KMITL Sci. Tech. J. Vol. 6 No. 1 Jan. - Jun. 2006

cDNA MICROARRAY TECHNOLOGY FOR THE ANALYSIS OF GENE EXPRESSION

Choowong Auesukareę

Department of Applied Biology, Faculty of Science,
King Mongkut's Institute of Technology Ladkrabang,
Chalongkrung Rd., Ladkrabang, Bangkok, 10520, Thailand

ABSTRACT

DNA microarray technology is a new and powerful tool for studying a genome-wide gene expression in many organisms. This technology is a high throughput, highly parallel RNA expression assay technique that permits quantitative analysis of RNAs transcribed from both known and unknown genes. This review describes array manufacturing, experimental design, and its application to gene discovery and expression. Microarray technology is rapidly becoming a central platform for functional genomics.

KEYWORDS: microarray, gene expression, functional genomics

1. INTRODUCTION

The ability of cells to store and translate the genetic information is essential to create and maintain a living organism. DNA stores this information for the synthesis of proteins which are the active working components of all the cellular machinery. Each cell within an organism contains a copy of the same genome. RNA carries out the expression of the information encoded in DNA and messenger RNAs (mRNAs) are the transcripts that carry the specific information for the sequence of amino acids in proteins. Since changes in the physiology of an organism or a cell during its development or the stimulation by physical and chemical reagents will be accompanied by changes in the pattern of gene expression, the analysis of gene expression is therefore important in many fields of biological research. To better understand the development of organisms, the onset of genetic diseases, or the concerted functions of genes in regulating cell function, it is necessary to study the changes in genome-scale gene expression patterns. Currently, due to advances in genome-scale technologies, biological researchers have shifted their interest to the functional aspects of thousands of genes, instead of the structural information of a single gene.

Several techniques for the analysis of gene expression at the mRNA-level are available, such as Northern blotting [1] and differential display [2]. However, these methods each have their disadvantages for the quantitative and simultaneous analysis of large numbers of expression products. Northern blot analysis only allows limited numbers of mRNAs to be studied at the same time. Differential display can simultaneously detect multiple differences in gene expression. However, this method is not quantitative and only a limited number of different conditions can be compared.

Recently, the DNA microarray technology has been developed and has become a powerful tool for high-throughput study of the expressions of all genes in the genome simultaneously in many organisms. This method is based on hybridization of mRNA to a high-density array of immobilized DNA probes, each corresponding to a specific gene. Instead of working on a gene-by-gene basis, scientists can now study tens of thousands of genes at once. This technology has been steadily developed since Patrick Brown and his

*Corresponding author. Tel: 662-201-5481 Fax: 662-354-7161 E-mail: choowonga@hotmail.com
Present address: Department of Biology, Faculty of Science, Mahidol University, Thailand

colleagues first published their work in 1995 [3]. Although this technology is new, it has already led to significant research advances and its use is spreading rapidly (Figure 1).

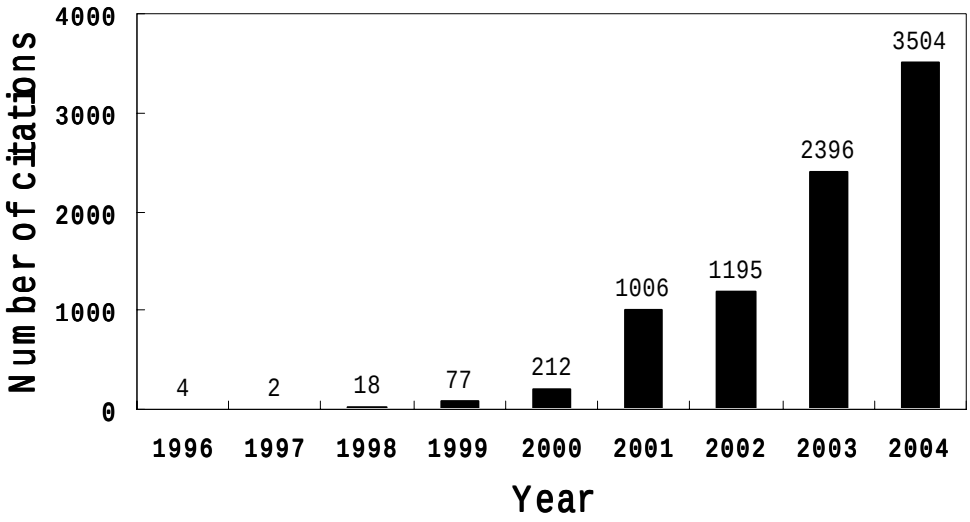

Figure 1 *Pubmed* citations containing the word "microarray" have increased rapidly in the past few years, which shown acceptance of this technology.

2. DNA MICROARRAYS

DNA microarrays consist of DNA probes attached to the surface of a glass microscope slide at extremely high density [3-4]. More than 20,000 genes (a half of expressed human genes) can be arrayed on a surface of a glass slide. Microarrays allow the simultaneous monitoring of gene expression of thousands or tens-of-thousands of genes in one hybridization experiment because of two basic principles of nucleic-acid hybridization: DNA and RNA will specifically bind to their complementary sequence, and this binding occurs in proportion to the abundance of a sequence.

The DNA probes that are used in a DNA microarray could be amplified cDNA fragments or synthesized DNA oligonucleotides that have sequences that complement the target sequences. Thus, DNA microarrays are categorized into cDNA microarrays and oligonucleotide arrays [4-7]. These DNA oligomers are either directly synthesized on the surface of the microarray or deposited onto the surface mechanically and covalently immobilized on the surface [3]. cDNA microarrays which were originally developed in the Brown Laboratory at Stanford University [3, 6] have been the predominant method for the parallel analysis of gene expression in various biological processes.

In DNA microarray experiments, the mRNA is normally transformed by reverse transcription into cDNA, which is more stable. mRNA or cDNA from a sample is applied to the array surface and allowed to hybridize. Sample mRNAs are labeled as a complex mixture, usually by incorporation of a fluorescent nucleotide (Cy3 and Cy5 labeled dCTP or dUTP are often used for this purpose) by oligo(dT)-primed reverse transcription. The labeled pool of sample mRNAs is subsequently hybridized to the array, where each messenger will quantitatively hybridize to its complementary target sequence. After washing, the

fluorescence at each spot on the array is a quantitative measure corresponding to the expression level of the particular gene. The strength of the fluorescent signal from the captured targets reflects the abundance of the target molecules and/or the binding compatibility between the probe and the target molecules. A charge-coupled device (CCD) or laser scanner can be used to record these fluorescence signals quantitatively. The use of two differently labelled mRNA samples allows quantitative comparison of gene expression in a pair of samples (Figure 2). The major advance of DNA microarray technology, as compared to conventional techniques, results from the small size of the array, which allows for a higher sensitivity, enables the parallel screening of larger numbers of genes and the use of smaller amounts of starting material. Mainly, the introduction of fluorescent probes has made miniaturisation of arrays possible [8]. The scale of gene expression analysis is not only extended by the simultaneous analysis of large numbers of genes, but also microarrays can be produced in series facilitating comparative analysis of a large number of samples.

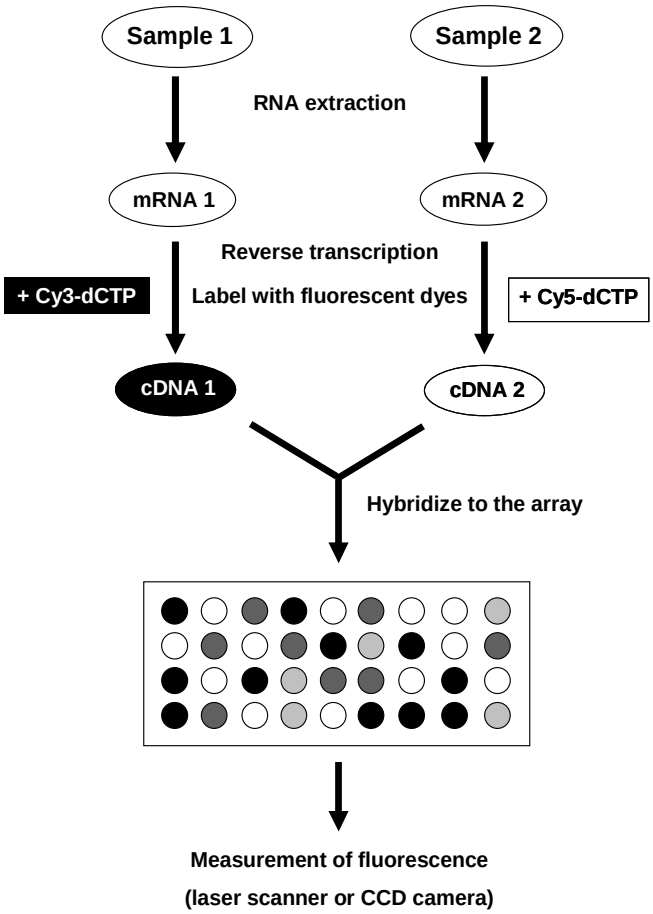

Figure 2 Outline of gene expression analysis by DNA microarray technology. RNA is extracted from the two comparison specimens and reverse transcribed to cDNA which is labeled with different fluorescence (Cy3 and Cy5 labeled dCTP or dUTP are often used for this purpose). The two samples are then hybridized to the array. The array is scanned by a laser scanner or a CCD camera. The final microarray images are analyzed by various computer programs.

3. MANUFACTURING DNA MICROARRAYS

Two main approaches for manufacturing DNA microarrays exist. The first approach is the direct synthesis of oligonucleotides on a solid surface. The method allows the manufacturing of microarrays with very high densities (>250,000 oligonucleotide spots per cm^2) and facilitates the production of large series of identical arrays. However, the design of this type of array, also named a DNA chip, is prohibitively expensive, has no flexibility in design and only a limited number of different arrays is currently on the market. The second approach, DNA micro-dispensing, is more flexible and can be performed in a regular molecular biology laboratory [3]. Small quantities of DNA solution, with a minimum volume of approximately 50 pl, are dispensed onto a solid surface. The number of commercially available micro-dispensing robots is quickly increasing and the performance of these machines is continually improving.

cDNA arrays are composed of PCR-amplified cDNA clones arranged on a non-porous surface. A typical cDNA array is printed onto a 30 mm by 15 mm glass microscope slide by a computer-controlled robotic arm; each spot is about 50-150 μm in diameter. Other flexible and porous surfaces such as nylon membranes can also be used [9]. Arrays of short oligonucleotides (about 25 bp) made by photolithography in situ are also available commercially and form the basis of GeneChip® technology [4]. Arrays of longer oligonucleotides manufactured in situ by use of ink jet printing technology are available through Agilent Technologies [10]. The longer oligonucleotide arrays have the advantage of not requiring the laborious process of PCR-amplification and clone-insert purification before arraying.

4. MICROARRAY EXPERIMENTAL DESIGN

Design of a microarray is largely dependent on the research question that has to be answered. Diagnostic questions will often involve smaller arrays with well defined target sequences, while elucidation of metabolic pathways or identification of novel responding genes requires the use of large arrays often with undefined target sequences. If the number of genes of an organism is not too large and all the genes are known, sequences corresponding to all open reading frames (ORFs) can be spotted on the array which allows simultaneous expression analysis of all mRNAs. At present, the complete genome sequences of several organisms such as *Escherichia coli*, *Saccharomyces cerevisiae* and *Caenorhabditis elegans* are available. For yeast, arrays containing all ORFs are commercially available and have been widely used in expression screening. Although existing sequence databases give a great advantage for selecting relevant genes, cDNA libraries can be used as source for the identification of new, unknown genes.

Since relative mRNA levels change rapidly in response to changes in regulatory signals, a DNA microarray experiment is often a measurement of differences in mRNA abundance between two conditions. Essentially, there are three types of the two-condition experimental design; differential response to growth parameters; treated versus untreated samples; and wild-type versus mutant strains. Examples of the first of these include differential gene expression studies of *E. coli* grown in minimal versus rich media [11] and of *S. cerevisiae* grown in minimal versus rich media [12]. Comparison of treated versus untreated samples is a strategy for measuring differential gene expression resulting from exposure to agents causing deleterious effects on growth or induction of global regulatory networks. Differential gene expression of Arabidopsis thaliana after the addition of nitrate led to the identification of novel nitrate-induced genes [13]. The expression profiling study of *E. coli* after the addition of acetate identified differentially expressed genes involved in the acid tolerance response [14]. Comparison of mutant strains with the wild-type strain is a powerful tool for the investigation of regulatory networks. The genome-wide expression patterns of

S. cerevisiae mutants defective in Mec1 signaling under normal growth conditions and in response to the methylating-agent methylmethane sulfonate (MMS) and ionizing radiation identified a set of genes which appears to represent an Mec1-dependent expression signature of DNA damage [15].

A simple comparison of steady-state mutant and wild-type cultures is appropriate for some cases but this is not a reasonable design for investigating all regulatory factors. For example, with inducible stress response networks, the loss of the regulatory factor would cause a failure to respond to the stress signal. Comparison of treated versus untreated wild-type cells reveals the normal global response to the treatment, whereas the same experiment run in the regulatory mutant identifies target genes in specific regulons. An example of this strategy is the analysis of genome-wide transcription profile of *E. coli* deletion strains of OxyR, the peroxide response regulator, treated with hydrogen peroxide. By this strategy, this work was successful in the identification of several new OxyR-activated and OxyR-repressed genes [16].

5. APPLICATIONS

DNA microarrays have a large number of applications that will expand and diversify over time. At present, the measurement of expression levels for thousands of genes in parallel appears to be one of the most biologically informative applications of this new technology. Since the expression data serves as a direct link to function, this technology can be applied to many biological fields. The DNA microarray technology was used to profile complex diseases and discover novel disease-related genes. Perou *et al.* [17] used the cDNA microarray to analyze gene expression in breast cancer and identified groups of genes associated with the cell proliferative state, with the response to interferon gamma, and with the presence of stromal cells and lymphocytes in tumor biopsies. De Risi *et al.* [18] studied the metabolic and genetic control of gene expression in yeast by using the yeast microarrays. They examined the effects of the diauxic shift from anaerobic to aerobic metabolism under glucose limitation and the concomitant switch to ethanol as a carbon source. Marton [19] performed drug validation studies and identified secondary drug target effects. In conclusion, significant advances in the DNA microarray technology have made the analysis of whole-genome expression possible and have led to a new level of understanding of the mechanism of transcription and of cellular functions.

REFERENCES

[1] Alwine, J.C., Kemp, D.J. and Stark, G.R. **1977** Method for Detection of Specific RNAs in Agarose Gels by Transfer to Diazobenzyloxymethyl-paper and Hybridization with DNA Probes, *Proc Natl Acad Sci U S A, 74*(12), 5350-5354.

[2] Liang, P. and Pardee, A.B. **1992** Differential Display of Eukaryotic Messenger RNA by means of the Polymerase Chain Reaction, *Science*, *257*(5072), 967-971.

[3] Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. **1995** Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray, *Science*, *270*(5235), 467-470.

[4] Lipshutz, R.J., Fodor, S.P., Gingeras, T.R. and Lockhart, D.J. **1999** High Density Synthetic Oligonucleotide Arrays, *Nat Genet*, *21*(1 Suppl), 20-24.

[5] Shalon, D., Smith, S.J. and Brown, P.O. **1996** A DNA Microarray System for Analyzing Complex DNA Samples using Two-color Fluorescent Probe Hybridization, *Genome Res.*, *6*(7), 639-645.

[6] DeRisi, J., Penland, L., Brown, P.O., Bittner, M.L., Meltzer, P.S., Ray, M., Chen, Y., Su, Y.A. and Trent, J.M. **1996** Use of a cDNA Microarray to Analyse Gene Expression Patterns in Human Cancer, *Nat. Genet*, *14*(4), 457-460.

[7] Lockhart, D.J. and Winzeler, E.A. **2000** Genomics, Gene Expression and DNA Arrays, *Nature*, *405*(6788), 827-836.

[8] Jordan, B.R. **1998** Large-scale Expression Measurement by Hybridization Methods: from High-density Membranes to "DNA chips", *J. Biochem (Tokyo)*, *124*(2), 251-258.

[9] Whitney, L.W., Becker, K.G., Tresser, N.J., Caballero-Ramos, C.I., Munson, P.J., Prabhu,V.V., Trent, J.M., McFarland, H.F. and Biddison, W.E. **1999** Analysis of Gene Expression in Multiple Sclerosis Lesions using cDNA Microarrays, *Ann Neurol*, *46*(3), 425-428.

[10] Hughes, T.R., Mao, M., Jones, A.R., Burchard, J., Marton, M.J., Shannon, K.W., Lefkowitz, S.M., Ziman, M., Schelter, J.M., Meyer, M.R., Kobayashi, S., Davis, C., Dai, H., He, Y.D., Stephaniants, S.B., Cavet, G., Walker, W.L., West, A., Coffey, E., Shoemaker, D.D., Stoughton, R., Blanchard, A.P., Friend, S.H. and Linsley, P.S. **2001** Expression Profiling using Microarrays Fabricated by an Ink-jet Oligonucleotide Synthesizer, *Nat. Biotechnol.*, *19*(4), 342-347.

[11] Tao, H., Bausch, C., Richmond, C., Blattner, F.R. and Conway, T. **1999** Functional Genomics: Expression Analysis of *Escherichia coli* Growing on Minimal and Rich Media, *J. Bacteriol.*, *181*(20), 6425-6440.

[12] Washburn, M.P., Koller, A., Oshiro, G., Ulaszek, R.R., Plouffe, D., Deciu, C., Winzeler, E. and Yates, J.R., 3rd, **2003** Protein Pathway and Complex Clustering of Correlated mRNA and Protein Expression Analyses in *Saccharomyces cerevisiae*, *Proc. Natl. Acad. Sci. U S A*, *100*(6), 3107-3112.

[13] Wang, R., Guegler, K., LaBrie, S.T. and Crawford, N.M. **2000** Genomic Analysis of a Nutrient Response in Arabidopsis Reveals Diverse Expression Patterns and Novel Metabolic and Potential Regulatory Genes Induced by Nitrate, *Plant Cell*, *12*(8), 1491-1509.

[14] Arnold, C.N., McElhanon, J., Lee, A., Leonhart, R. and Siegele, D.A. **2001** Global Analysis of *Escherichia coli* Gene Expression during the Acetate-induced Acid Tolerance Response, *J. Bacteriol.*, *183*(7), 2178-2186.

[15] Gasch, A.P., Huang, M., Metzner, S., Botstein, D., Elledge, S.J. and Brown, P.O. **2001** Genomic Expression Responses to DNA-damaging Agents and the Regulatory Role of the Yeast ATR Homolog Mec1p, *Mol. Biol. Cell*, *12*(10), 2987-3003.

[16] Zheng, M., Wang, X., Templeton, L.J., Smulski, D.R., LaRossa, R.A. and Storz, G. **2001** DNA Microarray-mediated Transcriptional Profiling of the *Escherichia coli* Response to Hydrogen Peroxide, *J. Bacteriol.*, *183*(15), 4562-4570.

[17] Perou, C.M., Jeffrey, S.S., van de Rijn, M., Rees, C.A., Eisen, M.B., Ross, D.T., Pergamenschikov, A., Williams, C.F., Zhu, S.X., Lee, J.C., Lashkari, D., Shalon, D., Brown, P.O. and Botstein, D. **1999** Distinctive Gene Expression Patterns in Human Mammary Epithelial Cells and Breast Cancers, *Proc. Natl. Acad. Sci. USA*, *96*(16), 9212-9217.

[18] DeRisi, J.L., Iyer, V.R. and Brown, P.O. **1997** Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale, *Science*, *278*(5338), 680-686.

[19] Marton, M.J., DeRisi, J.L., Bennett, H.A., Iyer, V.R., Meyer, M.R., Roberts, C.J., Stoughton, R., Burchard, J., Slade, D., Dai, H., Bassett, D.E., Jr., Hartwell, L.H., Brown, P.O. and Friend, S.H. **1998** Drug Target Validation and Identification of Secondary Drug Target Effects using DNA Microarrays, *Nat. Med.*, *4*(11), 1293-1301.