[SI-3]

Application of DNA Microarray Technology to Molecular Microbial Ecology

Jae-Chang Cho

Department of Environmental Sciences, Hankuk University of Foreign Studies

ABSTRACT

There are a number of ways in which environmental microbiology and microbial ecology will benefit from DNA microarray technology. These include community genome arrays, SSU rDNA arrays, environmental functional gene arrays, population biology arrays, and there are clearly more different applications of microarray technology that can be applied to relevant problems in environmental microbiology. Two types of the applications, bacterial identification chip and functional gene detection chip, will be presented. For the bacterial identification chip, a new approach employing random genome fragments that eliminates the disadvantages of traditional DNA-DNA hybridization is proposed to identify and type bacteria based on genomic DNA-DNA similarity. Bacterial genomes are fragmented randomly, and representative fragments are spotted on a glass slide and then hybridized to test genomes. Resulting hybridization profiles are used in statistical procedures to identify test strains. Second, the direct binding version of microarray with a different array design and hybridization scheme is proposed to quantify target genes in environmental samples. Reference DNA was employed to normalize variations in spot size and hybridization. The approach for designing quantitative microarrays and the inferred equation from this study provide a simple and convenient way to estimate the target gene concentration from the hybridization signal ratio.

There are number of direct applications of DNA microarrays to environmental microbiology and microbial ecology: 1) community genome arrays. Arrays constructed with genome of hundreds to thousands of environmental isolates would be used to study community composition and community dynamics of reactors, soils, sediments, water, gut, etc. The utilities of community genome arrays are equivalent to those of reverse sample genome probing (19). 2) SSU rDNA arrays. Oligonucleotide arrays constructed for different taxa could be used in community analysis studies. These could be designed in a phylogenetic frame work to survey different levels of sequence conservation, from highly conserved sequences giving broad taxonomic groupings, to hypervariable sequences giving genus (potentially species) level groupings. 3) Environmental functional gene arrays. These arrays could come in a variety

of styles. One concept would be to prepare oligonucleotide arrays for target gene expression, with genes of interest on the array. For example, oligonucleotide probes complementary to genes coding for key enzymes in all biogeochemical cycling could be arrayed. 4) Population biology arrays. Genetic diversity or polymorphisms within specific populations can be assessed with arrays. Oligonucleotides representing all open reading frames of a reference organism genome can be arrayed, then assayed against strain-level variants. Among above mentioned applications, Two types of the applications, bacterial identification chip and functional gene detection chip employing random genome fragments and non-competitive hybridization, respectively, are described below.

Bacterial identification chip based on shotgun array.

Bacterial identification methods currently used include analysis of morphological, physiological, biochemical, and genetic data. In the last two decades molecular methods, especially 16S rRNA gene sequencing, have been a reliable aid to the identification of diverse bacteria. Although the 16S rRNA method has served as a powerful tool for finding phylogenetic relationship between bacteria (23) because of its molecular clock properties and the large database for sequence comparison, the molecule is too conserved to provide good resolution at the species and subspecies levels (7, 8, 14, 17, 22). The relationship between 16S rRNA gene similarity and % DNA-DNA reassociation is a logarithmic function in which the sequence similarity within a species (>70% DNA relatedness) is expected to be > 98%, and the similarity among different species in a genus, e.g., fluorescent *Pseudmonas* is 93.3 ~ 99.9% (1, 2, 15). Considering the high sequence conservation and relative standard errors at 98% and 90% sequence similarities of 19% and 8%, respectively (13), 16S rDNA analysis results on closely related strains could be inaccurate and inconsistent with the results obtained by other methods. Since many important ecological and clinical characteristics of bacteria such as pathogenicity, competitiveness, substrate range, bioactive molecule production, vary below the species level, methods with higher resolution than 16S rDNA sequence are needed.

DNA-DNA hybridization is one method that provides more resolution than 16S rDNA sequencing, and the 70% criterion (20) has been a cornerstone for describing a bacterial species. In spite of these values, the method is not popular. Major disadvantages are i) the laborious nature of pairwise cross-hybridizations, ii) the requirement of isotope use, and iii) it is impossible to establish a central database. Here a new approach is proposed to identify and type bacteria based on genomic DNA-DNA homology that eliminates the above disadvantages. The method takes advantages of the capacity provided by microarray technology. Bacterial genomes are fragmented randomly and representative fragments are spotted on a glass slide, and then hybridized to test genomes. Resulting hybridization profiles are used in statistical procedures to identify test strains. Importantly, a database of hybridization profiles can be established.

We have developed a method based on random genome fragments and DNA microarray technology

that overcomes the disadvantages of whole genomic DNA-DNA hybridization. Reference genomes of four fluorescent *Pseudomonas* species were fragmented, and 60 to 96 genome fragments of approximately 1 kb from each strain were spotted on microarrays. Genomes from twelve well-characterized fluorescent *Pseudomonas* strains were labeled with Cy dyes and hybridized to the arrays. Cluster analysis on the hybridization profiles revealed taxonomic relationships between bacterial strains tested at species to strain level resolution, suggesting that this approach is useful for the identification of bacteria as well as determining the genetic distance among bacteria. Since arrays can contain thousands of DNA spots, a single array has the potential for broad identification capacity. In addition, the method avoids laborious cross-hybridizations and can provide an open database of hybridization profiles, which are limitations of traditional DNA-DNA hybridization.

Quantitative detection of microbial genes in environmental samples using DNA microarray.

DNA-microarrays provide a powerful tool for the parallel analysis of many genes. Most DNA-microarray studies conducted so far evaluate gene expression (12, 21, 24) by competitive hybridizations between different populations of mRNA expressed under different culture conditions. The relative extents of hybridizations of target genes to probes on the microarray provide information on the degree of expression of genes of interest. Recently, several research groups applied the DNA-microarray-based approach to other fields, such as single nucleotide polymorphism (SNP) and mutation detection (11), genetic linkage analysis and population genetics (9), comparative genomics (16), and identification of bacterial species (3, 4). However, the quantitative characteristics of microarray hybridization, other than competitive hybridization, have not been investigated. Studies on the quantitative properties and kinetics of microarray hybridization are needed to fully evaluate the potential uses of the DNA microarray method. DNA microarray technology holds promise for microbial ecology (18), for example, detecting and quantifying different gene families involved in biogeochemical cycling, biodegradation, and pathogenesis in a high-throughput manner. However, the approaches used for gene expression analysis or other previously reported applications are inappropriate for the titration of genes or DNA sequences in environmental samples, because the probe sizes (printed spots) vary and evenness of hybridization cannot be assured. Hence, interpretation of hybridization profiles obtained from one-color hybridization (similar to conventional Southern hybridization) is not accurate. On the other hand, gene expression analysis, which uses two-color competitive hybridization, is not affected by above two factors, since it uses the ratio values from the competitive hybridizations of each probe. However, DNA microarray hybridization to quantify gene amount requires direct binding of target sequence to the probe DNA rather than competitive hybridizations, and the extent of this direct binding should be normalized against the concentration variations in the probe DNA and spatial variation in the extent of hybridization (5). An alternative approach with a different microarray design and hybridization scheme is required to quantify target genes in biological samples. This need led us to develop the new format described below.

To quantify target genes in biological samples using DNA microarrays, we employed a reference DNA to normalize variations in spot size and hybridization. This method was tested using nitrate reductase (*nirS*), naphthalene dioxygenase (*nahA*), and *E. coli* O157 O-antigen biosynthesis genes as model genes, and lambda DNA as the reference DNA. We observed a good linearity between log (signal ratio) and log (DNA concentration ratio) at DNA concentrations above the method's detection limit, which was approximately 10 pg. This approach for designing quantitative microarrays and the inferred equation from this study provide a simple and convenient way to estimate the target gene concentration from the hybridization signal ratio.

REFERENCES

1. Cho, J.-C., and J. M. Tiedje. 2000. Biogeography and degree of endemicity of fluorescent Pseudomonas in soil. Appl. Environ. Microbiol. 66:5448-5456.
2. Cho, J.-C., and J. M. Tiedje. 2000. DNA relatedness of world-wide collection of fluorescent Pseudomonas genotypes, pp. 489-490, abstr. N-171., Abstracts of the 100th General Meeting of the American Society for Microbiology 2000. American Society for Microbiology, Washington, D. C.
3. Cho, J.-C., and J. M. Tiedje. 2001. Bacterial species determination from DNA-DNA hybridization using genome fragments and DNA microarrays. Appl Environ Microbiol. 67:3677-3682.
4. Cho, J.-C., and J. M. Tiedje. 2002. Microbial Identification Chip. US Patent Serial No. 60/296,982.
5. Cho, J.-C., and J. M. Tiedje. 2002. Quantitative detection of microbial genes by using DNA microarrays. Appl Environ Microbiol 68:1425-1430.
6. Cho, J.-C., D. Shooltz, and J. M. Tiedje. 2002. Electronic hybridization and signal detection for DNA microarrays. US Patent (pending). Invention Disclosure MSU-06871.
7. DeParasis, J., and D. A. Roth. 1990. Nucleic acid probes for identification of phytobacteria: identification of genus-specific 16S rRNA sequences. Phytopathology. 80:618-621.
8. Fox, G. E., J. D. Wisotzkey, and P. Jurtshuk, Jr. 1992. How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. Int. J. Syst. Bacteriol. 42:166-170
9. Gentalen, E., and M. Chee. 1999. A novel method for determining linkage between DNA sequences: hybridization to paired probe arrays. Nucleic Acids Res. 27:1485-1491
10. Gerry, N. P., N. E. Witowski, J. Day, R. P. Hammer, G. Barany, and F. Barany. 1999. Universal DNA microarray method for multiplex detection of low abundance point mutations. J Mol Biol. 292:251-262
11. Hacia, J. G., J. B. Fan, O. Ryder, L. Jin, K. Edgemon, G. Ghandour, R. A. Mayer, B. Sun, L. Hsie, C. M. Robbins, L. C. Brody, D. Wang, E. S. Lander, R. Lipshutz, S. P. Fodor, and F. S. Collins. 1999. Determination of ancestral alleles for human single-nucleotide polymorphisms using high-density oligonucleotide arrays. Nat Genet. 22:164-167
12. Harrington, C. A., C. Rosenow, and J. Retief. 2000. Monitoring gene expression using DNA

microarrays. Curr Opin Microbiol. 3:285-291

13. Keswani, J., S. Orkand, U. Premachandran, L. Mandelco, M. J. Franklin, W. B. whitman. 1996. Phylogeny and taxonomy of mesophilic Methanococcus spp. and comparison of rRNA, DNA hybridization and phenotypic method. Int. J. Syst. Bacteriol. 46:727-735
14. Martinez-Murcia, A. J., S. Benlloch, and M. D. Collins. 1992. Phylogenetic interrelationships of members of the genera Aeromonas and Plesiomonas as determined by 16S ribosomal DNA sequencing: lack of congruence with results of DNA-DNA hybridizations. Int. J. Syst. Bacteriol. 42:412-421
15. Moore, E. R. B., M. Mau, A. Arnscheidt, E. C. Böttger, R. A. Hutson, M. D. Collins, Y. van de Peer, R. de Wachter, and K. N. Timmis. 1996. The determination and comparison of the 16S rRNA gene sequences of species of the genus Pseudomonas (sensu stricto) and estimation of the natural intrageneric relationship. Syst. Appl. Microbiol. 19:478-492
16. Murray, A. E., D. Lies, G. Li, K. Nealson, J. Zhou, and J. M. Tiedje. 2001. DNA/DNA hybridization to microarrays reveals gene-specific differences between closely related microbial genomes. Proc Natl Acad Sci U S A. 98:9853-9858
17. Stackebrandt, E., and B. M. Goebel. 1994. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int. J. Syst. Bacteriol. 44:846-849
18. Tiedje, J. M., J. C. Cho, A. Murray, D. Treves, B. Xia, and J. Zhou. 2000. Soil teeming with life: new frontiers for soil science, p. 393-412. In B. M. Rees, B. C. Ball, C. D. Campbell, and C. A. Watson (ed.), Sustainable management of soil organic matter. CAB International, Oxon, UK.
19. Voordouw, G., Y. Shen, C. S. Harrington, A. J. Telang, T. R. Jack, and D. W. S. Westlake. 1993. Quantitative reverse sample genome probing of microbial communities and its application to oil field production waters. Appl Environ Microbiol. 59:4101-4114
20. Wayne, L. G., D. J. Brenner, R. R. Colwell, P. A. D. Grimont, O. Kandler, M. I. Krichevsky, and H. G. Truper. 1987. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. Int. J. Syst. Bacteriol. 37:463-464
21. Wei, Y., J. M. Lee, C. Richmond, F. R. Blattner, J. A. Rafalski, and R. A. LaRossa. 2001. High-density microarray-mediated gene expression profiling of Escherichia coli. J Bacteriol. 183:545-556
22. Weisburg, W. G., S. M. Barns, D. A. Pelletier, and D. J. Lane. 1991. 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol. 173:697-703
23. Woese, C. R. 1987. Bacterial evolution. Microbiol. Rev. 51:221-271
24. Ye, R. W., W. Tao, L. Bedzyk, T. Young, M. Chen, and L. Li. 2000. Global gene expression profiles of Bacillus subtilis grown under anaerobic conditions. J Bacteriol. 182:4458-4465