 31:35-36.
- 16 10. Kirkpatrick, S, Gelatt, CD, and Vecchi, M (1983) Optimization by simulated annealing.  
17 *Science* 220:671-680.
- 18 11. Kreader, CA (1996) Relief of amplification inhibition in PCR with bovine serum albumin  
19 or T4 gene 32 protein. *Appl. Environ. Microbiol.* 62:1102-1106.
- 20 12. Lane, DJ (1991) 16S/23S rRNA Sequencing, p. 115-175. *In* E. Stackebrandt and M.  
21 Goodfellow (ed.), *Nucleic acid techniques in bacterial systematics.* Wiley, New York.

- 1 13. Qiu, X, Wu, L, Huang, H, Mcdonel, P, Palumbo, A, Tiedje, J, and Zhou, J (2001)  
2 Evaluation of PCR-Generated Chimers, Mutations, and Heteroduplexes with 16S rDNA  
3 Gene-Based Cloning. *Appl. Environ. Microbiol.* 67:880-887.
- 4 14. Stommel, JR, Panta, GR, Levi, A, and Rowland, LJ (1996) Effects of Gelatin and BSA  
5 on the Amplification Reaction for Generating RAPD. *Biotechniques* 22:1064-1066.
- 6 15. Thompson, JR, Marcelino, LA, and Polz, MF (2002) Heteroduplexes in mixed-template  
7 amplifications: formation, consequence and elimination by 'reconditioning PCR'. *Nucleic  
8 Acids Res.* 30:2083-2088.
- 9 16. Tuomela, M, Vikman, M, Hatakka, A, and Itavaara, M (2000) Biodegradation of lignin in  
10 a compost environment: a review. *Bioresource Technol.* 72:169-183.
- 11 17. Valinsky, L, Della Vedova, G, Jiang, T, and Borneman, J (2002) Oligonucleotide  
12 fingerprinting of ribosomal RNA genes for analysis of fungal community composition.  
13 *Appl. Environ. Microbiol.* 68:5999-6004.
- 14 18. Valinsky, L, Della Vedova, G, Scupham, AJ, Alvey, S, Figueroa, A, Yin, B, Hartin, J,  
15 Chrobak, M, Crowley, DE, Jiang, T, and Borneman, J (2002) Analysis of bacterial  
16 community composition by oligonucleotide fingerprinting of rRNA genes. *Appl.  
17 Environ. Microbiol.* 68:3243-3250.

1 Table 1. Probe sequences and array washing conditions.

Bacteria				Fungi			
Probe No.	Probe Sequence	Washing Buffer (X SSC)	Washing Time (min)	Probe No.	Probe Sequence	Washing Buffer (X SSC)	Washing Time (min)
1	GTTGGGTAA	0.1	30	1	ATAGGGATAG	1	30
2	GTAACCTGCC	1	30	2	CTGGCTTCTT	1	30
3	GAAAGCCTGA	1	30	3	GTCTTTGGGT	1	30
4	AATTCGATGC	1	30	4	GATTTGTCTG	4	15
5	TTCGGATTGT	1	30	5	AGGGATCGGG	0.1	30
6	CGAAAGCGTG	4	5	6	GCTACACTGA	0.1	30
7	CGGCCAGAC	4	5	7	AAATAGCCCG	4	15
8	TTGATCCTGG	4	5	8	CGGTTCTATT	1	30
9	CACATGCAAG	1	30	9	TGATAGCTCT	1	30
10	GGTAATGGCC	1	30	10	CGCGCGCTAC	1	30
11	GGGCGCAAGC	1	30	11	GTTGGTGGAG	0.1	30
12	TGAAATGCGT	1	30	12	CTGGGTAATC	0.1	30
13	ATTCGATGCA	4	15	13	AATCAAAGTC	1	15
14	GCAAGCCTGA	4	30	14	GCCGTTCTTA	0.1	30
15	TCAGTTCGGA	1	30	15	GGCTTCTTAG	1	30
16	GAGGATGGCC	1	15	16	CAGAGCCAGC	0.1	30
17	GGGTAAAGGC	4	15	17	CAGACATAAC	1	15

18	CACCATGGGA	4	5	18	TTTGAGGCAA	4	15
19	AGCTAACGCG	4	5	19	GCACCTTACG	1	30
20	GTTGGTGAGG	0.1	30	20	CCAGACACAA	0.1	30
21	GTGAAAGCCC	1	15	21	TATGCCGACT	1	15
22	GTAAACGATG	1	15	22	CTTAACCTGC	1	15
23	ATGGCCCTTA	0.1	30	23	TTGATAGCTC	1	15
24	GAACGGGTGA	0.1	30	24	AAATTCTTGG	4	15
25	ACACCATGGG	0.1	30	25	TACTGCGAAA	1	15
26	GAAGCTAACG	4	30	26	TCAAAGTCTT	4	15
27	AAGTGGGGGA	1	15	27	GGTGAGTTTCCC	1	30
				*(ref)			
28	GCTGCTGGCA	1	15				
*(ref)							

- 1 After hybridization, the arrays are washed twice using the indicated buffers and times.
- 2 \*(ref) indicates reference probes, which have been designed to hybridize to all clones within a
- 3 specified taxon (ie. bacteria or fungi).
- 4