Int J Hum Genet, 7(1): 49-56 (2007)
PRINT: ISSN 0972-3757 ONLINE: 2456-6360
DOI: 10.31901/24566330.2007/07.01.04

DNA Microarray and Breast Cancer-A Review

Ramanathan S. Uma and T. Rajkumar

Department of Molecular Oncology, Cancer Institute (W.I.A), Chennai 600 020, Tamil Nadu, India

KEYWORDS Gene expression; breast cancer; DNA microarray

ABSTRACT Microarray is a powerful tool used widely to characterize tumors and has greatly improved the ability to subclassify tumors according to shared molecular characteristics and clinical behavior. It is a method to measure the expression of a large number of genes in any specimen simultaneously. Researchers at Stanford University were the first to describe and use DNA microarray to study gene expression in various diseases including cancer. Breast cancer is the second most common cancer among Indian women. Multiple factors like age, diet, obesity, parity, age at first childbirth, oral contraceptives, exogenous estrogens, genetics, environment, geographic location influence the development of this heterogeneous disease. Gene expression in these cancers by microarray is fast gaining in popularity in providing better prognostic and predictive information on the disease. This review is an attempt to look at the recent advances in breast cancer research with DNA microarray technology.

INTRODUCTION

Breast cancer management is a major concern to the clinician due to its extensive heterogeneity. The markers currently associated with the clinical implications include estrogen receptor (ER), progesterone receptor (PR), oncogene her2/neu; others such as cell proliferation marker Ki-67 antigen, proliferating cell nuclear antigen (PCNA), vascular endothelial growth factor (VEGF), CD31, Factor VIII etc. are yet to be proven clinical value (Desai et al. 2002). It would help patients if prognostic markers are available which can help identify patients who are likely to fail surgery. Similarly, predictive markers which can help identify a patient who is unlikely to respond to a particular drug/drugs, will be useful in avoiding failures due to lack of efficacy of the drug and avoiding unnecessary morbidity. Molecular biological methods like cytogenetics, comparative genomic hybridization, whole genome allelotyping, single nucleotide polymorphism studies on candidate genes, differential display, immunohistochemistry among others have shown great promise in addressing the issue at genetic levels. Studying genes at level of transcription (RNA) will give an idea of those that are expressed and analysis of multiple genes simultaneously rather than one at a time will be time saving, less laborious and more predictive. Two methods commonly used now to look at a large number of genes at the mRNA levels are microarray and multiplex quantitative real-time PCR (Rouzier et al. 2005). Apart from gene expression, microarray also finds application in disease diagnosis, drug discovery and toxicological research.

Corresponding Author: Dr. T. Rajkumar, *Director & Scientific Director, Head*, Department of Molecular Oncology, Cancer Institute (W.I.A), Adyar, Chennai 600 020, Tamil Nadu, India
Telephone: + 91-44-22350131; +91-44-22350241
Fax: +91-44-24912085
E-mail: cancer_institute_wia@vsnl.com

WHAT IS A MICROARRAY ?

A microarray is an orderly arrangement of known genes attached to a solid support. Each gene on the solid support referred to as spot or probe is usually less than 200μM in diameter. Each spot has a unique sequence different from the others in the array and will hybridize only to its complementary strand. The first arrays with cDNA clones were spotted onto nylon membrane (Desai et al. 2002) and probed with radiolabeled RNA. These should ideally be able to measure gene expression but slide based arrays first adapted by printing cDNA clones onto glass slides that are specially treated with an adherent polylysine or aminosilane have proven to be smaller, convenient and facilitate higher throughput (Jeffrey et al. 2002). There are two strategies for formatting of the slide/chip: - *in-situ* oligo synthesis or DNA may be directly applied to the surface with nibs or inkjet (Cooper 2001). The entire array can contain anything between 200-40,000 such spots/genes. The slides or chips, as they are referred to, can be high density chips or low density chips depending on the number of genes spotted on them.

Affymetrix Inc.'s GeneChip technology (for

DNA and RNA) has become the industry standard in microarray-based research. Due to the high density of content per array that can be achieved today (an industry leading 6.5 million features) the GeneChips can be used for high-throughput mutation detection, single nucleotide polymorphism genotyping, expression profiling and detection of chromosomal aberrations. This in turn opens the way for clinical applications in genetics, cytogenetics, pharmacogenetics, oncology and pathogen recognition (http://www.affymetrix.com). However theirs is a single dye detection system, hence only a single sample can be processed on a chip unlike many others like Amersham Biosciences who feature a dual dye detection system. The GeneChip system is a closed system and the cost ranges around 0.3 million US dollars in India.

TYPES OF MICROARRAY PLATFORM

1. Classical DNA microarray also called expression chips can itself be of two types - (a) cDNA microarray where cDNAs 500-5000 bases long are immobilized to a solid surface or (b) oligo where 20-80mer oligos are spotted on the chip.
2. Microarray CGH can be used to look for genomic gains and losses or for a change in the number of copies of a particular gene involved in a disease state (Oostlander et al. 2004).
3. Researchers can use SNP microarray technology (Huang et al. 2001) to test an individual for a disease expression pattern to determine whether he or she is susceptible to (at risk of developing) that disease.
4. A protein microarray (Stoll et al. 2005) consisting of antibodies, proteins, protein fragments, peptides or carbohydrate elements are used to screen and assess patterns of interaction with samples containing distinct proteins or classes of proteins.
5. In CpG arrays (Yan et al. 2000) GC-rich sequences derived from human CpG island genomic library are arrayed onto solid supports. Amplicons derived from pools of methylated CpG DNA are hybridized to the arrays.
6. ChIP-on-chip (Oberley et al. 2004) pairs chromatin immunoprecipitation (ChIP) with glass slide microarrays (chip) to analyze how regulatory proteins interact with the genome of living cells. It provides insight into key mechanisms of methylation, histone modification, as well as DNA replication, modification, and repair.

PRINCIPLE AND METHODOLOGY– DNAMICROARRAY

The principle behind microarray is affinity and specificity of complementary base pairing and involves hybridization of cDNA molecule to the DNA template from which it has originated. In microarray experiments a reference sample (normal tissue) is studied along with the experimental sample (cancerous sample). Both total and mRNA can be used; if RNA quantity is insufficient, in vitro transcription based linear amplification may be performed, which will generate enough amplified antisense RNA (Zhao et al. 2002)

25 μg of total RNA or 5 μg of mRNA from experimental and reference samples are reverse transcribed, labeled with different fluorescent dyes (Cy dyes – Amersham Biosciences) and mixed. The mixture is then hybridized overnight to the microarray and scanned at two different wavelengths. The intensities of the fluorescence are measured for each spot on the microarray and the ratio of fluorescent intensity reveals the abundance of RNA expressed by the experimental sample in relation to the reference. Two different microarrays for the reference and experimental sample can be used, in which case the labeled samples need not be mixed. Biotin labeling of samples followed by detection with silver is also gaining popularity (Eppendorf). Many bioinformatics approaches are available to analyze the data generated.

Solid tumors such as breast cancer are highly heterogeneous and hence it is difficult to compare expression pattern between samples. Laser capture microdissection (LCM) is an effective tool to isolate defined population of cells from tissue sections (Sgroi et al. 1999; Luzzi et al. 2001). Assersohn et al. (2002) studied the feasibility of using fine needle aspiration from primary breast cancers for cDNA microarray. But gene expression profiles (without RNA amplification techniques) from only 15% of the patients could be obtained due to lack of enough mRNA in the others. Gene expression profiling of RNA's from formalin fixed paraffin embedded samples is also possible by use of cDNA mediated annealing,

selection, extension and ligation (DASL) assay and universal microarrays (Bibikova et al. 2004).

CHALLENGES IN BREAST CANCER

Hereditary factors, diet, age, race, environmental factors, early menarche, late pregnancies, late menopause are factors associated with this disease. Estrogen plays a major role in the development, growth and differentiation of the normal breast as well as in the development of breast cancer. Tamoxifen and Raloxifene - antagonists of this hormone have been successfully used in its treatment. However only about 50% of breast cancers are ER positive at presentation (Desai et al. 2002). Down-regulation or loss of ER expression may occur via methylation of ER gene promoter (Ottaviano et al. 1994) or these ER negative tumors arise from a population of ER negative precursor cells (Russo et al. 1999). In addition to chemotherapy, surgery and radiation are the other therapeutic strategies, but although combinatorial therapies have decreased the risk of relapse by 20-40%, significant side-effects and development of chemo-resistance may occur (Desai et al. 2002). The first successful gene-based targeting was the development of a therapeutic approach against her2/neu/erbB2, a tyrosine kinase receptor, which is amplified or over-expressed in 30% of breast cancers. The expression of erbB2 is inversely correlated to patient survival and hence a humanized monoclonal antibody called herceptin was developed against it (Slamon et al. 1987). A major challenge is to identify new potential therapeutic targets and prognostic markers by understanding the multiple molecular events and mechanisms by which this cancer develops.

GENE EXPRESSION PROFILING IN BREAST CANCER

Various studies with microarray have been carried out to identify a breast cancer profile on the basis of gene expression. Perou et al. (1999) used cDNA microarray and clustering algorithm to identify patterns of gene expression in human mammary epithelial cells growing in culture and in human primary breast tumors. Two clusters of genes with coherent expression pattern correlated with variation in cell proliferation rates and with activation of the IFN-regulated signal transduction pathway. Membrane based arrays spotted with tags representing 124 genes were used by Martin et al. (2000) to group tumor samples into categories based on expression patterns and cluster analysis and conclude that expression levels of gene clusters have the potential to improve prognostic accuracy and therapeutic outcomes.

A 176-candidate genes microarray was used to study 34 cases of breast cancer (Bertucci et al. 2000). They identified two molecularly distinct subgroups of tumors within the histoclinical category of poor prognosis breast cancer and who are treated with adjuvant chemotherapy. These two subgroups would actually need distinct chemotherapy regimens. They further extended their results (Bertucci et al. 2002) by studying 1000 candidate genes in 55 clinical poor prognosis breast cancer tissue samples treated with adjuvant therapy and in 11 cell lines. They identified and confirmed a predictor set of 23 genes for tumor classification into three classes – stromal, immune and proliferation with significantly different long-term survival. 5-year overall survival and metastasis-free survival rates were 100% and 75% in class I, 65% and 56% in class II and 40% and 20% in class III. The group (Bertucci et al. 2004) also compared inflammatory breast carcinoma (IBC) with non-inflammatory breast cancer (NIBC) using DNA microarrays containing 8000 genes. Supervised analysis identified a set of 109 genes (related to aggressiveness, signal transduction, cell motility, adhesion and angiogenesis) the expression of which discriminated IBC from NIBC. Another 85-gene set was identified that divided IBC patients with significantly different pathological complete response to chemotherapy.

In breast cancer, p53 mutations are associated with worse overall and disease-free survival and even resistance to anticancer therapies. Oligonucleotide arrays covering >30,000 genes studied (Miller et al. 2005) in 251 primary invasive breast carcinomas and a 32-gene p53 signature derived which was associated with patient survival independent of other risk factors. In most cases, tumors with mutated p53 and wild type p53 could easily be distinguished by this expression profile.

PREDICTION OF CLINICAL DISEASE COURSE IN BREAST CANCER

Gene profiles obtained after microarray can

also help predict the disease in its clinical course. Gruvberger et al. (2001, 2004) classified breast tumors and generated a list of genes which discriminate tumors according to ER status. They also studied the prediction of prognostic markers such as S phase of cell cycle (SPF), DNA ploidy status and PR status. Zajchowski et al. (2001) found that the gene expression pattern of a cell line predicted accurately the aggressiveness of the cells by study of differences in gene expression between weakly invasive and highly invasive breast cancer was carried out with a 588 gene array.

A 'poor prognosis' signature consisting of genes regulating cell cycle, invasion, metastasis and angiogenesis was identified by van de Vijver et al.(2002) and van 't Veer et al. (2002) using 25,000 genes DNA microarray analysis on primary breast tumors. They applied supervised classification to identify a gene expression signature (set of 70 genes) strongly predictive (accuracy of 89% in 20 tumors) of a short interval to distant metastases in patients with lymph nodes negative and positive at diagnosis. Soulez et al. (2001) identified novel estrogen-responsive genes in a study with human ZR75-1 breast cancer cells. In continuation, Inoue et al. (2002) analyzed expression profiles of 9000 genes in human MCF-7 breast cancer cells in response to estrogen. Based on the results, they selected estrogen-responsive genes to develop a custom-made microarray system and analyzed effects of estrogen antagonists and estrogen-responsive genes expression profiles in other ER-positive cell lines derived from breast, ovary and stomach. In an attempt to distinguish between invasive lobular (ILC) and invasive ductal (IDC) breast carcinomas, Korkola et al. (2003) used cDNA microarrays with 10,368 genes. They carried out prediction analysis (PAM), significant analysis for microarrays (SAM) and MaxT permutation analysis and identified 11 genes (E-cadherin, survivin, cathepsin B, TPI1, SPRY1, SCYA14, TFAP2B, thrombospondin 4, osteopontin, HLA-G and CHC1) to be differentially expressed. Zhao et al. (2004) carried out a similar study using cDNA microarrays with 42,000 clones. Many of the differentially expressed genes in this study were involved in cell adhesion, motility, fatty acid transport and metabolism, immune responses and electron transport. Over 50% of ILC's differed from IDC's in transcription programs.

Weigelt et al. (2003) studied matching primary breast and their distant metastatic tumors by using a 18,336 human cDNA microarray. Their data showed that similar gene expression profiles of primary breast tumors were maintained in their distant metastasis even when metastasis developed after several years (15 years). Therefore, therapy recommendations based on expression profiles of the primary tumor are a rational approach towards preventing outgrowth of micrometastasis. But there have been many studies contrary to this. Kang et al. (2003) identified genes that promote breast carcinoma metastasis to bone. They compared parental MDA-MB-231 human breast carcinoma cell line with a variant selected for bone colonization. They identified 43 over-expressed genes and 59 under-expressed and referred it as 'bone metastasis signature'. Another study by Huang et al. (2003) also identified aggregate patterns of gene expression (metagenes) that associate with lymph node status and recurrence, and that are capable of predicting outcomes in individual patients with about 90% accuracy. The metagenes defined distinct groups of genes, suggesting different biological processes underlying these two characteristics of breast cancer. Wang et al. (2005) analyzed, with Affymetrix Human U133a GeneChips, the expression of 22000 transcripts from total RNA of frozen tumor samples from 286 lymph-node-negative patients who had not received adjuvant systemic treatment. In the first set of 115 tumors, they identified a 76-gene signature consisting of 60 genes for ER positive patients and 16 genes for ER negative patients. This signature showed 93% sensitivity and 48% specificity in a subsequent independent testing set of 171 lymph-node-negative patients. The gene profile was highly informative in identifying patients who developed distant metastases within 5 years.

CLASSIFICATION OF BREAST CANCER

Perou et al. (2000) and Lonning et al. (2001) studied cDNA microarrays representing 8,102 human genes on matched pairs of breast tumors before and after doxorubicin treatment. The tumors were segregated into four major groups based upon their expression patterns:-

(1) An ERBB2 cluster group expressing high levels of ERBB2 gene and other genes in the same amplicon. ER expression was low or absent.
(2) 'Normal-like' breast cluster expressing many

genes characteristic of normal breast epithelium. ER expression was low or absent.

(3) 'Basal-cell-like' cluster expressing genes characteristic of basal breast cells, in particular with respect to keratin 5 and 17. ER expression was low or absent.

(4) 'Luminal-cell-cluster' characteristic by relatively high expression of many genes including ER, LIV-1 protein, GATA-binding protein 3, prolactin receptor and carnitine palmitoyl-transferase II.

Sorlie et al. (2001) continued this study on larger number of samples and found that this classification based on gene expression patterns can be used as a prognostic marker with respect to overall and relapse-free survival in a subset of patients who received uniform therapy. Another finding was separation of ER positive tumors into at least two distinctive groups with characteristic expression profiles and different prognosis. In order to duplicate these results in a large, prospective trial, Sotirious et al. (2003) analyzed breast cancer patients with known clinical outcome. Their results were in concordance with Sorlie et al. (2001) and van de Vijver et al. (2002) despite differences in patient population, treatment modules and technology employed. ER status of the tumor was the most important discriminating factor for expression subtypes followed by tumor grade. Hierarchical clustering grouped the tumors into two clusters based on their basal (mostly ER negative) and luminal (mostly ER positive) characteristics. Smaller subgroups were identified within each cluster. A group of 485 genes were also identified that were associated with survival. Thus molecular signatures could be generalized to populations and across many microarray technologies.

MICROARRAY IN HEREDITARY BREAST CANCER

Many cases of hereditary breast cancer are due to mutations in either the BRCA1 or BRCA2 gene. Hedenfalk et al. (2001) studied these cancers using 6512 cDNA microarray and found that tumors with germ-line mutations in BRCA1 and those with mutations in BRCA2 differ significantly in their global patterns of gene expression. DNA repair and apoptosis genes (e.g. MSH2) were upregulated in BRCA1 mutation positive samples. In addition, these samples displayed increased expression of genes like PDCD5 associated with apoptosis induction and decreased expression of apoptosis suppressing genes like CTGF. The finding suggested that mutation of BRCA1 leads to constitutive stress-type state.

MICROARRAY IN CHEMOTHERAPY AND DRUG RESISTANCE MECHANISMS

Chang et al. (2003) took core biopsy samples from 24 primary breast tumors before treatment and then assessed tumor response to neoadjuvant docetaxel by cDNA analysis of RNA extracted from biopsy samples using HgU95-Av2 GeneChip. Sensitive tumors had higher expression of genes involved in cell cycle, cytoskeleton, adhesion, protein transport, protein modification, transcription, and stress or apoptosis; whereas resistant tumors showed increased expression of some transcriptional and signal transduction genes. These molecular profiles could allow development of a clinical test for patients with docetaxel sensitivity. Pawitan et al. (2005) obtained gene expression profiles of 159 breast cancer patients and used hierarchical clustering to identify the signature associated with prognosis and impact of adjuvant therapies. A subset of 64 genes from the Affymetrix high density oligonucleotide array human HG-U133 set were found to identify patients with good and poor outcomes. 3 subgroups were identified: patients who did well with therapy, patients who did well without therapy and patients who failed to benefit from therapy. Three genes out of the 64 were among the 70 gene set identified in the study by van't Veer et al. (2002).

Kudoh et al. (2000) used microarray technology on membrane arrays with 5760 genes spotted to study drug-resistance mechanisms (doxorubicin) in breast cancer cell lines. Genes involved in cell cycle mechanisms, transcription factors, signal transduction pathways and various metabolic pathways were found to differentially regulated.

LIMITATIONS OF MICROARRAY TECHNOLOGY

Precise sampling and the ability to sample homogeneous tissue samples are crucial steps in obtaining an accurate analysis. At the same time, the amount of tissue required can also be a problem as a relatively large amount of RNA is

required. However, more recent techniques to amplify RNA have been developed to allow extraction from minute tissue samples. The data obtained from microarray experiments has to be validated by real time PCR or northern blots before they can be proposed as predictive or prognostic markers. This technology does not provide information about post-transcriptional changes or post-translational protein modifications (Brown et al. 2003). In the present state, the technology is expensive and hence may not be affordable to all laboratories.

FUTURE DIRECTIONS

The power of this technology lies in the fact that it allows global analysis of patterns of simultaneous gene expression. Expression patterns can be combined with clinical data to identify new markers for prediction of drug sensitivity and to unravel the molecular basis of drug action. Genomic information has the potential to improve the prognosis in the clinical context by improving patient management based on personalized predictions. The combination of expressed and repressed genes within a tumor reflects its state, revealing proliferative status, metabolic rates and molecular interactions between malignant cells and surrounding fibroblasts, adipocytes, endothelial cells, macrophages and lymphocytes. Therefore, identification of specific patterns of gene expression will undoubtedly improve tumor classification, prognosis and treatment schemes. With respect to breast cancer, the aim should be to identify women 'at risk' at the time of screening, occult micrometastasis at the time of diagnosis and also to prevent unnecessary toxic and expensive adjuvant therapies for patients with non-metastasizing tumors. In an ideal situation especially in the study of cancer, a small population of cells in a tissue from each step of tumor progression should be isolated and microarray analysis should be performed to understand the exact molecular biology of the disease.

Further, developments in nanotechnology, have enabled the production of a 'single molecule array' (http://www.solexa.com/Technology/overview.htm) which should allow the complete sequencing of an entire individual in one reaction and on one array chip.

REFERENCES

Assersohn L, Gangi L, Zhao Y, Dowsett M, Simon R, Powles TJ, Liu ET 2002. The feasibility of using fine needle aspiration from primary breast cancers for cDNA microarray analyses. *Clin Cancer Res,* **8:** 794-801.

Bertucci F, Houlgatte R, Benziane A, Granjeaud S, Adelaide J, Tagett R, Loriod B, Jacquemier J, Viens P, Jordan B, Birnbaum D, Nguyen C 2000. Gene expression profiling of primary breast carcinomas using arrays of candidate genes. *Human Mol Genet,* **9:** 2981-91.

Bertucci F, Nasser V, Granjeaud S, Eisinger F, Adelaide J, Tagett R, Loriod B, Giaconia A, Benziane A, Devilard E, Jacquemier J, Viens P, Nguyen C, Birnbaum D, Houlgatte R 2002. Gene expression profiles of poor-prognosis primary breast cancer correlate with survival. *Human Mol Genet,* **11:** 863-72.

Bertucci F, Finetti P, Rougemont J, Charafe-Jauffret E, Nasser V, Loriod B, Camerlo J, Tagett R, Tarpin C, Houvenaeghel G, Nguyen C, Maraninchi D, Jacquemier J, Houlgatte R, Birnbaum D, Viens P 2004. Gene expression profiling for molecular characterization of inflammatory breast cancer and prediction of response to chemotherapy. *Cancer Res,* **64:** 8558-65.

Bibikova M, Talantov D, Chudin E, Yeakley JM, Chen J, Doucet D, Wickham E, Atkins D, Barker D, Chee M, Wang Y, Fan JB 2004. Quantitative gene expression profiling in formalin-fixed, paraffin-embedded tissues using universal bead arrays. *Am J Pathol,* **165:** 1799-1807.

Brown I, Heys SD, Schofield AC 2003. From peas to "chips" - the new millennium of molecular biology: a primer for the surgeon. *World J Surg Oncol,* **1:** 21.

Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, Mohsin S, Osborne CK, Chamness GC, Allred DC, O'Connell P 2003. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet,* **362:**362-9.

Cooper CS 2001. Applications of microarray technology in breast cancer research. *Breast Cancer Res,* **3:**158-175.

Desai KV, Kavanaugh CJ, Calvo A, Green JE 2002. Chipping away at breast cancer: insights from microarray studies of human and mouse mammary cancer *Endocr Relat Cancer,* **9:** 207-220.

Gruvberger S, Ringner M, Chen Y, Panavally S, Saal LH, Borg A, Ferno M, Peterson C, Meltzer PS 2001. Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns *Cancer Res,* **61:** 5979-84

Gruvberger-Saal SK, Eden P, Ringner M, Baldetrop B, Chebil G, Borg A, Ferno M, Peterson C, Meltzer PS 2004. Predicting continuous values of prognostic markers in breast cancer from microarray gene expression profiles. *Mol Cancer Ther,* **3:** 161-68.

Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, Meltzer P, Gusterson B, Esteller M, Kallioniemi OP, Wilford B, Borg A, Trent J 2001.

Gene expression profile in hereditary breast cancer. *N Engl J Med*, **344:** 539-48.

Huang JX, Mehrens D, Wiese R, Lee S, Tam SW, *et al*. 2001. High-throughput genomic and proteomic analysis using microarray technology. *Clinical Chem*, **47:** 1912-16.

Huang E, Cheng SH, Dressman H, Pittman J, Tsou MH, Horng CF, Bild A, Iversen ES, Liao M, Chen CM, West M, Nevins JR, Huang AT 2003. Gene expression predictors of breast cancer outcomes *Lancet*, **361:** 1590-96.

Inoue A, Yoshida N, Omoto Y, Oguchi S, Yamori T, Kiyama R, Hayashi S Development of cDNA microarray for expression profiling of estrogen-responsive genes 2002. *J Mol Endocrinol*, **29:** 175-192.

Jeffrey SS, Fero MJ, Borresen-Dale AL, Botstein D 2002. Expression array technology in the diagnosis and treatment of breast cancer. *Mol Interv*, **2:** 101-09.

Kang Y, Siegel PM, Shu W, Drobnjak M, Kakonen SM, Cordon-Cardo C, Guise TA, Massague J 2003. A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell*, **3:** 537-549.

Korkola JE, DeVries S, Fridlyand J, Hwang ES, Estep ALH, Chen YY, Chew KL, Dairkee SH, Jensen RM, Waldman FM 2003. Differentiation of lobular versus ductal breast carcinomas by expression microarray analysis. *Cancer Res*, **63:** 7167-75.

Kudoh K, Ramanna M, Ravatn R, Elkahloun AG, Bittner ML, Meltzer PS, Trent JM, Dalton WS, Chin KV 2000. Monitoring the expression profiles of doxorubicin-induced and doxorubicin-resistant cancer cells by cDNA microarray. *Cancer Res*, **60:** 4161-66.

Lonning PE, Sorlie T, Perou CM, Brown PO, Botstein D, Borresen-Dale AL 2001. Microarrays in primary breast cancer – lessons from chemotherapy studies. *Endocr Relat Cancer*, **8:** 259-263.

Luzzi V, Holschlag V, Watson MA 2001. Expression profiling of ductal carcinoma *in situ* by laser capture microdissection and high-density oligonucleotide arrays. *Am J Pathol*, **158:** 2005-10.

Martin KJ, Kritzman BM, Price LM, Koh B, Kwan CP, Zhang X, Mackay A, O'Hare MJ, Kaelin CM, Mutter GL, Pardee AB, Sager R 2000. Linking gene expression patterns to therapeutic groups in breast cancer. *Cancer Res*, **60:** 2232-38.

Miller LD, Smeds J, George J, Vega VB, Vergara L, Ploner A, Pawitan Y, Hall P, Klaar S, Liu ET, Bergh J 2005. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci U S A*, **102:** 13550-555.

Oberley MJ, Tsao J, Yau P, Farnham PJ 2004. High-throughput screening of chromatin immunoprecipitates using CpG-island microarrays. *Methods Enzymol*, **376:** 315-334.

Oostlander AE, Meijer GA, Ylstra B 2004. Microarray-based comparative genomic hybridization and its applications in human genetics. *Clin Genet*, **66:** 488-495.

Ottaviano YL, Issa JP, Parl FF, Smith HS, Baylin SB, Davidson NE 1994. Methylation of the estrogen receptor gene CpG island marks loss of estrogen receptor expression in human breast cancer cells. *Cancer Res*, **54:** 2552-55.

Pawitan Y, Bjohle J, Amler L, Borg AL, Egyhazi S, Hall P, Han X, Holmberg L, Huang F, Klaar S, Liu ET, Miller L, Ploner HNA, Sandelin K, Shaw PM, Smeds J, Skoog L, Wedren S, Bergh J 2005. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts *Breast Cancer Res*, **7:** R953 – R964.

Perou CM, Jeffrey SS, van de Rijn M, Rees CA, Eisen MB, Ross DT, Pergamenschikov A, Williams CF, Zhu SX, Lee JCF, Lashkari D, Shalon D, Brown PO, Botstein D 1999. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc Natl Acad Sci U S A*, **96**: 9212-17.

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, Botstein D 2000. Molecular portraits of human breast tumors *Nature*, **406:** 747-52.

Rouzier R, Wagner P, Morandi P, Pusztai L 2005. Gene expression profiling of primary breast cancer. *Curr Oncol Rep*, **7:** 38-44.

Russo J, Ao X, Grill C, Russo JH 1999. Pattern of distribution of cells positive for estrogen receptor alpha and progesterone receptor in relation to proliferating cells in the mammary gland. *Breast Cancer Res Treat*, **53:** 217-227.

Schena M, Shalon D, Davis RW, Brown PO 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, **270:** 467-70.

Sgroi DC, Teng S, Robinson G, LeVangie R, Hudson JR, Elkahloun AG 1999. *In Vivo* gene expression profile analysis of human breast cancer progression. *Cancer Res*, **59:** 5656-61.

Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL 1987. Human breast cancer: correlation of relapse and survival with amplification of the Her-2/neu oncogene. *Science*, **235:** 177-182.

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Hilde J, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lonning PE, Borresen–Dale AL 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A*, **98:** 10869-74.

Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, Martiat P, Fox SB, Harris AL, Liu ET 2003. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A*, **100:**10393-98.

Soulez M, Parker MG 2001. Identification of novel oestrogen receptor target genes in human ZR75-1 breast cancer cells by expression profiling. *J Mol Endocrinol*, **27:** 259-74.

Stoll D, Templin MF, Bachmann J, Joos TO 2005. Protein microarrays: applications and future challenges. *Curr Opin Drug Discov Devel*, **8:** 239-252.

van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AAM, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A,

Glas A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH, Bernards R 2002. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med*, **347:** 1999-2009.

van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH 2002. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*, **415:** 530-36.

Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J, Jatkoe T, Berns EM, Atkins D, Foekens JA 2005. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet*, **365:** 671-79.

Weigelt B, Glas AM, Wessels LFA, Witteveen AT, Peterse JL, van't Veer LJ 2003. Gene expression profiles of primary breast tumors maintained in distant metastases. *Proc Natl Acad Sci U S A*, **100:** 15901-905.

Yan PS, Perry MR, Laux DE, Asare AL, Caldwell CW, Huang TH 2000. CpG island arrays: an application toward deciphering epigenetic signatures of breast cancer. *Clinical Cancer Research*; **6:** 1432-38.

Zajchowski DA, Bartholdi MF, Gong Y, Webster L, Liu HL, Munishkin A, Beauheim C, Harvey S, Ethier SP, Johnson PH 2001. Identification of gene expression profiles that predict the aggressive behavior of breast cancer cells. *Cancer Res*, **61:** 5168-78.

Zhao H, Hastie T, Whitfield ML, Borresen-Dale AL, Jeffrey SS 2002. Optimization and evaluation of T7 -based RNA linear amplification protocols for cDNA microarray analysis. *BMC Genomics*, **3:** 31.

Zhao H, Langerod A, Ji Y, Nowels KW, Nesland JM, Tibshirani R, Bukholm IK, Karesen R, Botstein D, Borresen-Dale AL, Jeffrey SS 2004. Different gene expression patterns in invasive lobular and ductal carcinomas of the breast. *Mol Biol Cell*, **15:** 2523-36.