Acta Med. Okayama, 2001
Vol. 55, No. 6, pp. 319-328

Acta Medica Okayama

http://www.lib.okayama-u.ac.jp/www/acta/

Review

Recent Advances in DNA Microarrays

Tohru Nakanishi[a]*, Takashi Oka[b], and Tadaatsu Akagi[b]

[a]Department of Biochemistry and Molecular Dentistry, and
[b]Department of Pathology, Okayama University Graduate School of Medicine and Dentistry, Okayama 700-8525, Japan

The structure of the human genome is almost completely elucidated and the life sciences will now aim for a general and integrated study of gene expressions and the functional elucidation of proteins. In such a study, various new techniques have been developed, and DNA microarray technology is the most representative one. As for the DNA microarray techniques, several thousands to tens of thousands of gene segments are immobilized on a glass slide at high density, and cDNA probes prepared from specific cells or tissues are hybridized on the slides from which gene expression profiles are obtained at one sweep in a short time. The present development of this technique and its possible application to medicine-related fields are described.

Key words: DNA microarray, DNA chip, human genome, embryonic stem (ES) cell, single nucleotide polymorphism (SNP)

Background of DNA Microarray: Aspects of Post-Human Genome Analysis

Before describing DNA microarrays, some comments are needed to address the background of DNA microarrays in relation to the aspects of post-genome analysis.

In "Nature" and "Science" journals, the organization of the chromosomes deduced from the draft sequence of human genome was recently announced. In "Nature" which has become the main method of announcements from public institutes, a special issue of the human genome project was released with posters and a CD-ROM, and this historical achievement was immortalized [1]. In "Science", a special issue was also released, and the Celera genome assembly was described [2]. However, inconsistency in the way of thinking about possession of gene sequence information between public bodies and private enterprise comes to the fore.

It is clear that these results and information will become clues towards epoch-making individual disease therapy, but it is necessary to recognize that some problems may develop including commercial profit making and personal privacy that are easily connected in the sorting and dominance theory. In other words, it may lead to political problems that may include the interests of countries. Under these conditions, it will become the next major problem to advance such studies regarding the information of genes (transcriptomes) expected from these sequences and the functional information of the deduced proteins (proteomes).

Naturally, many researchers and companies are currently entering this field, and it is suggested that a study that obtains general and fundamental information occupies the mainstream of life science. Genome information tied to the function of proteins can produce vast profits and benefits. Although the situation concerning the establish-

Received May 22, 2001; accepted June 5, 2001.
*Corresponding author. Phone: +81-86-235-6646; Fax: +81-86-235-6649
E-mail: naka3285@md.okayama-u.ac.jp (T. Nakanishi)

ment of embryonic stem (ES) cells, especially those of human origin are apparently still confusing [3-5], it will soon become possible, as for the therapy of a disease, that mutated or unsuitable genes are corrected to normal or more suitable genes by gene therapy before birth or before implantation at the level of embryos or ES cells. However, it will also become theoretically possible to make genetically manipulated humans as portrayed in some novels or movies.

Under these conditions, the orientation of life science studies will be controlled by company profits and social or political events in addtion to a researcher's personal will, and, thus, researchers should pay attentions to these problems described above. Therefore, it will also become important that researchers explain their works and their thoughts widely to the public at large.

The Principle of DNA Microarrays

In a promotion of the general and integrated study described above, it will be a key point to develop a technique that produces large quantities of information with high speed and high efficiency. DNA microarray is a comprehensive gene expression analysis technique that has spread out rapidly since it was developed recently in the United States [6-8].

Information provided by this technique and the available speed are special because expressed cDNA probes are hybridized to many DNA fragments which are immobilized to a slide glass at one time, whereas the differential display method technique, to compare expressed gene fragments, is amplified from a limited number of cells or tissues. For example, the DNA microarray technique can be applied to characteristic analysis of cancer cells or an estimation of prognosis [9] so that information about many specimens, containing a large quantity of genes, is provided in a short time by arranging the profile, and the therapy, which fits the characteristics of the cancer cells, can be performed by classifications of the drug sensitivity of the cells using this technique. In addition, in DNA microarrays, single nucleotide polymorphisms (SNPs) can be detected in addition to gene expression information, and identification of SNPs of a specified gene or detection of the variation in a viral genome has already been done by using a microarray of the oligonucleotide type [10-12].

There are currently two following types of DNA microarrays that are commercially available (Table 1).

Table 1 Types of DNA microarrays

	Pasting type	Synthetic type
Developer	Stanford Univ.	Affymetrix
Methods	spotted on a slide glass	synthesized on a print board
DNA	cDNA fragments	oligonucleotides
Arrays	10,000-15,000/array	200,000-300,000/chip
Application	gene expression profiling	gene polymorphisms variations in oncogenes

One is a synthesis type DNA chip composed of oligonucleotides synthesized on the print board, and the other is a pasting type DNA microarray which carries cDNA fragments as hybridized arrays. The former is used for a gene expression profiling of tens of thousands of genes, and the latter is used for detection of gene polymorphysm such as SNPs and for detection of genomic mutations.

The former was developed as a slide glass pasting type DNA microarray by a group of Brown and Davis of Stanford University [13]. It is the system where a DNA solution is spotted on a slide glass using a DNA spotter, a kind of distributer with numerous pins that is already marketed as DNA chip production equipment. Although it has an upper array density limit (it has an upper array number limit, about 10,000 arrays per 4 cm^2 on a slide glass even if it is assumed that a spot with a diameter of 0.1 mm was displayed at 0.1 mm intervals), researchers can perform their analysis depending on their needs when they have a unique cDNA library and can make original DNA chips by spotting this library by themselves or by tip makers.

The latter was called DNA chips based on the image of IC tips of a semiconductor because DNA was synthesized on the print board of a semiconductor producted using a photolithography technique. At first, the paste type array was called DNA microarray and this latter type was called DNA tips, but currently the paste type array can also be called DNA chips. Extremely highly precise processing is possible using this method, and this synthesis type chip can produce density of the array to 0.01 mm^2 per spot using the latest print technology. In addition, it is effective for the analysis of gene polymorphisms or variations in oncogenes and viral genomes because detection of a difference of 1 bp becomes possible using this type of chip. Furthermore, there is an example which used this chip for analysis of gene expression (details will be described in the last chapter). However, there is one important point where attention should be paid in experi-

ments using this technique. It is caused by the high density characters of this method. This method is also expensive because of the exclusive production and marketing by the Affymetrix company.

The experiment principle of DNA microarrays is as follows. In the pasting type arrays, arbitrary cDNA library or PCR-amplified DNA fragments are spotted on a poly-L-lysine or silane-coated slide glass using a DNA spotter (spotting equipment). The quality of these microarrays depends on the spotting equipment, especially the top shape of the spotting pins that requires durability and accuracy of spotting volume without contamination. In addition, mRNA is extracted from samples such as a cells and tissues of which gene expression should be monitored, and fluorescent-labeled probes are prepared for hybridization to DNA microarrays. Recently, fluorescent labeling was performed using a double fluorescence system. In this method, two mRNA samples are labeled with different fluorescence (Cy3 and Cy5) and hybridized to the same array. The results of gene expression are analyzed by laser emission of different wavelengths, and compared to calculate the expression ratio (Fig. 1). After the hybridization, signals on the array spots are detected using an array scanner and analyzed en masse with exclusive software. There are two types of array scanners, the focal type and non-focal type. The former has an advantage in low background conditions and the latter has the advantage in precise reading. Table 2 shows a summary of websites of DNA microarrays currently available.

Although the findings obtained from expression profiles of the DNA microarray are enormous, it remains necessary to use many staff to obtain data from many samples, to analyze them statistically and search for useful genes all at high cost. An imbalance between the cost and staff required for analysis and the quality and value of the results to be provided is a problem of this current technique. Many researchers have already experienced this point even in experiments using a DNA array of lower density (microarray). In this respect, there

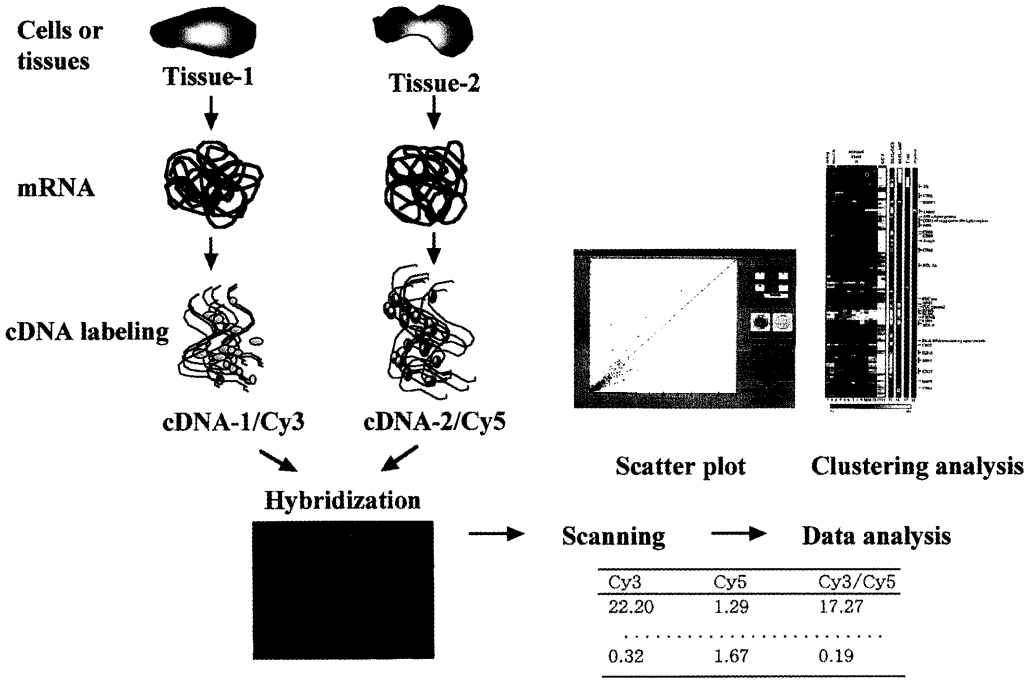

Fig. 1 cDNA microarray schema. Total RNA from both the test and reference tissues or cell samples is fluorescently labeled with either Cy3 or Cy5-dUTP using single round reverse transcription. The fluorescent probes are used for hybridization under stringent conditions to the clones on the microarray. Laser excitation of the hybridized probes yields an emission with characteristic spectra, which is measured using a scanning fluorescent detector. Monochrome images from the scanner are imported into computer software in which the images are pseudo-colored and merged. Data from a single hybridization is viewed as a normalized ratio (Cy3/Cy5) that indicates no change (~ 1), increase (>1) or decrease (<1) of gene expression relative to the control. Accumulated data can be used for the further analysis such as scatter plot or clustering analysis.

Talbe 2 Relational Web sites of DNA microarrays

Research Institutes	
Stanford Univ.	http://cmgm.stanford.edu/pbrown/array.html
Washington Univ.	http://chroma.mbt.washington.edu/mod_www/
NHGRI/NIH	http://www.nhgri.nih.gov/DIR/LCG/15K/HTML/
DNA microarrays	
Affymetrix	http://www.affymetrix.com/
Clontech	http://www.clontech.com/atlas/index.html
Stanford Genmic Resources	http://genome-www.stanford.edu/
Research Genetics	http://www.resgen.com/products/index.php3
Amersham pharmacia	http://www.jp.apbiotech.com
Kurabo	http://www.bio.kurabo.co.jp/
Toyobo	http://toyobo.co.jp/
Scanners, Softwares	
Axon	http://www.axon.com/
Silicon Genetics	http://www.sigenetics.com
Spotfire	http://www.spotfire.com
Scanalytics	http://www.scanalytics.com/
Molecular Dynamics	http://www.apbiotech.com/
Genetic Microsystems	http://www.geneticmicro.com/
General Scanning	http://www.genscan.com/
Gene Machines	http://www.genemachines.com/
BioDiscovery	http://www.biodiscovery.com/
BioRobotics	http://www.biorobotics.com/
Takara	http://www.takara.co.jp/

remains considerable interest in this method. These problems are expected to improve in the near future.

The Present Situation in the Research Applications of DNA Microarray

Diagnostic Gene Expression Profilings. Genome-wide analysis techniques such as chromosome painting [14], comparative genomic hybridization (CGH) [15], representational difference analysis (RDA) [16], restriction landmark genome scanning [17] and high-throughput analysis of loss of heterozygosity (LOH) [18] are now accelerating high-resolution genome aberration localizations in various diseases especially in human tumors. These techniques are complemented largely by cDNA expression microarray analysis. The most common application of microarray technology is transcript profiling. Microarrays are increasingly being used to characterize the differences in mRNA populations between tumors and normal cells. Such information will not only provide a signature of diagnostic purposes but also reveal new insights into the biology of tumorigenesis.

One of the most promising applications of cDNA microarray is in the classification of diseases especially cancer by gene expression profiles. The challenge of cancer treatment has been to target specific therapies to pathogenetically distinct tumor types, to maximize efficacy and minimize toxicity. One example was class discovery and class prediction of acute leukemia into those arising from lymphoid precursors (acute lymphoblastic leukemia, ALL) or from myeloid precursors (acute myeloid leukemia, AML) with cDNA expression array [19]. These findings demonstrate the feasibility of cancer classification based on gene expression monitoring and suggest a general strategy for discovering and predicting cancer classes for other types of cancer, independent of previous biological knowledge. In some cases, previous distinguishing characteristics of different classes and possibly not even the number of different classes was unknown. A particularly interesting example of class discovery was performed in the identification of two distinct forms of diffuse large B cell lymphoma (DLBCL), the most common subtype of non-Hodgkin's lymphoma [9]. This type of malignant lymphoma is clinically heterogeneous. It was proposed that variability in natural history reflects unrecognized molecular heterogeneity in the tumor and the two distinct forms of DLBCL were identified on the basis of gene expression patterns indicative of different stages of B-cell differentiation. One type expressed genes characteristic of germinal center B cells (germinal center B-like DLBCL); the second type expressed genes normally induced during *in vitro* activation of peripheral blood B cells (activated B-like DBCL). Remarkably, this molecular classification of tumors on the basis of gene expression can identify previously undetected and significant subtypes of cancer; patients with germinal center B-like DLBCL had a significantly better overall survival than those with activated B-like DLBCL. Although histology alone can be used to predict outcome for many cases of DLBCL, this gene expression data provides a framework for a more objective and comprehensive classification system that could be applied by pathologist everywhere. Importantly, cDNA expression array technique of class identification as well as SNP analysis can be applied to distinctions relating to future clinical outcome, such as drug response or survival. Class prediction provides an unbiased, general approach to constructing prognostic tests, provided a database of

tumor samples is available for which eventual outcome is known. In addition to hematopoietic malignancies, molecular classification of cutaneous malignant melanoma by cDNA expression profiling also have been reported, indicating that 31 melanoma patients could be divided into 2 distinct subgroups based on their gene expression profiles, even though there were no obvious pathological distinctions between the patients' tumors [20]. Various types of cancers have been investigated and ongoing with cDNA microarrays such as breast cancer [21, 22], colon cancer [23], bladder tumor [24, 25] hepatocellular carcinoma [26, 27], and rhabdomyosarcoma [28-30]. Systemic investigation of 60 human cancer cell lines have been performed in the National Cancer Institute's screen for anti-cancer drugs [31]. This kind of pharmacogenomic analysis should reveal gene-drug relationships for the clinical agents and clarify how variations in the transcription levels of particular genes relate to mechanisms of drug sensitivity and resistance [32, 33].

Understanding Pathogenesis. Microarrays promise to accelerate our understanding of the etiology of various kinds of diseases such as malignant tumors, inflammatory diseases, immune disorders, degenerative disorders and the host-pathogen interaction [34, 35]. Microarray technology has been used to identify signaling pathways regulated by key genes implicated in tumorigenesis as well as physiological responses [36]. Recently, we investigated the NK/T lymphoma, one of the most aggressive types of malignant lymphoma, with cDNA expression array to elucidate the mechanism of NK/T lymphomagenesis [37]. We found strong suppression of hematopoietic cell specific protein-tyrosine-phosphatase SH-PTP1 (SHP1) mRNA by cDNA expression array and RT-PCR. Further analysis with standard immunohistochemistry and tissue microarray, which utilized 207 paraffin-embedded specimens of various kinds of malignant lymphomas, showed that 100% of NK/T lymphoma specimens and more than 95% of various types of malignant lymphoma were negative for SHP1 protein expression. In addition, various kinds of hematopoietic cell lines-particularly the highly aggressive lymphoma/leukemia lines-lacked SHP1 expression *in vitro,* suggesting that loss of SHP1 expression may relate to not only malignant transformation, but also tumor cell aggressiveness. These findings suggest that loss of SHP1 gene expression plays an important role in multistep tumorigenesis, possibly as an anti-oncogene in the wide range of lymphomas/leukemias as well as NK/T lymphoma. In the field of cancer research, one of the next major challenges is the identification of molecular changes that progress cells to the metastatic state, with the ultimate purpose being to develop treatments which inhibit metastasis. Clark *et al.* used microarray to compare the gene expression profiles of highly metastatic melanoma cells with those of less metastatic cells from which they were derived. They found from these analyses that small GTPase RhoC is one of key molecules of metastasis and RhoC enhances metastasis when overexpressed, whereas a dominant negative Rho inhibits metastasis [38]. Harkin *et al.* carried out microarray based expression profiling to identify downstream transcriptional targets of BRCA1 tumor-suppressor gene, which is related to breast and ovarian cancer, as a means of defining its function [39]. Thereafter, BRCA1 was implicated in at least three functional pathways, namely, mediating the cellular response to DNA damage, as a cell cycle checkpoint protein and in the regulation of transcription [40]. Many applications of the microarray technique to clarify carcinogenesis are ongoing and should spread to various fields of life science [41-43].

The large-scale cDNA microarray or DNA chip methods are powerful and promising for the systematic and comprehensive analysis of gene expression profiles of known and unknown genes, genetic linkage studies with SNP, mutation detection and analysis with oligo arrays, analysis of disease responsible genes and so on. It is essential to carry out multiple independent experiments in a large cohort of samples to isolate biologically relevant changes and establish the diagnostic, prognostic and therapeutic importance of each of the many oncogene or anti-oncogene candidates from spurious results that may arise as a result of genetic heterogeneity between individuals. To do that with cDNA microarrays will incur high costs. To overcome this difficulty, an array-based high-throughput tissue microarray technique has been established to facilitate analysis of *in situ* hybridization, immunohistochemical staining and DNA analysis using a micro-dissecting method from large numbers of tissue specimens [44-46]. Presently, a combination of the low cost middle-scale cDNA macroarray and tissue microarray is more realistic and also more informative. This strategy, which was used in our previous investigation [37], may have a great advantage especially in the rapid and easy identification of the disease-responsible or closely related genes with quick screening of multiple genes and multiple tumor types under the same condition, thereby

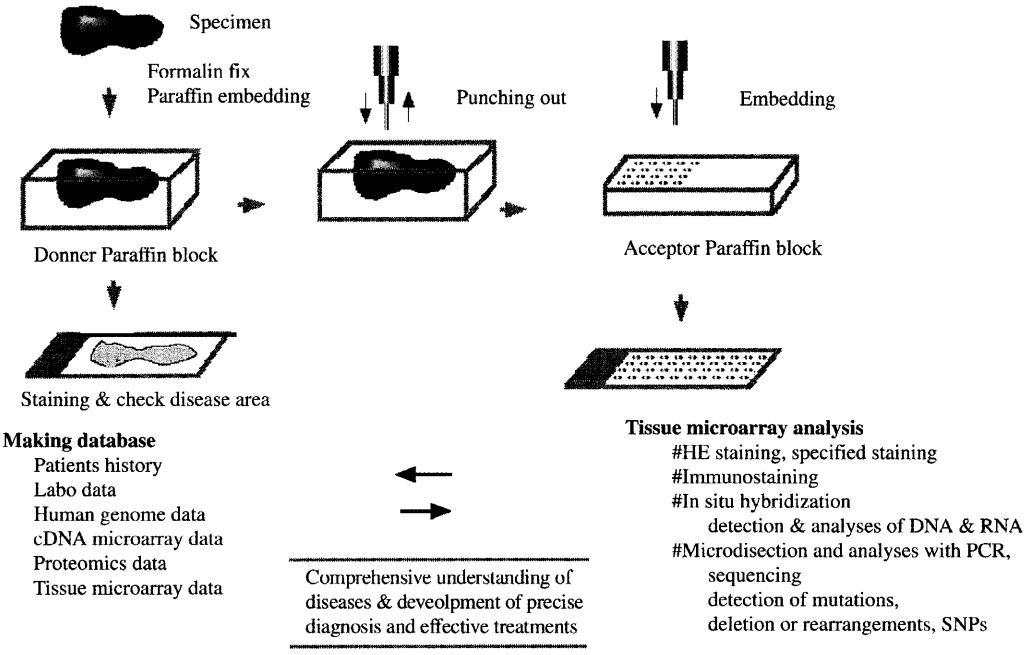

Fig. 2 Schematic diagram of tissue microarray. Core tissue biopsies (diameter, 0.8mm) are taken from selected regions of individual paraffin-embedded specimens (donor block) and precisely arrayed into a new recipient paraffin block (50mm×23mm) using a custom-built instrument. After the block construction is completed, 5μm sections of the resulting tissue microarray block are cut with a microtome. These specimens are used for hematoxylin-eosin staining, immunostaining, *in situ* hybridization and sequencing for analysis and detection of mutation, deletion or rearrangements, SNPs using microdissection. Tissue microarray could be used for the rapid identification of specific changes of mRNA and protein expression with the large number of specimen of the focused diseases followed by the detailed analysis of genetic abnormalities in the specific cells of the specimen using micro-dissection technique.

leading to a more unbiased reliable analysis (Fig. 2). This approach is useful for identification of genes from various types of diseases where a particular molecular alteration is very important.

Cell cycle analysis with microarrays. In 1998, the availability of the yeast whole genome sequence allowed genome-wide surveys for cell cycle regulated genes with cDNA microarrays. The cell populations are synchronized by arresting them at a homogeneous cell cycle state, then releasing from the arrested state and sampling at various time intervals throughout cell cycles. Cho *et al.* [47] and Spellman *et al.* [48] found at least 800 genes, which was more than 10% of the total number of genes in the yeast genome, that oscillated at the mRNA level as a function of the mitotic cell cycle position. Spellman *et al.* used the hierarchical clustering method to analyze 800 cell cycle regulated genes [48, 49]. This approach revealed about 10 convincing gene-clusters of cell cycle regulated genes encompassing about 400 genes. The behavior of these genes in the cell cycle can be found at this website (http: //genome-www.stanford.edu/cellcycle/). Related studies examined gene expression in the meiotic cell cycle [50] and the gene expression in responce to treatment with alpha-factor [51].

Proteomics: New biomedical perspectives. In the 20 th century, massive progress was made in the knowledge of the processes underlying the genetic basis of humans, ranging from the structure of DNA to the identification of disease-specific genetic abnormalities, and their clinical application as the basis of trials of gene therapy or prognosis markers. Presently, the identification and sequencing of the 40,000 to 60,000 genes of the human genome in addition to those of several other species have been or are almost accomplished. The next phase of the projects concerning the structural and functional genomics and proteomics have begun. Proteins are functional outputs of the cell and might be expected to provide very important information. There is a poor correlation between mRNAs and their respective proteins, generally lower than 0.5 [52], because degradation rates of individual mRNAs and proteins differ and transcript

can be spliced in various ways to produce different proteins. In addition, extensive modification can be introduced during and after translation: *e.g.* phosphorylation, glycosylation, lipid attach ment, peptide cleavage, complex formation, and translocation within the cell. Thus, the challenge is to develop high throughput approaches to systematic and global protein analysis to conduct the functional classification of proteins into a biological context. The array format is now an established method for global analysis of cDNA expression and has recently been adapted for protein investigations. The peptide and protein arrays described so far can be divided either chemical/non-living arrays, composed of synthetic peptides using combinatorial synthesis/purified proteins [53, 54], or biological/living arrays, in which peptides and proteins are expressed in the living organisms, such as phage display method [55], modified two-hybrid system of budding yeast [56, 57] or the *in vitro* virus method [58]. New visions of proteomics need to develop technologies for ultra-high throughput analysis of at least the majority or all (if possible) of the proteins in the cells. Our interests in proteomics are not only for the identification of proteins in the cells but also their functional state, posttranslational modification and the dynamic changes in the rates of synthesis and turnover. These approaches should be supported by the various technological advances including mass spectrometry, particularly matrix-assisted laser desorption mass spectrometry (MALDI), electrospray ionization tandem mass spectrometry (ESI-MS/MS), the interface of these methods with gel-based separation methods, high-speed data-processing routines and data analysis argorithms based on correlating peptide masses and collision-induced dissociation spectra obtained from mass spectrometry with predicted proteins/peptides from the DNA sequence database. Proteomics promises exciting insights into the working mechanisms of the cells or living organisms. However many technical hurdles remain to be overcome.

The Future Prospects of DNA Microarray

As mentioned above, the reputation of the innovative technology advances in the case of this DNA microarray technology, and there are many current cases which have not yet reached a useful result because the result is too confused and unreliable. In the field of clinical testing in particular, the characteristics of this technique; that it brings enormous expression information of more than 10,000 different genes in a short time has attracted much attention. Unfortunately, users of these instruments often feel that this technology is still at a stage of development and also feel that the price, quality and value of the results obtained from these instruments are not satisfactory.

Current new developments can be seen on the following websites (Table 2). Materials of DNA carriers or fixing methods will be improved. In addition, various improvements in the methods of data analysis have been reported. Using this microarray system effectively, there are some trials to integrate the analyzed data of allied researchers (KEGG or CELL) [59].

Among the recent applications of ongoing research, applications to cancer research is the most noteworthy. The DNA microarray has been applied to the classification of cancer, a search of oncogenes, and further advances as described above, and the integration of databases regarding the expression profiles and drug sensitivities of cancer cells are now ongoing in Japan using the arrays of more than 30,000 genes. This information will be useful for predicting the drug sensitivity of cancer cells and the development of new drugs. In addition, analysis technology to measure changes in chromosome copy numbers by DNA microarray using mRNA (cDNA) extracted from neoplasm as probes (CGH: comparative genomic hybridization) has been developed, and it was examined whether information about abnormal copy numbers was useful for diagnosis of cancer or a prediction of therapy prognosis [60, 61].

In addition, applications of this technique have spread explosively into the study of development and regeneration. For example, using an array made by mRNA extracted from early phase embryos of mice, a gene expression pattern in embryogenesis will be analyzed. Microarray analysis of gene expression in many functional stem cells and in the process of their differentiation will also enable isolation or characterization of multi-functional stem cells from the adult body at the gene level [62]. Comparing with the case of other typical methods of expression profiling by SAGE [63], DNA microarray requires comparatively large amounts of RNA but it has the advantage of speed of analysis that is brought by omitting DNA sequencing. Therefore, the DNA microarray will be expected to produce a greater impact in the field of tissue engineering in the near future.

Furthermore, this technology can be applied to protein and tissue microarrays as described above. For example, the proteome chip was developed on which

antibodies are immobilized by plasma polymerization methods to measure the protein that specifically appears in cancer organization, and also the chips which immobilize various receptors are also being developed, and they will be applied to clinical diagnosis of cancer cells.

The chip technology that will accumulate as described above will greatly influence trends in the life science in the future, but it is noteworthy that the United States is preceding with patent fields with applications in the technological development and industrial applications of this technology, where a specified company or university will have superiority over others. Therefore, the development of future chip technology depends on the original idea that will be planned by these major leaders and especially other companies that follow these leaders. Chip technology will become one of the big trends in life science of 21 th century.

References

1. The human genome. Nature (2001) **409**, 813-958.
2. The human genome. Science (2001) **291**, 1279-1351.
3. Nakanishi T: Japan may seek embryo cells from overseas. Nature (2000) **405**, 882.
4. Dickson D: European panel rejects creation of human embryos for research. Nature (2000) **408**, 277.
5. Antoniou M: Embryonic stem cell research. The case against. Nat Med (2001) **7**, 397-399.
6. Marshall A and Hodgson J: DNA chips: An array of possibities. Nat Biotechnol (1998) **16**, 27-31.
7. Ramsay G: DNA chips: State-of-the art. Nat Biotechnol (1998) **16**, 40-44.
8. Editorial: Getting hip to the chip. Nat Genet (1998) **18**, 195-197.
9. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R and Staudt LM: Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature (2000) **403**, 503-511.
10. Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, Shaw N, Lane CR, Lim EP, Kalyanaraman N, Nemesh J, Ziaugra L, Friedland L, Rolfe A, Warrington J, Lipshutz R, Daley GQ and Lander ES: Characterization of single-nucleotide polymorphisms in coding regions of human genes. Nat Genet (1999) **22**, 231-238.
11. Halushka MK, Fan JB, Bentley K, Hsie L, Shen N, Weder A, Cooper R, Lipshutz R and Chakravarti A: Patterns of single-nucleotide polymorphisms in candidate genes for blood-pressure homeostasis. Nat Genet (1999) **22**, 239-247.
12. Kozal MJ, Shah N, Shen N, Yang R, Fucini R, Merigan TC, Richman DD, Morris D, Hubbell E, Chee M and Gingeras TR: Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. Nat Med (1996) **2**, 753-759.
13. Schena M, Shalon D, Davis RW and Brown PO: Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science (1995) **270**, 467-470.
14. Verdorfer I, Neubauer S, Letzel S, Angerer J, Arutyunyan R, Martus P, Wucherer M and Gebhart E: Chromosome painting for cytogenetic monitoring of occupationally exposed and non-exposed groups of human individuals. Mutat Res (2001) **491**, 97-109.
15. Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F and Pinkel D: Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science (1992) **258**, 818-821.
16. Lisitsyn N and Wigler M: Cloning the differences between two complex genomes. Science (1993) **259**, 946-951.
17. Rickman DS, Tyagi R, Zhu XX, Bobek MP, Song S, Blaivas M, Misek DE, Israel MA, Kurnit DM, Ross DA, Kish PE and Hanash SM: The gene for the axonal cell adhesion molecule TAX-1 is amplified and aberrantly expressed in malignant gliomas. Cancer Res (2001) **61**, 2162-2168.
18. Liloglou T, Maloney P, Xinarianos G, Fear S and Field JK: Sensitivity and limitations of high throughput fluorescent microsatellite analysis for the detection of allelic imbalance: Application in lung tumors. Int J Oncol (2000) **16**, 5-14.
19. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD and Lander ES. Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. Science (1999) **286**, 531-537.
20. Bittner M, Meltzer P, Chen Y, Jiang Y, Seftor E, Hendrix M, Radmacher M, Simon R, Yakhini Z, Ben-Dor A, Sampas N, Dougherty E, Wang E, Marincola F, Gooden C, Lueders J, Glatfelter A, Pollock P, Carpten J, Gillanders E, Leja D, Dietrich K, Beaudry C, Berens M, Alberts D and Sondak V: Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature (2000) **406**, 536-40.
21. Perou CM, Jeffrey SS, van de Rijn M, Rees CA, Eisen MB, Ross DT, Pergamenschikov A, Williams CF, Zhu SX, Lee JC, Lashkari D, Shalon D, Brown PO and Botstein D: Distinct gene expression patterns in human mammary epithelial cells and breast cancers. Proc Natl Acad Sci USA (1999) **96**, 9212-9217.
22. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO and Botstein D: Molecular portraits of human breast tumors. Nature (2000) **406**, 747-752.
23. Alon U, Barkai N, Notterman DA, Gish K, Ybarra S, Mack D and Levine AJ: Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Pro Natl Acad Sci USA (1999) **96**, 6745-6750.
24. Chen CC, Shieh B, Jin YT, Liau YE, Huang CH, Liou JT, Wu LW, Huang W, Young KC, Lai MD, Liu HS and Li C: Microarray profiling of gene expression patterns in bladder tumor cells treated with genistein. J Biomed Sci (2001) **8**, 214-222.
25. Wikman FP, Lu ML, Thykjaer T, Olesen SH, Andersen LD, Cordon-Cardo C and Orntoft TF: Evaluation of the performance of a p53 sequencing microarray chip using 140 previously sequenced bladder tumor samples. Clin Chem (2000) **46**, 1555-1561.
26. Shirota Y, Kaneko S, Honda M, Kawai HF and Kobayashi K: Identification of differentially expressed genes in hepatocellular carcinoma with cDNA microarrays. Hepatology (2001) **33**, 832-840.
27. Okabe H, Satoh S, Kato T, Kitahara O, Yanagawa R, Yamaoka Y, Tsunoda T, Furukawa Y and Nakamura Y: Genome-wide analysis of gene expression in human hepatocellular carcinomas using cDNA microarray: Identification of genes involved in viral carcinogenesis and tumor progression. Cancer Res (2001) **61**, 2129-2137.

28. Khan J, Simon R, Bittner M, Chen Y, Leighton SB, Pohida T, Smith PD, Jiang Y, Gooden GC, Trent JM and Meltzer PS: Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. Cancer Res (1998) **58**, 5009–5013.
29. Pandita A, Zielenska M, Thorner P, Bayani J, Godbout R, Greenberg M and Squire JA: Application of comparative genomic hybridization, spectral karyotyping, and microarray analysis in the identification of subtype-specific patterns of genomic changes in rhabdomyosarcoma. Neoplasia (1999) **1**, 262–275.
30. Scholl FA, Betts DR, Niggli FK and Schafer BW: Molecular features of a human rhabdomyosarcoma cell line with spontaneous metastatic progression. Br J Cancer (2000) **82**, 1239–1245.
31. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, Iyer V, Jeffrey SS, Van de Rijn M, Waltham M, Pergamenschikov A, Lee JC, Lashkari D, Shalon D, Myers TG, Weinstein JN, Botstein D and Brown PO: Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet (2000) **24**, 227–235.
32. Scherf U, Ross DT, Waltham M, Smith LH, Lee JK, Tanabe L, Kohn KW, Reinhold WC, Myers TG, Andrews DT, Scudiero DA, Eisen MB, Sausville EA, Pommier Y, Botstein D, Brown PO and Weinstein JN: A gene expression database for the molecular pharmacology of cancer. Nat Genet (2000), **24**, 236–244.
33. Debouck C and Goodfellow PN: DNA microarray in drug discovery and development. Nat Genet (1999) **21**, 48–50.
34. Manger ID and Relma DA: How the host 'sees' pathogenes: Global expression responses to infection. Curr Opin Immunol (2000) **12**, 215–218.
35. Cummings CA and Relman DA: Using DNA microarrays to study host-microbe interactions. Emerg Infect Dis (2000) **6**, 513–525.
36. Iyer VR, Eisen MB, Ross DT, Schuler G, Moore T, Lee JC, Trent JM, Staudt LM, Hudson J Jr, Boguski MS, Lashkari D, Shalon D, Botstein D and Brown PO: The transcriptional program in the response of human fibroblasts to serum. Science (1999) **283**, 83–87.
37. Oka T, Yoshino T, Hayashi K, Ohara N, Nakanishi T, Yamaai Y, Hiraki A, Kondo E, Teramoto N, Takahashi K, Tsuchiyama J and Akagi T: Reduction of hematopoietic cell-specific tyrosine phosphatase SHP-1 gene expression in natural killer cell lymphoma and various types of lymphomas/leukemias: combination analysis with cDNA expression array and tissue microarray. Am J Pathol (2001) **159**, 1495–1505.
38. Clark EA, Golub TR, Lander ES and Hynes RO: Genomic analysis of metastasis reveals an essential role for RhoC. Nature (2000) **406**, 532–535.
39. Harkin DP, Bean JM, Miklos D, Song YH, Truong VB, Englert C, Christians FC, Ellisen LW, Maheswaran S, Oliner JD and Haber DA: Induction of GADD45 and JNK/SAPK-dependent apoptosis following inducible expression of BRCA1. Cell (1999) **97**, 575–586.
40. Harkin DP: Uncovering functionally relevant signaling pathways using microarray-based expression profiling. Oncologist (2000) **5**, 501–507.
41. Feske S, Giltnane J, Dolmetsch R, Staudt LM and Rao A: Gene regulation mediated by calcium signals in T lymphocytes. Nat Immunol (2001) **2**, 316–324.
42. Coller HA, Grandori C, Tamayo P, Colbert T, Lander ES, Eisenman RN and Golub TR: Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion. Proc Natl Acad Sci USA (2000) **97**, 3260–3265.
43. Shaffer AL, Yu X, He Y, Boldrick J, Chan EP and Staudt LM: BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control. Immunity (2000) **13**, 199–212.
44. Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G and Kallioniemi O: Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med (1998) **4**, 844–847.
45. Schraml P, Kononen J, Bubendorf L, Moch H, Bissig, H, Nocito A, Mihatsch MJ, Kallioniemi OP and Sauter G: Tissue microarrays for gene amplification surveys in many different tumor types. Clin Cancer Res (1999) **5**, 1966–1975.
46. Moch H, Schraml P, Bubendorf L, Mirlacher M, Kononen J, Gasser T, Mihatsch MJ, Kallioniemi OP and Sauter G: High-throughput tissue microarray analysis to evaluate genes uncovered by cDNA microarray screening in renal cell carcinoma. Am J Pathol (1999) **154**, 981–986.
47. Cho RJ, Campbell MJ, Winzeler EA, Steinmetz L, Conway A, Wodicka L, Wolfsberg TG, Gabrielian AE, Landsman D, Lockhart DJ and Davis RW: A genome-wide transcriptional analysis of the mitotic cell cycle. Mol Cell (1998) **2**, 65–73.
48. Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, Brown PO, Botstein D and Futcher B: Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol Biol Cell (1998) **9**, 3273–3297.
49. Eisen MB, Spellman PT, Brown PO and Botstein D: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA (1998) **95**, 14863–14868.
50. Chu S, DeRisi J, Eisen M, Mulholland J, Botstein D, Brown PO and Herskowitz I: The transcriptional program of sporulation in budding yeast. Science (2000) **282**, 699–705.
51. Spellman PT, Sherlock G, Zhang MQ, Iyer VR, Anders K, Eisen MB, Brown PO, Botstein D and Futcher B: Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization. Mol Biol Cell (1998) **9**, 3273–3297.
52. Abbott A: A post-genomic challenge: Learning to read patterns of protein synthesis. Nature (1999) **402**, 715–720.
53. Kramer A and Schneider-Mergener J: Synthesis and screening of peptide libraries on continuous cellulose membrane supports. Methods Mol Biol (1998) **87**, 25–39.
54. Reineke U, Sabat R, Kramer A, Stigler RD, Seifert M, Michel T, Volk HD and Schneider-Mergener J: Mapping protein-protein contact sites using cellulose-bound peptide scans. Mol Divers (1996) **1**,141–148.
55. Zozulya S, Lioubin M, Hill RJ, Abram C and Gishizky ML: Mapping signal transduction pathways by phage display. Nat Biotechnol (1999) **17**, 1193–1198.
56. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P, Qureshi-Emili A, Li Y, Godwin B, Conover D, Kalbfleisch T, Vijayadamodar G, Yang M, Johnston M, Fields S and Rothberg JM: A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae. Nature (2000) **403**, 623–627.
57. Ito T, Tashiro K, Muta S, Ozawa R, Chiba T, Nishizawa M, Yamamoto K, Kuhara S and Sakaki Y: Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. Proc Natl Acad Sci USA (2000) **97**, 1143–1147.
58. Nemoto N, Miyamoto-Sato E, Husimi Y and Yanagawa H: *In vitro* virus: Bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome *in vitro*. FEBS Lett (1997) **414**, 405–408.
59. Kanehisa M and Goto S: KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res (2000) **28**, 27–30.
60. Pinkel D, Segraves R, Sudar D, Clarks S, Poole I, Kowbel D, Collins C, Kuo W-L, Chen C, Zhai Y, Dairkee SH, Ljung B-M, Gray JW and Albertson DG: High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet

(1998) **20**, 207-211.
61. Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D and Brown PO: Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet (1999) **23**, 41-46.
62. Kelly DL and Rizzino A: DNA microarray analyses of genes regulated during the differentiation of embryonic stem cells. Mol Reprod Dev (2000) **56**, 113-123.
63. Velculescu VE, Zhang L, Vogelstein B and Kinzler KW: Serial analysis of gene expression. Science (1995) **270**, 484-487.