New Insights of DNA Microarray

Woo Chi

DNA Microarray is a recent molecular biotechnology, which can measure vast amount of gene expressions simultaneously (Schena et al., 1995). Microarray technology has been considered the most standout method for examining global aspects genomic data and gene expression profiling from the microarray data essentially added another dimension in bioinformatics with its infinite possibilities of expression change comparison in various conditions. With on-going DNA sequencing endeavors of many different model organisms including humans, now easily available DNA sequencing information can be used for monitoring almost the entire, if not all, gene expression of those organisms using the microarray.

Microarray takes advantage of two basic technologies. One is binding between single stranded DNA sequence with its complementary sequence. Another is using fluorescent probe to visualize difference in cDNA level which in tern represent mRNA level. A typical microarray would consist of a regular microscope glass slide that contains thousands of microscopic quantities of PCR products of cDNA (the Stanford method) or synthetic oligonuclotides of genes. Each spot should represents one specific exon or one gene in the genome. Thanks to recent development of high-speed robotic printing, once all the nucleotide sequences are ready, mass production of microarray slide is now possible for many different experiments.

Preparation of probe DNA usually includes use of fluorescence dyes such as Cy3 (green, but red under visible light) and Cy5 (red, but blue under visible light). These dyes are used to distinguish cDNA pools from reverse transcription of two different population of mRNA. For example, one can use Cy5 to label RNA from unaffected cells and Cy3 for affected or diseased cells (Figure 1).In microarray, it is the competitive hybridization of the differently labeled collection of cDNAs that gives us useful information. It is the relative intensities of Cy5/Cy3 probes that generate different colors in each spot. Thus, from the fluorescence intensities and colours for each spot, the relative expression levels of the genes in both samples can be estimated. It is quite important, however, that red and green light detection channel to be normalized so that two data sets can be compatible for intensity ratio-calculation. Analyzing thousands of spots with different intensity is an enormous task that requires power of computers. Many different image-processing software have been developed that allows scientist analyze and process their data fast and efficiently. Perhaps microarray is useless without the data analyzing softwares. List of thousands of genes with their expression profile then further processed to look for up and down regulation of genes with significance (usually 2 fold or higher expression change). These gene lists are the potentially important transcriptional targets or expression phenotype of disease.

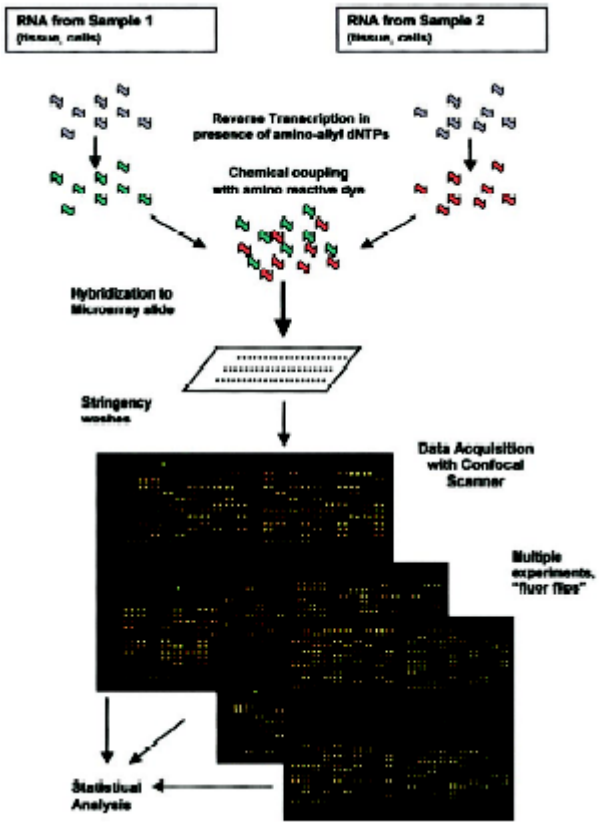

Fig. 1. Schema of a DNA microarray experiment showing how RNA from 2 samples are reverse transcribed and labeled with different fluors, hybridized to the microarrays slide, washed, and scanned to acquire data. The tiled scan images indicate the necessity for replicates and multiple experiments, producing data amenable to analysis.

Figure 1 adopted from (Panlilio et al)

Application of Microarray

As alluded before, microarray technology is used to answer the question like what genes are expressed in a particular cell type of an organism, at a particular time, under particular conditions. This is very useful when comparing mutant or diseased cells with unaffected cells and look for differences in gene expression that can be the potentially critical factor leading to disease. The most prominent study example with microarray seems to be cancer research. DNA microarrays has been used to virtually all different types of cancer research. Many of the researchers are looking at gene expression pattern in different developmental stages in cancer, such as in breast cancer (Sgroi et al), and prostate cancer (Luo et al) and trying to isolate key genes that are critical for pathway of tumor genesis. Microarray is also being used for screening genes that shows sensitivity to particular cancer drug (Zembutsu et al). It is conceivable that in the future, microarray will be standard cancer screening or detection method. The study of Raffeld and colleagues already showed identification of a tumor-specific marker in lymphoid neoplasms, setting the milestone of enormous potential of gene array technologies for diagnostic marker discovery (Wellmann et al). With increasing number of human cDNA and EST available on website, more global aspects of disease gene pattern should be identified soon.

Microarray technology is useful not only in clinical field, but also in basic science research. Development biology in particular, has been benefiting of this technology. Many developmental stages of an organism are regulated by differential gene expression, directing cells to behave appropriately at each stage. Many groups are studying model organisms and try to understand the complex developmental pathways that are under direct influence of genetic network. Kevin White from Yale had studied differential gene expression of *Drosophila melanogaster* during

metamorphosis with utility of DNA microarray (White et al). His group was able to generate 35 different clusters of genes that behave similar ways during metamorphosis and identified possible unknown gene function that may be involved in the control and execution of the process. Valerie Reinke, also from Yale used microarrays to compare various germline mutants with wild type in *C. elegans* and identified 1416 germline-enriched genes, many of which were previously unknown (Reinke et al). Although definite gene function should be further tested and confirmed with other method, microarray can certainly provide direction of what we can expect about the gene.

Microarray technology also has been applied to monitor cellular metabolic changes including aging. Although it has been speculated what factors might influencing the aging process to occur, the mechanisms that underlie longevity of cells and the age-related phenotypes are not well understood. In a recent study by Prolla and colleagues, high-density oligonucleotide arrays were used to analyze the gene expression profile of the aging process in skeletal muscle of mice. In their study, they found aged skeletal muscle in mouse showed similar array profile to that of stress response, coupled with lower metabolic and biosynthetic gene expression. Interestingly, these alterations were reduced with the aged mice under caloric restriction.

Shortcomings of Using Microarray

As with all other technologies, microarray is not a perfect system. Perhaps the biggest shortcoming of microarray is that mRNA level does not necessarily correlates with its protein level. Therefore, it can be misleading to think that high expression of a gene is a indication that the cell needs the activity of gene product at that specific time. Another concern is cross-hybridization of sequences with high sequence identity, multicopy genes in organism. For gene expression studies, each spot ideally should identify one gene or one exon in the genome, however, in practice this is not always so clear-cut due to families of similar genes in a genome witch can give false information about gene expression. Usually, microarray experiments are repeated several times and only the consistent expression pattern are considered to be reliable. While it is relatively easy to do a microarray experiment, data analysis after the experiment is another story. Even with help of computer, it takes fair amount of time and effort to really understand what the array tells. This has been major rate limiting step of getting the informative results.

Although it may not be perfect system yet, extra care in sequence selection and handling of array, clear experimental design and experimental annotation should substantially improve the quality of microarray data.

Future of Microarray

The field of microarray continues to be increasing and becomes more comprehensive. In the near future, complete DNA microarray chip of human will be available and that would likely bring huge impact on field of medicine and society in general. Conventional method of detection of disease, coupled with in depth, comprehensive molecular level of screening using microarray should revolutionize preventative medicine. With its ability to monitor even single gene expression level, sub typing disease according to their specific pattern of gene expression might be very plausible defining new classification of disease. This will also push drug markets to design and develop drugs that can target key gene activity according to very specific type of disease. Different drugs might be applied to individuals with same disease but showing different gene expression, or drug sensitivity. In other words, individualized drug will be available with simple microarray test.

For non-clinical science, microarray would be big help generating many different RNA expression profiles, clustering of gene function in time and space to generate more meaningful hypotheses and models that can be tested. Predicting unknown gene function and finding out various biochemical, metabolic, and cellular pathways using microarray is well on its way. Global

perspective studies will not only widen many scientists' scope of studying organism, but also help understanding organism in big picture.

DNA microarray might not be the only type of microarray in the future. Recent success of using yeast protein on glass slide instead of DNA began the new era of proteome. This protein microarrays can be used to screen protein-drug interactions and to detect posttranslational modifications (Zhu et al). The post genomic era should certainly be enjoyable with new techniques like microarray. Only the enormous degree of impact is awaiting us.

Selected Reading

Lee CK, Klopp RG, Weindruch R, Prolla TA. (1999) Science Aug 27;285(5432):1390-3

Luo, J., Duggan, D. J., Chen, Y., Sauvageot, J., Ewing, C. M., Bittner, M. L., Trent, J. M., Isaacs, W. B. (2001). Human Prostate Cancer and Benign Prostatic Hyperplasia: Molecular Dissection by Gene Expression Profiling. *Cancer Res* 61: 4683-4688

Panlilio, M., Robbins, S. (2002) *Am J Physiol Gastrointest Liver Physiol* 282: G397–G402.

Reinke V, Smith HE, Nance J, Wang J, Van Doren C, Begley R, Jones SJ, Davis EB, Scherer S, Ward S, Kim SK. (2000) Mol Cell Sep;6(3):605-16

Schena, A., Shalon, D., Davis, W. R., Brown P. O. (1995). Science 270, 467-470.

Sgroi, D. C., Teng, S., Robinson, G., LeVangie, R., Hudson, J. R. Jr., Elkahloun, A. G. (1999). In Vivo Gene Expression Profile Analysis of Human Breast Cancer Progression. *Cancer Res* 59: 5656-5661

Wellmann, A., Thieblemont, C., Pittaluga, S., Sakai, A., Jaffe, E. S., Siebert, P., Raffeld, M. (2000). Detection of differentially expressed genes in lymphomas using cDNA arrays: identification of clusterin as a new diagnostic marker for anaplastic large-cell lymphomas. *Blood* 96: 398-404

White KP, Rifkin SA, Hurban P, Hogness DS. (1999) Science Dec 10;286(5447):2179-84

Zembutsu H, Ohnishi Y, Tsunoda T, Furukawa Y, Katagiri T, Ueyama Y, Tamaoki N, Nomura T, Kitahara O, Yanagawa R, Hirata K, Nakamura Y. **(**2002) Cancer Res Jan 15;62(2):518-27

Zhu H, Bilgin M, Bangham R, Hall D, Casamayor A, Bertone P, Lan N, Jansen R, Bidlingmaier S, Houfek T, Mitchell T, Miller P, Dean RA, Gerstein M, Snyder M. (2001) *Science* Sep 14: 2101-2105.