Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Biotechnology

Development of functional gene microarrays for microbial community analysis

Zhili He, Ye Deng and Jizhong Zhou

Functional gene arrays (FGAs) are a special type of microarrays containing probes for key genes involved in microbial functional processes, such as biogeochemical cycling of carbon, nitrogen, sulfur, phosphorus and metals, virulence and antibiotic resistance, biodegradation of environmental contaminants, and stress responses. FGAs have been demonstrated to be a specific, sensitive, and quantitative tool for rapid analysis of microbial communities from different habitats, such as waters, soils, extreme environments, bioreactors, and human microbiomes. In this review, we first summarize currently reported FGAs, and then focus on the FGA development. We will also discuss several key issues of FGA technology as well as challenges and directions in future FGA development.

Address
Institute for Environmental Genomics, Department of Botany and Microbiology, University of Oklahoma, Norman, OK 73019, USA

Corresponding author: He, Zhili (zhili.he@ou.edu)

Current Opinion in Biotechnology 2011, **23**:1–7

This review comes from a themed issue on
Analytical biotechnology
Edited by Wei E. Huang and Jizhong Zhou

0958-1669/$ – see front matter
Published by Elsevier Ltd.

DOI 10.1016/j.copbio.2011.11.001

Introduction

Microbes are the most diverse groups of organisms known in terms of phylogeny and functionality, and mediate almost every biological process on the planet, such as biogeochemical cycling of carbon (C), nitrogen (N), sulfur (S), phosphorus (P), and metals (e.g. iron, copper, and zinc), and biodegradation or stabilization of environmental contaminants. However, characterization, identification, and quantitative analysis of such diverse microbial communities in natural setting are of great challenge with conventional molecular biology tools [1–3]. The recent advance of metagenomic technologies such as high throughput sequencing [4,5•,6] and functional gene arrays (FGAs) [7,8••,9•,10] provides powerful high throughput tools for analyzing microbial communities.

FGAs are a special type of microarrays containing probes for key genes involved in microbial functional processes of interest, such as biogeochemical cycling of C, N, S, P and metals, antibiotic resistance, biodegradation of environmental contaminants, and stress responses. In the past decade, a variety of FGAs have been developed (Table 1), and among them, GeoChip has been the most comprehensive FGA [7,8••,9•,11–13]. Various versions of GeoChip have been developed and used to characterize microbial communities in terms of functional diversity, composition, structure and metabolic activity/capability from a variety of habitats, such as groundwater [7,14–17], soil [8••,18,19,20••], extreme environments [21•,22], bioreactor systems [23], and oil-contaminated sites [9•,10,24]. The results demonstrate that FGAs are a powerful tool to analyze microbial communities and their linkages with environmental factors and ecosystem functioning.

In this review, we first summarize currently reported FGAs, and then focus on the FGA development with GeoChip as an example. We also discuss several key issues of FGA technology and future challenges and directions in FGA development.

History of functional gene arrays

Representatives of currently reported FGAs are summarized in Table 1. The first FGA was constructed with 89 PCR-amplicon probes targeting four functional genes (*nirS*, *nirK*, *amoA*, and *pmoA*) derived from pure culture isolates and clone libraries [11]. Then various FGAs were developed with PCR amplicons or oligonucleotides to target specific functional processes, such as N cycling [25–28], mathanotroph [29,30], virulence, antibiotic resistance, and antimicrobial resistance [31–36]. Recently, two small scale FGAs for bioleaching systems [37] and diverse microbial communities [38••] were developed. As comprehensive FGAs, GeoChips constructed with 50-mer oligonucleotides have evolved several generations. The first generation of GeoChip (GeoChip 1.0) was constructed with 763 probes involved in nitrogen cycling, methane oxidation, and sulfite reduction [12] and 2402 probes involved in organic contaminant biodegradation and metal resistance [13]. Since GeoChip 1.0 only has a limited probe set, a comprehensive FGA, GeoChip 2.0 was developed with 24 243 probes from 150 gene categories [7]. Three years later, GeoChip 3.0 was developed with additional gene categories and several distinct features, such as a common oligonucleotide reference standard (CORS) for data normalization and comparison [39•], a software package for data management and future update [8••]. Based on GeoChip 3.0, the latest generation,

Please cite this article in press as: He Z, et al. Development of functional gene microarrays for microbial community analysis, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.11.001

Table 1

Summary of representatives of various functional gene arrays

Functional process (FGA)	No. functional gene families	Probe type	No. functional gene probes	Refs
N cycling and methanotroph	4 (amoA, nirS, nirK, pmoA)	Amplicons	89	[11]
N cycling	4 (amoA, nifH, nirS, nirK)	70-mer oligos	61 and 64	[25]
Antibiotic resistance	2 (tet, bla_{TEM-1})	Amplicons	18	[31]
N cycling (N fixation)	1 (*nifH*)	Amplicons	88	[26]
N cycling, methanotroph, S reduction	6 (*amoA*, *nirS*, *nirK*, *nifH*, *pmoA*, *dsrAB*)	50-mer oligos	763	[12]
Contaminant degradation, metal resistance	NS[a]	50-mer oligos	2042	[13]
N cycling (nodulation)	1 (*nodC*)	41–50 meroligos	130	[27]
Methanotroph	1 (*pmoA*)	~20-mer oligos	59 and 68	[29,30]
Virulence	2 (*invA*, *sopB*)	70-mer oligos	4	[32]
Virulence, antibiotic resistance	NS[a]	Amplicons	120	[33]
Comprehensive (GeoChip 2.0)	>150	50-mer oligos	24 243	[7]
Bioleaching	NS[a]	50-mer oligos	501	[37]
N fixation	1 (*nifH*)	~20-mer oligos	194	[28]
Virulence	>30	Oligos	791 and 2034	[35]
Virulence, antibiotic resistance (NimbleGen)	160	Oligos	1245	[34]
Antimicrobial resistance	NS[a]	Amplicons	800	[36]
Comprehensive(GeoChip 3.0)	292	50-mer oligos	27 812	[8••]
Comprehensive	NS[a]	cDNA clones	13 056	[38••]
Comprehensive (GeoChip 4.0, NimbleGen)	539[b]	50-mer oligos	120 054	[9•,10]

[a] NS: not specify.
[b] GeoChip also contains genes targeting human microbiomes with 36,062 probes in 139 functional gene families.

GeoChip 4.0 in the NimbleGen format has been developed, which not only contains functional categories from GeoChip 3.0, but also includes additional functional categories, such as genes involved in stress responses, bacterial phages and virulence. In addition, GeoChip 4.0 contains 36 062 probes derived from human microbiomes.

Current development of functional gene arrays

FGA development involves three major steps: first, gene selection, sequence retrieval and verification, second, oligonucleotide probe design and validation, and third, array construction and evaluation (Figure 1).

Selection of target genes and sequence retrieval and verification

A variety of genes can be used as functional markers targeting different microbial functional processes. Generally, genes are chosen for key enzymes or proteins with the corresponding function(s) of interest. Once a functional gene is selected, seed sequences are identified. Sequence retrieval and verification are performed generally by a pipeline [8••]. For each selected functional gene, a query of key words is first used to search public databases (e.g. GenBank Protein Database) to fetch all candidate amino acid sequences. Second, to remove unrelated sequences retrieved by key words, all candidate sequences are verified by HMMER 2.3.2 [40] with the selected seed sequences. Finally, all confirmed protein sequences are used to retrieve their corresponding nucleic acid sequences for oligonucleotide probe design (Figure 1a).

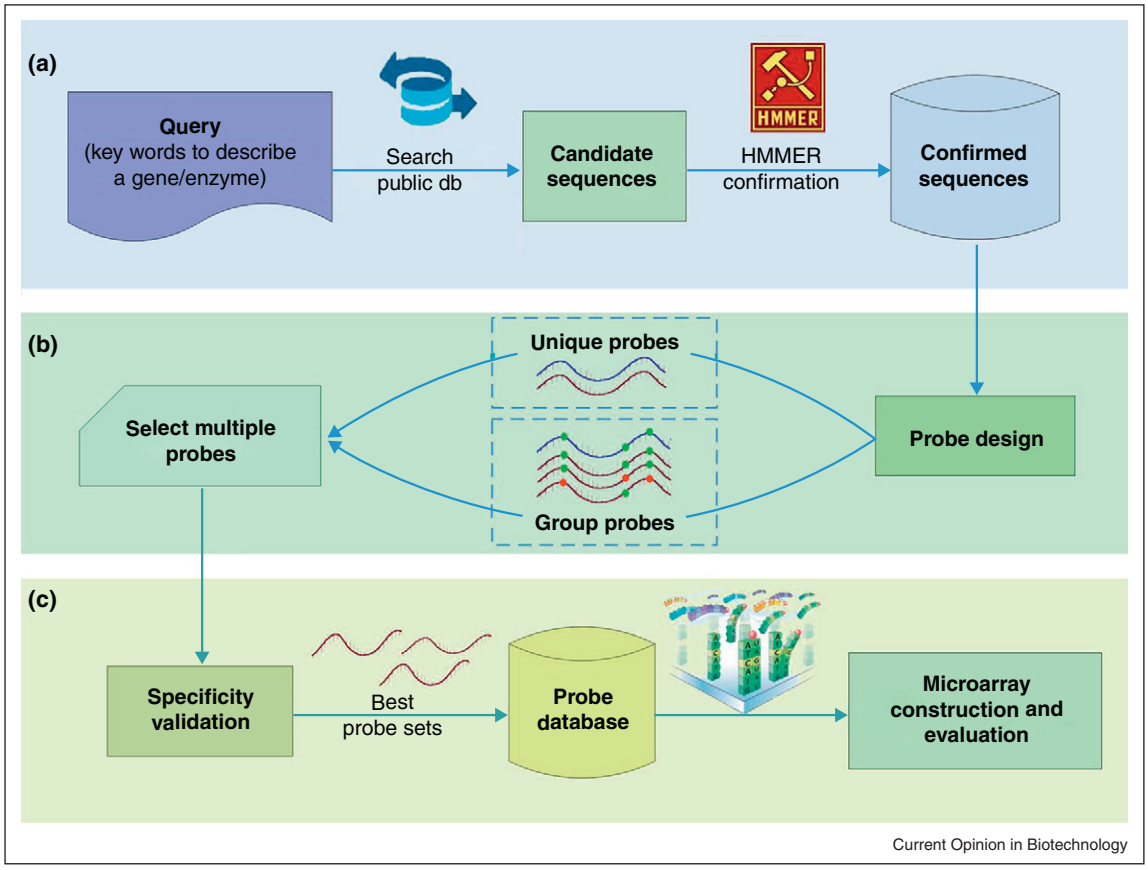
Oligonucleotide probe design and validation

A selection of specific oligonucleotide probes for FGAs faces a few challenges. First, sequences for each functional gene are generally homologs with high similarities. Second, many functional gene sequences are obtained from environmental samples by PCR amplification, and hence they are short sequences instead of full-length sequences. Third, sequences for many functional genes of interest are increasing exponentially due to high throughput sequencing approaches, such as 454 pyrosequencing and Illumina technologies. Therefore, many software tools for microarray probe design may not be suitable for FGAs. Currently, CommOligo [41] and its new versions have been considered as one of most popular tools for FGA probe design, which were used for GeoChip probe design [7,8••,9•,10]. Generally, 50-mer oligonucleotide probes that showed good specificity and sensitivity [12,13,42] are chosen for GeoChips [7,8••,9•,12,13].

CommOligo can select probes with different specificities based on different criteria: first, gene-specific probes: $\leq$90% sequence identity, $\leq$20-base continuous stretch, and ≥ -35 kcal/mol free energy with non-targets [43]; and second, group-specific probes: a group-specific probe has to meet the above requirements and must also have $\geq$96% of sequence identity, $\geq$35-base of continuous stretch, and ≤ -60 kcal/mol of free energy within the group [44]. Other oligonucleotide design software tools, such as HPD [45], ProDesign [46], and HiSpOD [47••] may be also used for FGA probe design. Since all designed probes are based on a reduced data set, their specificity must be

Please cite this article in press as: He Z, et al. Development of functional gene microarrays for microbial community analysis, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.11.001

Figure 1



Schematic presentation of functional gene array development. **(a)** Selection of target genes, sequence retrieval, and verification; **(b)** oligonucleotide probe design; **(c)** FGA construction and evaluation.

verified against large databases. Normally, multiple probes for each sequence or each group of sequences are designed so that only the best probe set is finally chosen for array construction (Figure 1b).

FGA construction and evaluation

FGAs can be constructed in-house [7,8••,11–13] or commercially [9•,10,34]. For in-house array construction, all designed probes are synthesized commercially and adjusted to a final concentration (e.g. 100 pmol/μl). Also, positive control probes (e.g. 16S rRNA gene) and negative control probes (e.g. human genes) should be synthesized and spotted on each subgrid in triplicate or more. In addition, some standards should be applied for spot and array normalization and comparison [8••,38••,39•]. All oligonucleotide probes, controls and standards are arrayed onto glass slides using an array spotter as described previously [7,8••]. For commercial array construction, different technologies are used. For example, NimbleGen (Roche NimbleGen, Madison, WI) uses micromirror array synthesis, which synthesizes all probes and internal controls directly onto the glass slide surface with different formats.

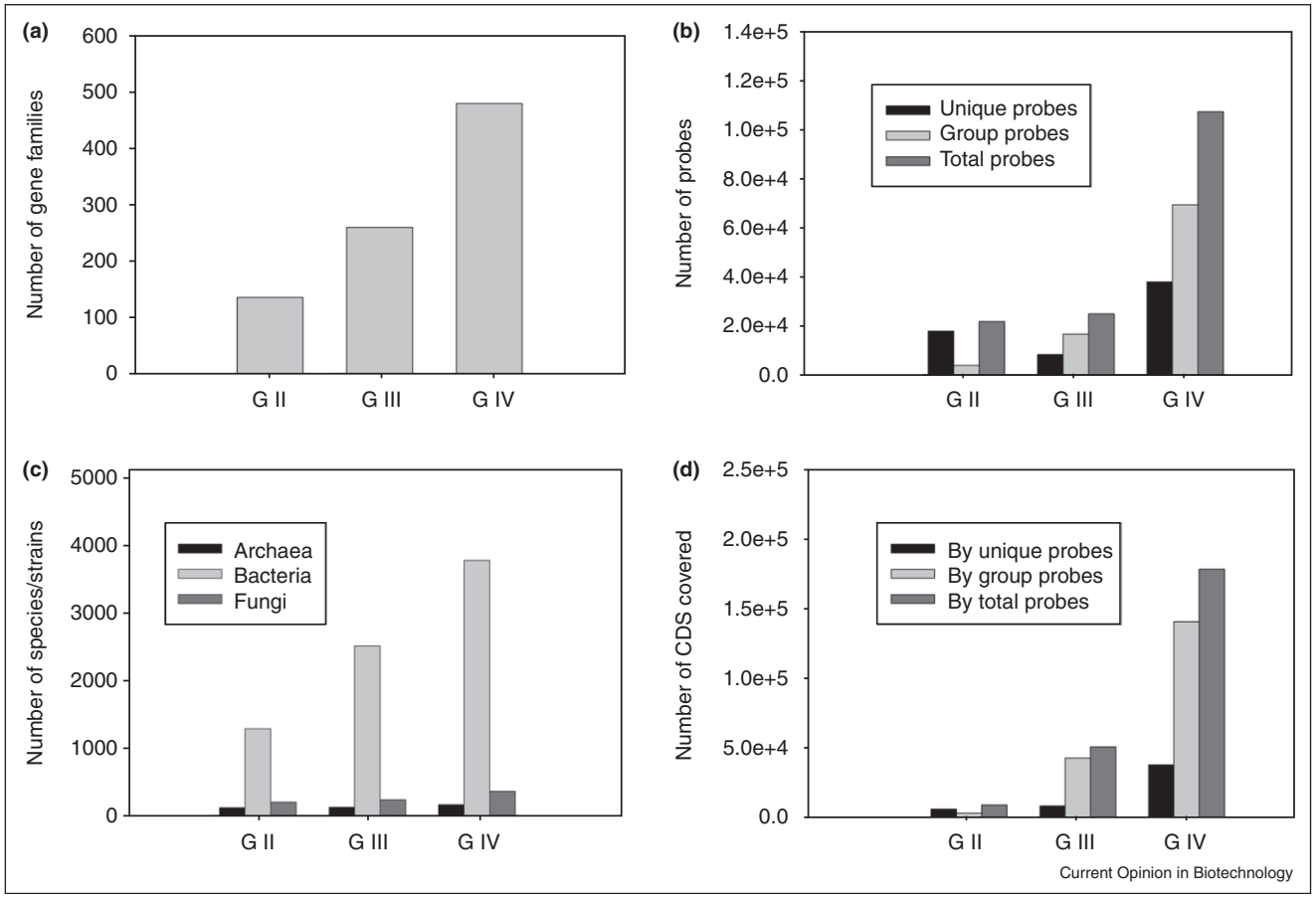
A newly developed FGA should be extensively evaluated in terms of specificity, sensitivity, and quantitative capability. First, the specificity of all designed probes is computationally evaluated in terms of maximum sequence identity, maximum stretch length, and minimum free energy [7,8••]. Second, with known targets (synthesized oligonucleotides, PCR products, and genomic DNAs), optimal hybridization conditions are identified by minimizing the number of false positives and false negatives [48], and probe specificity is further evaluated experimentally [7,8••]. Also, the array sensitivity is estimated using known targets [11–14,29,42]. Additionally, quantitative capability is assessed by quantitative real-time PCR [12–14,20••]. Finally, the application of the developed FGA is evaluated with environmental samples [7,8••,11–13].

Key issues of functional gene array technology

Although FGAs have been demonstrated to be specific, sensitive, and quantitative for profiling microbial communities, some challenges still remain, including probe

Please cite this article in press as: He Z, et al. Development of functional gene microarrays for microbial community analysis, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.11.001

Figure 2



Comparison of GeoChip 2.0 (G II), GeoChip 3.0 (G III), and GeoChip 4.0 (G IV) in terms of the number of covered functional gene families **(a)**, the number of total probes **(b)**, the number of covered species/strains **(c)**, and the number of covered coding sequences **(d)**.

coverage, specificity, sensitivity, and quantitative capability.

Coverage

Earlier FGAs are comprised a limited number of probes and cover a small number of functional groups, coding sequences (CDSs), and microorganisms, but recently developed FGAs greatly expand their coverage (Table 1). As the first comprehensive FGA, GeoChip 2.0 contains 24 243 probes for >10 000 CDSs in >150 functional gene families derived from 130, 1404, and 221 archaeal, bacterial, and fungal species/strains, respectively, while GeoChip 3.0 has 27 812 probes targeting about 57 000 CDSs in 292 functional gene families derived from 140, 2744, and 262 archaeal, bacterial, and fungal species/strains, respectively. The latest version GeoChip 4.0 is more comprehensive than GeoChip 3.0 with more than 120 000 probes targeting >200 000 CDSs in 539 functional gene families derived from 183, 4123, and 396 archaeal, bacterial, and fungal species/strains, respectively (Figure 2). A comparison of GeoChip 2.0, 3.0 and 4.0 shows that the total number of probes and covered CDSs appears to be exponentially increased with more and more CDSs covered by group-specific probes, which may be largely due to the recent development of high throughput sequencing technologies (Figure 2). Because of the limitation of microarray density, this trend will challenge the development of more comprehensive FGAs, leading to a variety of specific FGAs. Also, because of exponential increases in group-specific probes, the FGA specificity and quantitative capability may be affected.

Specificity

The most important issue with microarray technology is specificity, especially when environmental samples are analyzed with FGAs. Array specificity is controlled by probe design and hybridization conditions. Designing specific probes for sequences recovered from environmental samples with PCR amplification is challenging. The use of carefully selected design criteria can produce highly specific oligonucleotide probes [43,44]. Also, probe specificity can be improved with specific FGA probe design tools. For example, CommOligo uses global alignment algorithms and multiple criteria and filters for selection

Please cite this article in press as: He Z, et al. Development of functional gene microarrays for microbial community analysis, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.11.001

of optimal oligonucleotides, which could reduce cross-hybridization [41]. In addition, an optimization of hybridization conditions can increase specificity, which is generally controlled by temperatures and formamide concentrations [7,8••,12,43,49]. Finally, probe specificity can be improved by designing mismatch (MM) probes paired with their perfect match (PM) probes [49]. With more information on the diversity of the microbial communities of interest, it is expected that the prediction of probe specificity will be more accurate.

Sensitivity

The sensitivity for spotted oligonucleotide arrays using environmental samples is approximately 50–100 ng or 10^7 cells [12,13,29,50], or approximately 5% of the microbial community [29], providing a coverage of dominant community members. For complex microbial communities, many gene variants are expected to be in low abundances, and several strategies have been implemented to improve sensitivity. The most important approach is the amplification of DNA or RNA targets [14,36,51], which greatly improves the detection sensitivity (10 fg per reaction) [14]. Also, increasing the probe length improves sensitivity [42,52]. Another option is to reduce the volume of hybridization solution to increase target concentrations. For example, spotted FGAs generally use 40–120 μl hybridization solution, while NimbleGen arrays only need 7–30 μl. Based on our current studies, it appears that NimbleGen arrays have higher sensitivity. A total of 1–2 μg community DNA from environmental samples is enough for achieving good hybridization results. In addition, microbial populations detected by GeoChip 2.0 were significantly higher than those estimated by clone library analysis and were consistent with those by extensive sequencing analysis [4,21•]. Nevertheless, novel approaches, including, labeling, dye chemistry, slide surface structure, image processing, and data analysis are needed for further improving FGA sensitivity.

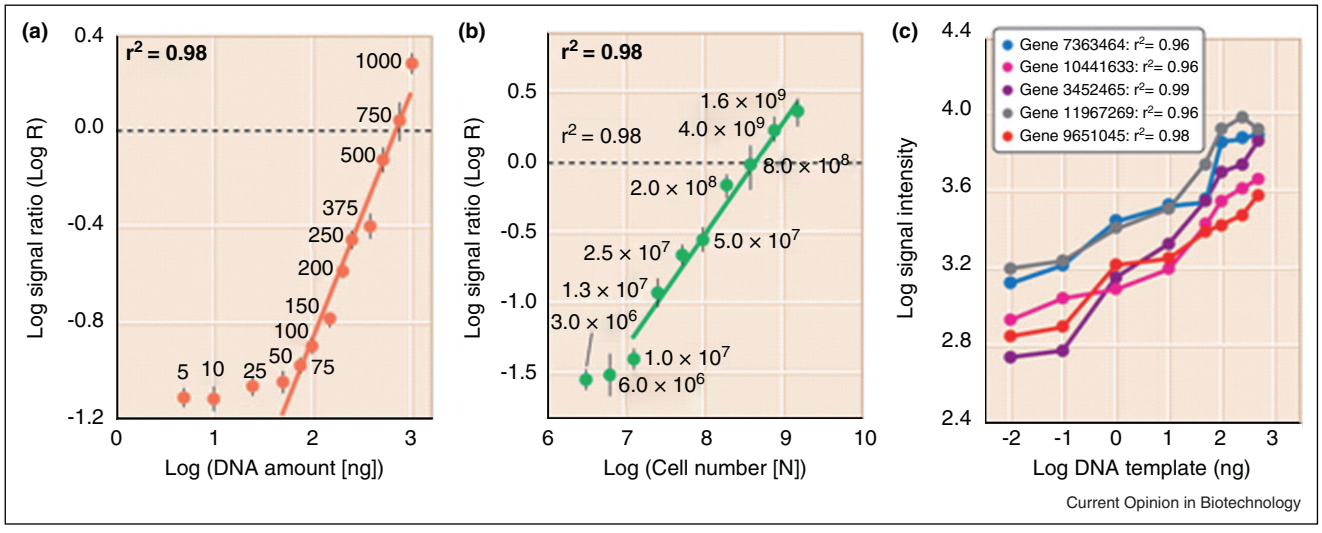
Quantitative capability

An important goal of microbial detection for environmental and ecological applications is to provide quantitative information. A linear relationship between target DNA and RNA concentrations and hybridization signal intensities was obtained with FGAs [11–13,51]. This linear relationship was observed for pure culture, mixed culture, and environmental samples with or without whole community genome amplification (WCGA) (Figure 3). In addition, quantitative real-time PCR analyses of representative genes showed a good correlation between the copy number of functional genes and the hybridization signal intensity [13,20••,21•]. However, quantitative capability can be affected by sequence divergence. For example, group-specific probes may reduce FGA quantitative ability as mentioned above. One of strategies is to design mismatch probes to distinguish signal from its noise [49]. Another strategy is to use relative comparisons between control and treatment samples so that any cross-hybridization can be cancelled [7,11,14].

Future directions

FGA technology has revolutionized the study of microbial ecology. It has been demonstrated to be a powerful tool to provide sensitive, specific, and potentially quantitative analysis of microbial communities from different ecosystems. Although next generation sequencing technologies

Figure 3



Evaluation of quantitative capability of functional gene arrays. **(a)** Quantitative relationship of FGA hybridization with pure genomic DNA in the presence of heterogeneous background DNA; **(b)** quantitative relationship of FGA hybridization with total RNA from a known number of pure culture cells in the presence of background DNA; **(c)** quantitative relationship between the signal intensity and the concentration of community DNA (0.01–250 ng) from a biostimulated groundwater sample.This figure is adapted from Figure 2 published in *Microbe* in 2010 [53].

Please cite this article in press as: He Z, et al. Development of functional gene microarrays for microbial community analysis, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.11.001

are available, FGAs are specifically useful for providing insights into our understanding of the functional diversity, composition and structure of microbial communities and their linkages with environmental factors and ecosystem functioning, and this is largely due to their unique features, including defined probe sets with known functions for community comparisons, rapid operation, and easy data analysis. However, to fully realize its potential, challenges for FGA development still remain, and this is especially true for the study of complex microbial communities. Therefore, the further development of FGAs should focus on the following key issues. First, it is necessary to develop more powerful software tools to design specific probes for large sets of sequence data generated from high throughput sequencing technologies. Second, because of sequence divergence in the environment, novel approaches, and strategies for improving sensitivity and quantitative accuracy are needed. In addition, future FGAs will be in two different types: more comprehensive FGAs for survey of diverse microbial communities, and more specific FGAs for specific detection and identification of microbial communities for particular ecosystems or functional processes of interest.

Acknowledgements

This work was part of ENIGMA, a Scientific Focus Area Program supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research, Genomics: GTL Foundational Science through Contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy, Environmental Remediation Science Program (ERSP), Office of Biological and Environmental Research, Office of Science, and Oklahoma Applied Research Support (OARS), Oklahoma Center for the Advancement of Science and Technology (OCAST), the State of Oklahoma through the Projects AR062-034 and AR11-035.

References and recommended reading

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

- • of special interest
- •• of outstanding interest

1. Zhou J, Thompson DK: **Challenges in applying microarrays to environmental studies**. *Curr Opin Biotechnol* 2002, **13**:204-207.

2. Zhou J: **Microarrays for bacterial detection and microbial community analysis**. *Curr Opin Microbiol* 2003, **6**:288-294.

3. Zhou J, He Q, Hemme CL, Mukhopadhyay A, Hillesland K, Zhou A, He Z, Van Nostrand JD, Hazen TC, Stahl DA *et al.*: **How sulphate-reducing microorganisms cope with stress: lessons from systems biology**. *Nat Rev Microbiol* 2011, **9**:452-466.

4. Huber JA, Mark Welch DB, Morrison HG, Huse SM, Neal PR, Butterfield DA, Sogin ML: **Microbial population structures in the deep marine biosphere**. *Science* 2007, **318**:97-100.

5. • Iwai S, Chai B, Sul WJ, Cole JR, Hashsham SA, Tiedje JM: **Gene-targeted-metagenomics reveals extensive diversity of aromatic dioxygenase genes in the environment**. *ISME J* 2009, **4**:279-285.

This paper describes an application of pyrosequencing technology to analyze functional genes.

6. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T *et al.*: **A human gut microbial gene catalogue established by metagenomic sequencing**. *Nature* 2010, **464**:59-65.

7. He Z, Gentry TJ, Schadt CW, Wu L, Liebich J, Chong SC, Huang Z, Wu W, Gu B, Jardine P *et al.*: **GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes**. *ISME J* 2007, **1**:67-77.

8. •• He Z, Deng Y, Van Nostrand JD, Tu Q, Xu M, Hemme CL, Li X, Wu L, Gentry TJ, Yin Y *et al.*: **GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity**. *ISME J* 2010, **4**:1167-1179.

This article marks the release of an updated GeoChip 3.0, with details for selected gene families, and describes its application to analyze the effects of plant diversity on the functional diversity and structure of soil microbial communities.

9. • Hazen TC, Dubinsky EA, DeSantis TZ, Andersen GL, Piceno YM, Singh N, Jansson JK, Probst A, Borglin SE, Fortney JL *et al.*: **Deep-sea oil plume enriches indigenous oil-degrading bacteria**. *Science* 2010, **330**:204-208.

This paper provides a summary of GeoChip 4.0 and details the use of GeoChip 4.0 and other technologies to examine the potential of indigenous microbial communities to degrade contaminants from the oil spill in the Gulf of Mexico.

10. Lu Z, Deng Y, Van Nostrand JD, He Z, Voordeckers J, Zhou A, Lee YJ, Mason OU, Dubinsky EA, Chavarria KL, *et al.*: **Microbial gene functions enriched in the deepwater horizon deep-sea oil plume**. *ISME J* 2011, doi:10.1038/ismej.2011.91

11. Wu L, Thompson DK, Li G, Hurt RA, Tiedje JM, Zhou J: **Development and evaluation of functional gene arrays for detection of selected genes in the environment**. *Appl Environ Microbiol* 2001, **67**:5780-5790.

12. Tiquia SM, Wu L, Chong SC, Passovets S, Xu D, Xu Y, Zhou J: **Evaluation of 50-mer oligonucleotide arrays for detecting microbial populations in environmental samples**. *Biotechniques* 2004, **36**:664-670.

13. Rhee S-K, Liu X, Wu L, Chong SC, Wan X, Zhou J: **Detection of genes involved in biodegradation and biotransformation in microbial communities by using 50-mer oligonucleotide microarrays**. *Appl Environ Microbiol* 2004, **70**:4303-4317.

14. Wu L, Liu X, Schadt CW, Zhou J: **Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification**. *Appl Environ Microbiol* 2006, **72**:4931-4941.

15. Van Nostrand JD, Wu W-M, Wu L, Deng Y, Carley J, Carroll S, He Z, Gu B, Luo J, Criddle CS *et al.*: **GeoChip-based analysis of functional microbial communities during the reoxidation of a bioreduced uranium-contaminated aquifer**. *Environ Microbiol* 2009, **11**:2611-2626.

16. Kimes NE, Van Nostrand JD, Weil E, Zhou J, Morris PJ: **Microbial functional structure of Montastraea faveolata, an important Caribbean reef-building coral, differs between healthy and yellow-band diseased colonies**. *Environ Microbiol* 2010, **12**:541-556.

17. Xu M, Wu W-M, Wu L, He Z, Van Nostrand JD, Deng Y, Luo J, Carley J, Ginder-Vogel M, Gentry TJ *et al.*: **Responses of microbial community functional structures to pilot-scale uranium in situ bioremediation**. *ISME J* 2010, **4**:1060-1070.

18. Yergeau E, Kang S, He Z, Zhou J, Kowalchuk GA: **Functional microarray analysis of nitrogen and carbon cycling genes across an Antarctic latitudinal transect**. *ISME J* 2007, **1**:163-179.

19. Zhou J, Kang S, Schadt CW, Garten CT: **Spatial scaling of functional gene diversity across various microbial taxa**. *Proc Natl Acad Sci U S A* 2008, **105**:7768-7773.

20. •• He Z, Xu M, Deng Y, Kang S, Kellogg L, Wu L, Van Nostrand JD, Hobbie SE, Reich PB, Zhou J: **Metagenomic analysis reveals a marked divergence in the structure of belowground microbial communities at elevated CO_2**. *Ecol Lett* 2010, **13**:564-575.

A good example of GeoChip 3.0 applications, which demonstrates that 10 years of field exposure of a grassland ecosystem to elevated CO_2 dramatically altered the structure and functional potential of soil microbial communities, and that changes in microbial structure were significantly correlated with soil C and N contents and plant productivity.

21. • Wang F, Zhou H, Meng J, Peng X, Jiang L, Sun P, Zhang C, Van Nostrand JD, Deng Y, He Z *et al.*: **GeoChip-based analysis of metabolic diversity of microbial communities at the Juan de**

Fuca Ridge hydrothermal vent. *Proc Natl Acad Sci U S A* 2009, **106**:4840-4845.
Applications of GeoChip to analyze micorbial communties from an extreme envrionment with low gene abundances, and evaluation of GeoChip quantiative capability by quantitative real-time PCR. The results indicate that GeoChip has high sensitivity for detecting low biomass microbial communities.

22. Xie J, He Z, Liu X, Liu X, Van Nostrand JD, Deng Y, Wu L, Qiu G, Zhou J: **GeoChip-based analysis of the functional gene diversity and metabolic potential of microbial communities in acid mine drainage**. *Appl Environ Microbiol* 2011, **77**:991-999.

23. Liu W, Wang A, Cheng S, Logan BE, Yu H, Deng Y, Van Nostrand JD, Wu L, He Z, Zhou J: **Geochip-based functional gene analysis of anodophilic communities in microbial electrolysis cells under different operational modes**. *Environ Sci Technol* 2010, **44**:7729-7735.

24. Liang Y, Van Nostrand JD, Deng Y, He Z, Wu L, Zhang X, Li G, Zhou J: **Functional gene diversity of soil microbial communities from five oil-contaminated fields in China**. *ISME J* 2011, **5**:403-413.

25. Taroncher-Oldenburg G, Griner EM, Francis CA, Ward BB: **Oligonucleotide microarray for the study of functional gene diversity in the nitrogen cycle in the environment**. *Appl Environ Microbiol* 2003, **69**:1159-1171.

26. Steward GF, Jenkins BD, Ward BB, Zehr JP: **Development and testing of a DNA macroarray to assess nitrogenase (*nifH*) gene diversity**. *Appl Environ Microbiol* 2004, **70**:1455-1465.

27. Bontemps C, Golfier G, Gris-Liebe C, Carrere S, Talini L, Boivin-Masson C: **Microarray-based detection and typing of the Rhizobium nodulation gene *nodC*: potential of DNA arrays to diagnose biological functions of interest**. *Appl Environ Microbiol* 2005, **71**:8042-8048.

28. Zhang L, Hurek T, Reinhold-Hurek B: **A *nifH*-based oligonucleotide microarray for functional diagnostics of nitrogen-fixing microorganisms**. *Microb Ecol* 2007, **53**:456-470.

29. Bodrossy L, Stralis-Pavese N, Murrell JC, Radajewski S, Weilharter A, Sessitsch A: **Development and validation of a diagnostic microbial microarray for methanotrophs**. *Environ Microbiol* 2003, **5**:566-582.

30. Bodrossy L, Stralis-Pavese N, Konrad-Koszler M, Weilharter A, Reichenauer TG, Schofer D, Sessitsch A: **mRNA-based parallel detection of active methanotroph populations by use of a diagnostic microarray**. *Appl Environ Microbiol* 2006, **72**:1672-1676.

31. Call DR, Bakko MK, Krug MJ, Roberts MC: **Identifying antimicrobial resistance genes with DNA microarrays**. *Antimicrob Agents Chemother* 2003, **47**:3290-3295.

32. Kostic T, Weilharter A, Sessitsch A, Bodrossy L: **High-sensitivity, polymerase chain reaction-free detection of microorganisms and their functional genes using 70-mer oligonucleotide diagnostic microarray**. *Anal Biochem* 2005, **346**:333-335.

33. Cleven BEE, Palka-Santini M, Gielen J, Meembor S, Kronke M, Krut O: **Identification and characterization of bacterial pathogens causing bloodstream infections by DNA microarray**. *J Clin Microbiol* 2006, **44**:2389-2397.

34. Jaing C, Gardner S, McLoughlin K, Mulakken N, Alegria-Hartman M, Banda P, Williams P, Gu P, Wagner M, Manohar C *et al.*: **A functional gene array for detection of bacterial virulence elements**. *PLoS ONE* 2008, **3**:e2163.

35. Miller SM, Tourlousse DM, Stedtfeld RD, Baushke SW, Herzog AB, Wick LM, Rouillard JM, Gulari E, Tiedje JM, Hashsham SA: **In situ-synthesized virulence and marker gene biochip for detection of bacterial pathogens in water**. *Appl Environ Microbiol* 2008, **74**:2200-2209.

36. Palka-Santini M, Cleven B, Eichinger L, Kronke M, Krut O: **Large scale multiplex PCR improves pathogen detection by DNA microarrays**. *BMC Microbiol* 2009, **9**:1.

37. Yin H, Cao L, Qiu G, Wang D, Kellogg L, Zhou J, Dai Z, Liu X: **Development and evaluation of 50-mer oligonucleotide arrays for detecting microbial populations in acid mine drainages and bioleaching systems**. *J Microbiol Methods* 2007, **70**:165-178.

38. •• McGrath KC, Mondav R, Sintrajaya R, Slattery B, Schmidt S, Schenk PM: **Development of an environmental functional gene microarray for soil microbial communities**. *Appl Environ Microbiol* 2010, **76**:7161-7170.
A comprehensive environmental functional gene array (E-FGA) was constructed with mRNA-enriched anonymous microbial clones from diverse microbial communities to profile microbial gene transcripts. Also, a new normalization method using internal spot standards was developed to overcome spotting and hybridization biases. Potential applications were evaluated with agricultural soils with a low or high flux of N_2O.

39. • Liang Y, He Z, Wu L, Deng Y, Li G, Zhou J: **Development of a common oligo reference standard (CORS) for microarray data normalization and comparison across different microbial communities**. *Appl Environ Microbiol* 2010, **76**:1088-1094.
This paper describes a common oligonucleotide reference standard (CORS) method for microarray data normalization and comparison across different microbial communities, and evaluates this method with other commonly used microarray normalization methods, genomic DNA and dye-swap approaches.

40. Eddy SR: **Profile hidden Markov models**. *Bioinformatics* 1998, **14**:755-763.

41. Li X, He Z, Zhou J: **Selection of optimal oligonucleotide probes for microarrays using multiple criteria, global alignment and parameter estimation**. *Nucleic Acids Res* 2005, **33**:6114-6123.

42. He Z, Wu L, Fields MW, Zhou J: **Use of microarrays with different probe sizes for monitoring gene expression**. *Appl Environ Microbiol* 2005, **71**:5154-5162.

43. Liebich J, Schadt CW, Chong SC, He Z, Rhee SK, Zhou J: **Improvement of oligonucleotide probe design criteria for functional gene microarrays in environmental applications**. *Appl Environ Microbiol* 2006, **72**:1688-1691.

44. He Z, Wu L, Li X, Fields MW, Zhou J: **Empirical establishment of oligonucleotide probe design criteria**. *Appl Environ Microbiol* 2005, **71**:3753-3760.

45. Chung W-H, Rhee S-K, Wan X-F, Bae J-W, Quan Z-X, Park Y-H: **Design of long oligonucleotide probes for functional gene detection in a microbial community**. *Bioinformatics* 2005, **21**:4092-4100.

46. Feng S, Tillier ERM: **A fast and flexible approach to oligonucleotide probe design for genomes and gene families**. *Bioinformatics* 2007, **23**:1195-1202.

47. •• Dugat-Bony E, Missaoui M, Peyretaillade E, Biderre-Petit C, Bouzid O, Gouinaud C, Hill D, Peyret P: **HiSpOD: probe design for functional DNA microarrays**. *Bioinformatics* 2011, **27**:641-648.
This paper presents a new efficient functional gene array probe design algorithm, high specific oligo design (HiSpOD) to design highly specific probes for individual nucleic sequences or consensus sequences from multiple alignments. Experimental evaluations were performed with genes involved in chlorinated solvent biodegradation.

48. He Z, Zhou J: **Empirical evaluation of a new method for calculating signal-to-noise ratio for microarray data analysis**. *Appl Environ Microbiol* 2008, **74**:2957-2966.

49. Deng Y, He Z, Van Nostrand J, Zhou J: **Design and analysis of mismatch probes for long oligonucleotide microarrays**. *BMC Genomics* 2008, **9**:491.

50. Chen Q, Yin H, Luo H, Xie M, Qiu G, Liu X: **Micro-array based whole-genome hybridization for detection of microorganisms in acid mine drainage and bioleaching systems**. *Hydrometallurgy* 2009, **95**:96-103.

51. Gao H, Yang ZK, Gentry TJ, Wu L, Schadt CW, Zhou J: **Microarray-based analysis of microbial community RNAs by whole-community RNA amplification**. *Appl Environ Microbiol* 2007, **73**:563-571.

52. Denef VJ, Park J, Rodrigues JLM, Tsoi TV, Hashsham SA, Tiedje JM: **Validation of a more sensitive method for using spotted oligonucleotide DNA microarrays for functional genomics studies on bacterial communities**. *Environ Microbiol* 2003, **5**:933-943.

53. Zhou J, He Z, Van Nostrand JD, Wu L, Deng Y: **Applying GeoChip analysis to disparate microbial communities**. *Microbe* 2010, **5**:60-65.

Please cite this article in press as: He Z, et al. Development of functional gene microarrays for microbial community analysis, Curr Opin Biotechnol (2011), doi:10.1016/j.copbio.2011.11.001