

Application of DNA array technology for diagnostic microbiology

Stephanie A Booth DPhil¹, Michael A Drebot PhD², Graham A Tipples PhD³, Lai King Ng PhD¹

Microarrays or DNA chips have been hailed as the ultimate experimental tool for research, drug discovery and diagnostics. They have the potential to perform a multitude of molecular tests simultaneously and to produce a wealth of information from a single clinical sample. Applications include genotyping, expression analysis and sequencing (1-4). The aim of this review is to provide a brief summary of current microarray technology and highlight the many ways in which it is being developed for use in clinical microbiology laboratories.

WHAT IS A DNA MICROARRAY?

A DNA microarray consists of a solid support, usually glass, with unique nucleic acid sequences (probes) bound at discrete positions on the slide. Nucleic acid hybridization is the central event in microarray technology. Sequences from a labelled target are recognized by the hybridization affinities of the oligonucleotide probes in a highly specific and sensitive manner. There are a number of different microarray formats, the biggest distinction among them being whether oligonucleotides are synthesized in situ onto the chip, or whether cDNA or oligonucleotides are spotted onto the chip with high precision (5). Recently a light-directed oligonucleotide synthesis process has been developed to generate probes directly onto a solid surface (6). This procedure uses combinations of unique masks to allow light to reach predefined sites on the array to build probes up to 30 basepairs in length accurately. It is possible to produce approximately 400,000 unique probes in squares 20 µm across on a small glass wafer (7). The numbers of probes are so high that they can be designed to interrogate every base within genes, in

effect allowing sequencing on the chip (4). The masks are unique for a given set of probe sequences, and it is time consuming and expensive to produce them, which limits the use of this technology to produce custom arrays. New innovations, such as the introduction of mirrors to direct the light, may improve flexibility and broaden the use of this type of array for diagnostics (8). One other way to synthesize probes directly onto a surface is by using ink jet printers to direct small volumes of oligonucleotide synthesis reagents onto predefined sites on the chip (9). This technology should be both flexible and relatively inexpensive for the production of custom arrays (10).

Presynthesized oligonucleotides can also be spotted onto a solid surface using a robot. This type of array may contain probes designed to pick out short unique sequences or detect point mutations for genotyping. Spotted arrays are attractive because they are very flexible, and more probes can be added and changes can be made to the chip as new targets are identified. There are two major practical problems that limit the development and use of arrays for genotyping. The first is the difficulty in designing oligonucleotide probes that discriminate well enough to recognize single base changes under uniform hybridization conditions (1,11). The other is in the preparation of the target sample. The assay requires nucleic acid amplification of different genomic targets from clinical samples before hybridization. In this 'multiplex polymerase chain reaction (PCR)' technique, the challenge is to design numerous sets of primers with similar annealing temperatures for amplification in the same tube (12-14).

Novel approaches to standardize the hybridization conditions for all the probes on a chip, even if they have different

¹National Laboratory for Sexually Transmitted Disease, ²National Laboratory for Zoonotic Diseases and Special Pathogens, ³National Laboratory for Viral Diagnostics, National Microbiology Laboratory, Population and Public Health Branch, Health Canada, Winnipeg, Manitoba

Correspondence and reprints: Dr Lai King Ng, National Laboratory for Sexually Transmitted Diseases, National Microbiology Laboratory, Population and Public Health Branch, 1015 Arlington Street, Winnipeg, Manitoba R3E 3R2. Telephone 204-789 2131, fax 204-789-2140, e-mail Lai_king_ng@hc-sc.gc.ca

guanine-cytosine (GC) contents and binding efficiencies, are under development. These include the production of DNA chips containing microelectronic devices to facilitate the hybridization discrimination of mismatches (15,16), alternative probe chemistries to equalize the DNA duplex stabilities of comprehensive sets of oligonucleotide probes (17) and the use of chaotropic agents in the hybridization reaction (18).

One of the most promising ways to increase specificity on the microarray is to include an enzymatic step in the assay. The most popular method using this approach is termed 'minisequencing'. Hybridization of an unlabelled target to the arrayed probes is followed by a primer extension reaction (19). Oligonucleotides are designed to be complementary to the target, ending immediately adjacent to the base under interrogation. The probe is then extended by a single-labelled base using a DNA polymerase, and the identity of the label is determined by scanning the array. Using the scanners with four-colour recognition capability, all four nucleotides can be determined on a single array (20). Another conceptually similar method to increase specificity is the inclusion of a ligation step to attach a probe containing either the mutant or wild type allele.

APPLICATIONS OF OLIGONUCLEOTIDE ARRAYS IN DIAGNOSTICS AND SURVEILLANCE

Traditionally many of the standard methods for looking at characteristics of pathogenic organisms, such as typing of isolates, resistance to drugs and the identification of virulence, were performed on laboratory cultures. Often the scope of such tests is limited by the slow growth and fastidiousness of the organism. As a result, it can be difficult to provide information about phenotypes fast enough to be clinically relevant. In addition this information does not reveal anything about the mechanisms behind particular characteristics. Many molecular tests used in microbiological diagnostics rely on DNA sequencing or nucleic acid hybridization to identify specific sequences or point mutations that vary between different clones (21-24). An array format expands on these methods by enabling the simultaneous detection of thousands of genomic targets. This will result in more accurate and efficient genotyping plus information on the expression of specific phenotypic characteristics, such as drug resistance (13,25).

The utility of high density DNA chips for typing has been demonstrated for both HIV and *Mycobacterium tuberculosis* using high density arrays of in situ synthesized oligonucleotides (25,26). In the *M. tuberculosis* studies, probes were designed to characterize the 16S rRNA for typing of tuberculosis strains and the *rpoB* gene for analyzing mutations linked to resistance against the antibiotic rifampicin (25,27). The chip consists of a 'tiled array' of oligonucleotide probes 20 to 25 bases in length, with four different probes designed to interrogate each base of the target sequence in turn (28). The base under interrogation is in the central position of the probe where a mismatch produces the greatest degree of hybridization instability. Quantitation of hybridization sig-

nals allows perfectly matched probe and targets to be discriminated over mismatched duplexes under stringent conditions. This study illustrates the utility of DNA arrays for identifying sequence changes leading to drug resistance. The use of a standard chip for identifying multiple bacterial pathogens and drug resistance markers would be a cost effective diagnostic tool (27). A universal sequencing chip is also under development, and this will have many potential uses for the sequencing of short genomic targets for typing and other diagnostic applications (10,29).

The detection of viral pathogens is another application for which microarrays have potential as a diagnostic tool. Low density, spotted oligonucleotide arrays that define either short sequences of interest or individual mutations could be used to identify virulence markers that distinguish viral vaccine strains from wild type isolates. Microarrays that include both generic and strain-specific probes could also be used for identifying both previously recognized strains and new or 'variant' viruses associated with outbreaks (30-32). DNA chips have the capability to genotype viral pathogens and may be useful in determining viral transmission pathways, and the source of outbreaks (33). Another application for arrays could include the identification of distinct subspecies of vectors and reservoirs that harbour zoonotic pathogens such as hantaviruses and various kinds of arboviruses (34).

APPLICATIONS OF GENOMIC DNA ARRAYS IN STUDIES OF GENE EXPRESSION AND PATHOGENS

Pathogen gene expression: The most common use of DNA microarrays is for monitoring expression levels of transcripts from cells, viruses and bacteria (35-37). Probes to be spotted onto slides can be made in three ways: PCR amplicons generated from genomic regions defining predicted open reading frames; cDNA generated from mRNA by reverse transcriptase (RT)-PCR; and oligonucleotides (35-37).

The mRNA target can be used directly from culture or RT-PCR amplified and subsequently labelled (35). One challenge in performing microarray analysis in bacteria is the need to extract mRNA quickly due to its short half-life in many organisms. Priming for cDNA synthesis can be achieved using random hexamers or short oligonucleotides (38-41). A recent study reports that the use of random hexamer priming resulted in the most accurate quantitation of expression levels (42).

Profiling gene expression levels in an isolate is useful to predict the function of uncharacterized genes or for analyzing the expression of virulence-associated genes. Monitoring the levels of expression of certain genes in pathogens may be important for diagnosing particularly virulent strains or comparing responses to different drugs (37). This type of assay may be useful in deciding on an effective drug treatment against organisms that are initially unresponsive to therapy. The use of whole genome microarrays is not limited to bacteria. Recent publications show how microarray analysis was used to study human cytomegalovirus and herpes simplex virus 1 gene expression (43,44).

Host gene expression: The response of a host to an invading

pathogen involves a variety of mechanisms. With the availability of arrays of human genes comes the opportunity to investigate in detail the relation between host and pathogen (45-47). As well as providing clues into the mechanisms of microbial pathogenicity, results from this type of study may provide indicators to identify those individuals most susceptible to infection and prognostic markers for the outcome of the infection. This type of information could be used to predict the most effective course of treatment during infection or possibly even to screen susceptible populations where preventative measures could be taken.

Comparison of gene expression patterns has already been validated as a means to classify tumour types in human cancer studies (48,49). Host expression patterns could also be used as a diagnostic tool in infectious disease and as an indicator of disease progression. This kind of application is supported by the finding that the patterns of gene expression induced in primary human monocytes infected by two closely related strains of Ebola viruses, Zaire and Reston, are significantly different. In this study microarray methodology was able to differentiate between host gene expression profiles generated during infection with the two Ebola strains (50).

One of the most important aspects of this type of approach involving the analysis of thousands of genes simultaneously is the development of algorithms for the analysis of the huge amounts of data produced. The term 'data-mining' has been adopted to encompass this field (2,51).

Comparative genomic hybridization: Comparative genomic hybridization has been used for a number of years. One important application is scanning for differences in DNA sequence copy number in tumour cells to characterize certain tumour types or stages of tumour progression (52). This type of assay can be readily transferred to an array format to produce information on the number of DNA copies and the persistence of particular genes between different pathogens (53,54). Amplicons representing some or all genes from a particular species are spotted onto arrays and interrogated with labelled genomic target sequences derived from different strains or disease states (54). An example of the utility of this method is provided in a study on bacilli Calmette-Guérin vaccine strains resulting in a reconstruction of the phylogeny of the vaccine over time (53).

One pitfall of this approach is that only known sequences can be represented on the array, which means that insertions in the target sequence are not detected. Another consideration is that the degree of sequence variation tolerated in the hybridization reaction on the microarray has not been characterized, and could lead to anomalies. The more sequence information that is gained from genomic sequencing projects and functional genomic studies the greater the potential for the design of reference arrays as diagnostic and surveillance tools.

CONCLUSION

The variety of approaches described above illustrate the enormous potential for microarray analysis in almost every

aspect of clinical microbiology. The huge amount of commercial and academic interest in microarray technology will guarantee continued innovation and improve the accessibility of arrays for routine diagnostics and surveillance. Although few studies are as yet published for many of the microarray approaches described, projects using DNA chip methodology will increase exponentially in the near future. The development of arrays for clinical use in infectious disease is limited by the amount of sequence information available for pathogenic organisms. The standardization of systematic methods to allow comparisons of array data and the computational tools necessary to analyze and model huge sets of biological data are also not sophisticated enough for routine use at present. Nevertheless the promise of flexible, simultaneous, high-throughput diagnostics will ensure that DNA array analysis will eventually become a standard part of the diagnostic microbiology laboratory.

REFERENCES

- Gerhold D, Rushmore T, Caskey CT. DNA chips: promising toys have become powerful tools. *Trends Biochem Sci* 1999;24:168-73.
- Brown PO, Botstein D. Exploring the new world of the genome with DNA microarrays. *Nat Genet* 1999;21(1 Suppl):33-7.
- Lockhart DJ, Winzler EA. Genomics, gene expression and DNA arrays. *Nature* 2000;405:827-36.
- Fan JB, Chen X, Halushka MK, et al. Parallel genotyping of human SNPs using generic high-density oligonucleotide tag arrays. *Genome Res* 2000;10:853-60.
- Lemieux B, Aharoni A, Schena M. Overview of DNA chip technology. *Mol Breeding* 1998;4:277-89.
- Fodor SP, Read JL, Pirrung MC, Stryer L, Lu AT, Solas D. Light-directed, spatially addressable parallel chemical synthesis. *Science* 1991;251:767-73.
- Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ. High density synthetic oligonucleotide arrays. *Nat Genet* 1999;21(1 Suppl):20-4.
- Singh-Gasson S, Green RD, Yue Y, et al. Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nat Biotechnol* 1999;17:974-8.
- Okamoto T, Suzuki T, Yamamoto N. Microarray fabrication with covalent attachment of DNA using Bubble Jet technology. *Nat Biotechnol* 2000;18:438-41.
- Drmanac R, Drmanac S. cDNA screening by array hybridization. *Methods Enzymol* 1999;303:165-78.
- Lipshutz RJ, Morris D, Chee M, et al. Using oligonucleotide probe arrays to access genetic diversity. *Biotechniques* 1995;19:442-7.
- Wang DG, Fan JB, Siao CJ, et al. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* 1998;280:1077-82.
- Gentalen E, Chee M. A novel method for determining linkage between DNA sequences: hybridization to paired probe arrays. *Nucleic Acids Res* 1999;27:1485-91.
- Wang E, Miller LD, Ohnmacht GA, Liu ET, Marincola FM. High-fidelity mRNA amplification for gene profiling. *Nat Biotechnol* 2000;18:457-9.
- Edman CF, Raymond DE, Wu DJ, et al. Electric field directed nucleic acid hybridization on microchips. *Nucleic Acids Res* 1997;25:4907-14.
- Edman CF, Mehta P, Press R, Spargo CA, Walker GT, Nerenberg M. Pathogen analysis and genetic predisposition testing using microelectronic arrays and isothermal amplification. *J Investig Med* 2000;48:93-101.
- Weiler J, Gausepohl H, Hauser N, Jensen ON, Hoheisel JD. Hybridisation based DNA screening on peptide nucleic acid

- (PNA) oligomer arrays. *Nucleic Acids Res* 1997;25:2792-9.
18. Nguyen HK, Fournier O, Asseline U, Dupret D, Thuong NT. Smoothing of the thermal stability of DNA duplexes by using modified nucleosides and chaotropic agents. *Nucleic Acids Res* 1999;27:1492-8.
 19. Syvanen AC. From gels to chips: "minisequencing" primer extension for analysis of point mutations and single nucleotide polymorphisms. *Hum Mutat* 1999;13:1-10.
 20. Pastinen T, Raitio M, Lindroos K, Tainola P, Peltonen L, Syvanen AC. A system for specific, high-throughput genotyping by allele-specific primer extension on microarrays. *Genome Res* 2000;10:1031-42.
 21. Dumler JS, Valsamakis A. Molecular diagnostics for existing and emerging infections. Complementary tools for a new era of clinical microbiology. *Am J Clin Pathol* 1999;112:S33-9.
 22. Wagar EA. Direct hybridization and amplification applications for the diagnosis of infectious diseases. *J Clin Lab Anal* 1996;10:312-25.
 23. Tang YW, Procop GW, Persing DH. Molecular diagnostics of infectious diseases. *Clin Chem* 1997;43:2021-38.
 24. El-Sayed NM, Hegde P, Quackenbush J, Melville SE, Donelson JE. The African trypanosome genome. *Int J Parasitol* 2000;30:329-45.
 25. Gingeras TR, Ghandour G, Wang E. Simultaneous genotyping and species identification using hybridization pattern recognition analysis of generic *Mycobacterium* DNA arrays. *Genome Res* 1998;8:435-48.
 26. Wilson JW, Bean P, Robins T, Graziano F, Persing DH. Comparative evaluation of three human immunodeficiency virus genotyping systems: the HIV-GenotypR method, the HIV PRT GeneChip assay, and the HIV-1 RT line probe assay. *J Clin Microbiol* 2000;38:3022-8.
 27. Troesch A, Nguyen H, Miyada CG, et al. *Mycobacterium* species identification and rifampin resistance testing with high-density DNA probe arrays. *J Clin Microbiol* 1999;37:49-55.
 28. Cronin MT, Fucini RV, Kim SM, Masino RS, Wespi RM, Miyada CG. Cystic fibrosis mutation detection by hybridization to light-generated DNA probe arrays. *Hum Mutat* 1996;7:244-55.
 29. Diamandis EP. Sequencing with microarray technology – a powerful new tool for molecular diagnostics. *Clin Chem* 2000;46:1523-5.
 30. Bellini WJ, Rota PA. Genetic diversity of wild-type measles viruses: implications for global measles elimination programs. *Emerg Infect Dis* 1998;4:29-35.
 31. Drebot MA, Campbell JJ, Lee SH. A genotypic characterization of enteroviral antigenic variants isolated in eastern Canada. *Virus Res* 1999;59:131-40.
 32. Drebot MA, Tipples GA. Application of molecular diagnostics for the surveillance of hantavirus and measles virus in Canada. *Can J Infect Dis* 1998;9:71-5.
 33. Bean P, Wilson J. HIV genotyping by chip technology. *Am Clin Lab* 2000;19:16-7.
 34. Clement JP, Ranst MV. The role of rodents in emerging and re-emerging human infections. *Infect Dis Rev* 2000;2:84-7.
 35. Gingeras TR, Rosenow C. Studying microbial genomes with high-density oligonucleotide arrays. *Am Soc Microbiol News* 2000;66:463-9.
 36. Harrington CA, Rosenow C, Retief J. Monitoring gene expression using DNA microarrays. *Curr Opin Microbiol* 2000;3:285-91.
 37. Wilson M, DeRisi J, Kristensen HH, et al. Exploring drug-induced alterations in gene expression in *Mycobacterium tuberculosis* by microarray hybridization. *Proc Natl Acad Sci USA* 1999;96:12833-8.
 38. de Saizieu A, Certa U, Warrington J, Gray C, Keck W, Mous J. Bacterial transcript imaging by hybridization of total RNA to oligonucleotide arrays. *Nat Biotechnol* 1998;16:45-8.
 39. Fislage R, Berceanu M, Humboldt Y, Wendt M, Oberender H. Primer design for a prokaryotic differential display RT-PCR. *Nucleic Acids Res* 1997;25:1830-5.
 40. de Saizieu A, Gardes C, Flint N, et al. Microarray-based identification of a novel streptococcus pneumoniae regulon controlled by an autoinduced peptide. *J Bacteriol* 2000;182:4696-703.
 41. Talaat AM, Hunter P, Johnston SA. Genome-directed primers for selective labeling of bacterial transcripts for DNA microarray analysis. *Nat Biotechnol* 2000;18:679-82.
 42. Arfin SM, Long AD, Ito ET, et al. Global gene expression profiling in *Escherichia coli* K12. THE EFFECTS OF INTEGRATION HOST FACTOR. *J Biol Chem* 2000;275:29672-84.
 43. Bresnahan WA, Shenk T. A subset of viral transcripts packaged within human cytomegalovirus particles. *Science* 2000;288:2373-6.
 44. Stingley SW, Ramirez JJ, Aguilar SA, et al. Global analysis of herpes simplex virus type 1 transcription using an oligonucleotide-based DNA microarray. *J Virol* 2000;74:9916-27.
 45. Cummings CA, Relman DA. Using DNA Microarrays to study host-microbe interactions. *Emerg Infect Dis* 2000;6:513-25.
 46. Manger ID, Relman DA. How the host 'sees' pathogens: global gene expression responses to infection. *Curr Opin Immunol* 2000;12:215-8.
 47. Zhu H, Cong JP, Mamtora G, Gingeras T, Shenk T. Cellular gene expression altered by human cytomegalovirus: global monitoring with oligonucleotide arrays. *Proc Natl Acad Sci USA* 1998;95:14470-5.
 48. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;403:503-11.
 49. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999;286:531-7.
 50. Xiang C, Young H, Alterson H, et al. Comparison of cellular gene expression in Ebola-Zaire and Ebola-Reston virus-infected primary human monocytes. *Nat Genet* 1999;23:82.
 51. Bassett DE Jr, Eisen MB, Boguski MS. Gene expression informatics – it's all in your mine. *Nat Genet* 1999;21(1 Suppl):51-5.
 52. Forozan F, Karhu R, Kononen J, Kallioniemi A, Kallioniemi OP. Genome screening by comparative genomic hybridization. *Trends Genet* 1997;13:405-9.
 53. Behr MA, Wilson MA, Gill WP, et al. Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* 1999;284:1520-3.
 54. Pollack JR, Perou CM, Alizadeh AA, et al. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* 1999;23:41-6.