CANCER GENOMICS & PROTEOMICS *4*: 135-146 (2007)

Review

Application of Array-based Genomic and Epigenomic Technologies to Unraveling the Heterogeneous Nature of Breast Tumors: On the Road to Individualized Treatment

MARK ABRAMOVITZ[1] and BRIAN LEYLAND-JONES[2]

[1]*VM Institute of Research, Montreal, Quebec, Canada;*
[2]*Winship Cancer Institute, Emory University, Atlanta, GA, U.S.A.*

Abstract. *The recent application of genomic microarray technology to the molecular profiling of breast tumors has clearly demonstrated their heterogeneous nature. Targeted treatment strategies are having a clear impact on patient survival. It has also become apparent that accumulated mutations, genomic instability, epigenetic phenomena, genetic variability and environmental factors all contribute to the uniqueness of a patient's tumor. Novel genomic and epigenetic-based technologies have been or are being developed in order to greatly enhance the analysis of tumor samples including those samples previously thought unusable due to the fixation process, such as archival formalin-fixed paraffin-embedded (FFPE) samples. Patients and their tumors can now be studied with regard to genetic variation, genomic instability, gene expression, gene mutations, and methylation patterns. These areas of research are being made more accessible through genome-wide screening technologies and will, in the near future, rapidly expand our understanding of what contributes to the unique properties of each tumor and lead to the identification of genes that could be potential therapeutic targets for specific tumor subtypes. Application of these technologies to our understanding of breast cancer will undoubtedly have an impact on the individualization of treatment for breast cancer patients in the not to distant future.*

Breast cancer is the leading cause of cancer death among women and it is estimated that in the United States approximately 40,000 women are expected to die from this disease in 2007 (29). Breast cancer will account for ~26% of new cancer cases among women in the United States in 2007 (29) and there is a great need, therefore, to improve early detection, develop targeted therapies and implement effective individualized treatment strategies given the heterogeneous nature of the disease. The current classification system for breast cancer is inadequate but by being able to combine it with a robust molecular characterization of the tumor has the potential to ensure that optimized individual treatment regimens for maximum benefit to the patient will become a reality.

Molecular medicine is exploiting pattern-based diagnostic discoveries in genomics and proteomics, with the ultimate aim of discovering new types of biomarkers/biomarker sets for targeted therapies with improved sensitivity and specificity (3). In breast cancer, current therapies specifically include the aromatase inhibitors/SERM's, trastuzumab (Herceptin) and other targeted HER2 (epidermal growth factor receptor 2) therapies, bevacizumab (Avastin) and other anti-angiogenic agents, erlotinib (Tarceva)/cetuximab (Erbitux) and other EGFR (epidermal growth factor receptor 1) targeted therapies, as well as agents directed towards two or more of these targets. But in order to maximize the effectiveness of targeted therapy and to develop novel targeted drugs tailored to the individual will require a much more detailed understanding of both the tumor as well as the patient's genetic makeup (13).

Great strides are rapidly being made in both genomic and epigenomic array-based platforms. Array-CGH with ultra-high density arrays that contain 2 million features will be available this year with unprecedented resolution across the entire genome. Single nucleotide polymorphism (SNP) analysis can now be performed on arrays that assay one million SNP loci. A molecular profiling platform is now available for analysis of over 500 genes using RNA prepared from previously unusable archival formalin-fixed tumor tissue specimens. Epigenetic-based methylation arrays have

Correspondence to: Brian Leyland-Jones, Winship Cancer Institute, Emory University, Atlanta, GA, U.S.A. e-mail: leyland@emory.edu

Key Words: Breast cancer, targeted therapies, DNA copy number alterations, microarrays, DASL assay, array-CGH, SNPs, epigenetics, methylation, review.

just become available to probe over 1,500 CpG loci selected from over 800 genes combined with high sample throughput. The ability to apply these novel genomic and epigenetic array-based technologies to rapidly characterize both the patient's genetic variability and their breast tumor with respect to prognosis and treatment outcome will be of immense benefit to the individual patient. An added benefit is that this type of detailed analysis could also potentially highlight genes and/or signaling pathways implicated in tumorigenesis.

Novel Genomic Technologies

Array-CGH. Genomic instability which can give rise to gene amplification and gene deletion, also known as DNA copy number alterations (CNAs), play important roles in tumor progression and it has become apparent that an increased number of CNAs correlate with a poor prognosis for the patient. Several genes that undergo CNAs have been identified that act as either oncogenes or tumor suppressors which are crucial factors in the disease process. In breast cancer, *ERBB2* (HER2) amplification, which occurs in 15-25% of women with early breast cancer and is associated with an aggressive form of the disease, has been linked to a poor prognosis. Amplification of *ERBB2* is also associated with responsiveness to trastuzumab, a humanized monoclonal antibody that targets the HER2 receptor (62).

Another oncogene, *MYC*, has been found to be amplified in a number of different tumors including breast. In fact, in one study (39), *MYC* amplification was found to occur in 15% of breast cancer patients. The authors also found that *MYC* amplification was associated with *ERBB2* amplification and cellular proliferation. *MYC* amplification also appears to be an early event in breast tumor development as opposed to *ERBB2* amplification which seems to occur at a later point in tumor progression (50). Co-amplified *ERBB2*/*MYC* correlated with hormone receptor negative status and, most strikingly, *ERBB2*/*MYC*-coamplified cancers had a worse prognosis than tumors with only one of these gene amplifications (4). Despite the poor prognosis, Dr. Paik and co-workers, have shown that patients, in whose tumors *ERBB2* and *MYC* are co-amplified, actually responded better to trastuzumab treatment than patients with only ERBB2 amplified (HR for 4-year recurrence of 0.24 *versus* 0.63, respectively), resulting in a 4-year recurrence-free survival of over 90% (34). Therefore, contrary to initial belief, cMYC expression and *MYC* amplification status in *ERBB2*-overexpressing primary breast tumors may be a marker of trastuzumab efficacy and not resistance.

In another interesting study, Dr. Slamon and colleagues found that co-amplification of the topoisomerase II alpha gene (*TOP2A*), which occurred in ~40% of *ERBB2*-positive patients in their study, may confer a therapeutic advantage to anthracycline-based, trastuzumab combination regimens (43). In addition, Slamon's group observed a trend towards benefit of AC-TH over both AC-T and TC (carboplatin)-H (N=744, 13 *vs.* 23 *vs.* 21 events) in the co-amplified group of patients (51). A significant benefit of AC-TH and TCH over AC-T (N=176, 45 *vs.* 54 *vs.* 92 events, $p<0.001$) was also observed in the non co-amplified group of patients. These results indicate that CNAs of *TOP2A* may be a candidate marker for response to trastuzumab.

PTEN (phosphatase and tensin homolog) is a tumor suppressor gene that when deficient, results in resistance of breast cancer cells to trastuzumab therapy (38). This is thought to be a consequence of deregulation of the PI3K (phosphoinositide-3-kinase)-AKT (v-akt murine thymoma viral oncogene homolog) -mTOR (FK506 binding protein 12-rapamycin associated protein) pathway. Translocation of PTEN to the nucleus has also recently been found to help maintain chromosomal stability (49) ascribing an additional tumor suppressor role to PTEN. Loss of heterozygosity (LOH) at the *PTEN* locus has been found to occur in breast cancer (~24%; 31 of 131 cases) and is correlated with HER2 overexpression (57). In prostate cancer, *PTEN* gene deletion, which was found to occur in 68% of cases (24/35), suggests that loss of the tumor suppressor gene, *PTEN*, is an important event in prostate cancer progression (63). The consequences of *PTEN* gene deletion in breast cancer have not been studied, but both the LOH study in breast cancer and the FISH study in prostate cancer suggest that this would be an important avenue to explore.

Thus, genes such as *MYC, TOP2A* and *PTEN* have potential, clinical utility as prognostic and/or predictive markers of trastuzumab response in primary *ERBB2*-amplified breast cancer patients. Other as yet unidentified amplified or deleted genes may play important roles as biomarkers of prognosis or act as putative biomarkers of response to treatment and/or relapse. Some may also play critical roles in the pathobiology of breast cancer. The ability to scan the entire genome for CNAs at high resolution would be invaluable towards discovering important prognostic and or predictive breast cancer genes.

Array-Comparative Genomic Hybridization (array-CGH), introduced back in 1998 (42), was a major advance in array-CGH technology because it enabled the analysis of chromosomal CNAs that occurred as a result of genomic instability without the need to culture tumor cells. The probe sets used on array-CGH are of two basic types, large insert clones such as bacterial artificial chromosomes (BACs) or more recently, oligonucleotides (generally 60 to 70-mers). In general, most published array-CGH studies have been conducted at a resolution of 1 to 2 Mb. Current technology has resulted in the manufacturing of high-density CGH-arrays with greatly increased resolution and in the

case of NimbleGen, Systems Inc. (Madison, WI, USA) their arrays are tiled with probes through genic and intergenic regions at a median probe spacing of 6 kb. A generalized array-CGH scheme is shown in Figure 1. This type of resolution will ensure that genome-wide CNAs will be possible to detect.

With regard to breast cancer, array-CGH has been used principally to identify novel breast cancer-related genes, genes that can be used to classify different types of breast cancer or to discover genes that yield prognostic or predictive information. With the availability of high-density array-CGH, disease-related gene aberrations, including translocation breakpoints, should be more easily identified. Previous CGH studies, either chromosome-based or BAC array-based, have identified chromosomal changes related to clinicopathological parameters such as tumor grade, estrogen receptor (ER) status, tumor subtype or overall survival. Loss on 5q and gain on 2p and 6p, for example, have been found to occur more frequently in high-grade tumors (45). In a BAC array-based CGH study, gain on 1q and loss on 5q occurred more frequently in ER-negative tumors (36). Several amplified loci, including 17q12 (*ERBB2*) and 20q13 (*ZNF217*) (zinc finger protein 217) (6, 12) have also been shown to correlate with overall survival (46, 56). BAC array-CGH was used to assess copy number CNAs in ER-positive breast tumor tissue in order to determine if the observed changes would be different for samples from patients who had distant recurrence within 5 years *versus* those who were without disease at least 5 years after diagnosis. What was found was that although CNAs occurred with similar frequencies, CNAs were significantly different between the recurrence group and non-recurrence group samples (25).

Array-CGH has been used in the classification of breast tumors and shown that subtypes can be discerned based on genomic instability patterns. In one study, array-CGH made up of 2464 genomic clones, was applied to the analysis of CNAs in 62 sporadic invasive ductal carcinomas and three breast tumor subtypes were found based upon genomic DNA CNAs (19). The breast cancer subtypes Luminal A, Luminal B, HER2 and Basal-like, as defined by expression profiling of 78 tumors (52), were also found to yield distinct CNA profiles suggesting that genomic instability mechanisms differ between subtypes (6).

The use of array-CGH will allow for the identification of chromosomal aberrations in breast cancer that have, heretofore, not been detected and should also help in the discovery of novel breast cancer-related genes that will be key to understanding the molecular mechanisms underlying the disease. The technology is at a point where high resolution arrays are now available and this will greatly enhance the utility of this technology. Another important advance is the ability to perform array-CGH on archival FFPE tumor samples. Recent studies have addressed this important issue and shown that DNA can be prepared from FFPE tissue samples and used successfully in array-CGH (31, 32, 35). Taken together, these technical improvements in array-CGH open up the possibility of using this platform in the clinical setting for the diagnosis of breast cancer.

SNP arrays. Breast cancer is a major disease that has a complex genetic component in which genetic susceptibility results from the combined effect of a number of variant genes that contribute to the disease etiology and can impact on response to treatment. Finding these genes and understanding the role that they play is an important goal for improving breast cancer prognosis. SNPs have been linked to individual variations in breast cancer susceptibility. Genetic variations in genes involved in ER and drug metabolism, genes involved in cell-cycle control as well as variations depending on ethnic background have been found to contribute to breast cancer susceptibility. The best evidence for linking SNPs to breast cancer comes from the Nurses' Health Study which examined a large and primarily healthy population of women in a prospective study with long term follow-up on the development of breast cancer. During a 10-year observation period the study identified seven polymorphisms that were overrepresented and could act as small but significant risk factors for non-hereditary breast cancer. Of the seven SNPs identified, three are in genes involved in the activation by or metabolism of sex steroids (*PGR*, progesterone receptor; *AR*, androgen receptor; *CYP19A1*, aromatase) (14, 22, 23), two in DNA repair (*XRCC1*, X-ray repair complementing defective repair in Chinese hamster cells; *XRCC2*) (24), one in metabolism of both sex steroids and xenobiotics (*CYP1A1*, aryl hydrocarbon hydroxylase) (28), and one in cell cycle control (*VDR*, vitamin D receptor) (11). In a meta-analysis of seven case-control studies that involved 4,284 individuals, a variant of the *TGFBR1* (transforming growth factor receptor beta I) gene (TGFBR1*6A), a gene involved in cell-cycle control, was determined to be a strong risk factor for breast cancer and well as other cancers such as ovarian, colon and hematological malignancies (33).

These genetic variants add another layer of complexity to the already heterogeneous nature of breast tumors but the question then becomes what is the most efficient way of identifying such variants? Most genotyping efforts have focused on the candidate gene approach, however, newer array-based technologies are allowing for empirical genome-wide studies to be conducted. In order to enable large-scale genotyping studies to be feasible, the whole-genome genotyping assay (WGA), which genotypes array-hybridized genomic targets directly, was developed at Illumina Inc. (San Diego, CA, USA) (21). The Infinium™ WGA is a

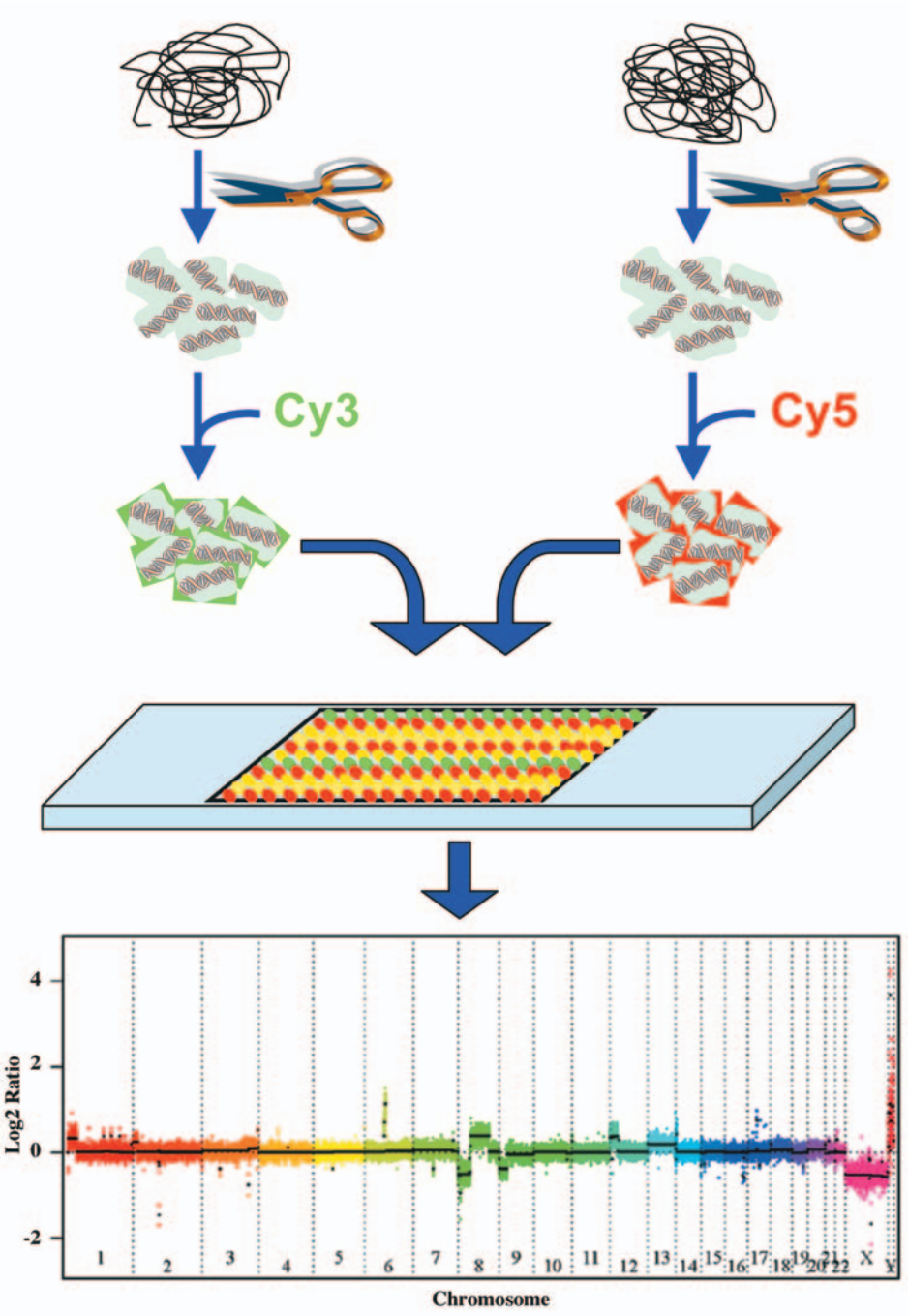

Figure 1. *Array-CGH assay. 1. Test genomic DNA is isolated from a patient's tumor sample and reference genomic DNA is isolated from normal tissue of the same patient. 2. The DNA is randomly fragmented into lower molecular weight species. 3. The test and reference DNA fragments are differentially labeled with fluorescent dyes, Cy3 (green) and Cy5 (red). 4. Both the test and reference labeled DNA fragments are hybridized to the array. In case of a deletion in the test DNA, less test DNA will bind to a particular array spot and the red label of the reference DNA will be more intense; gains in the test genome can be identified by a dominance of the green label of the test DNA. Spots, representing sequences with the same copy number in the test genome relative to the reference genome appear yellow. 5. The data from the array is processed such that genome-wide chromosomal gains and losses can be detected in a tumor sample. The array can be made of BACs or oligonucleotides tiled over the entire genome, however, the use of oligonucleotides, can greatly increase the resolution of the array.*

scalable genome-wide SNP genotyping assay that can interrogate a large number of SNPs, which is currently up to 1 million SNP loci. The assay scheme is described in Figure 2. Using this assay with 550,000 SNPs, a genome-wide association study in the Cancer Genetic Markers of Susceptibility project identified a new SNP association at

8q24 on prostate cancer susceptibility (61). This SNP (rs6983267) is independent of a previously identified SNP (rs1447295) within 8q24 that was found to contribute to prostate cancer in men of European ancestry (5). The application of these types of genome-wide arrays in future breast cancer studies will have the potential to identify novel variants that contribute to breast cancer.

Gene expression profiling of breast tumor tissue. The heterogeneous nature of breast tumors at the molecular level really became apparent with the application of microarray-based gene expression profiling platform to analyze RNA prepared from freshly frozen tumor tissue samples. Initial studies published by Perou and colleagues (40, 41) and Sorlie and co-workers (52) clearly identified breast tumor subtypes using non-supervised hierarchical clustering of tumor samples. They were able to establish that breast tumors could be phenotypically classified into subtypes distinguishable by differences in their expression profiles into at least four subtypes including ER-positive Luminal A; ER-positive Luminal B; HER2-positive, ER-negative; Basal-like (ER negative, PGR-negative, HER2-negative; so-called Triple Negative) and that these molecular subtypes were highly significantly correlated with overall survival (53).

Van't Veer and colleagues (59) applied supervised hierarchical clustering to identify a 70-gene-expression signature strongly predictive of a short interval to distant metastasis in primary, LN-negative, breast tumors of young patients (<55 yrs old). The utility of the 70-gene prognosis profile was subsequently confirmed in a follow-up study in which multivariate Cox regression analysis showed that the prognostic profile was a strong independent factor for predicting disease outcome and was a more powerful predictor of the outcome of disease in young patients with breast cancer than standard systems based on clinical and/or histologic criteria (58).

Piccart and colleagues (10) recently presented the validation of the Amsterdam 70-gene prognostic signature in untreated breast cancer using frozen archival tumor material of LN-negative patients aged <61 years old at the time of diagnosis. Their results provide evidence for the clinical value of this new genomic test but the definitive evidence for the clinical utility of the 70-gene prognosis signature compared with that of traditional prognostic indicators for assigning adjuvant chemotherapy to patients with node-negative breast cancer will be sought in the MINDACT clinical trial (9).

Applying the DASL assay to the molecular profiling of FFPE breast tumor tissue. In order to conduct expression profiling using microarrays, freshly-frozen tissue is required as the starting material for RNA extraction. This presents a problem for exploiting FFPE tumor tissue samples, by far the most abundant source of archival clinically well-annotated specimens available, for molecular profiling studies. The formalin fixation process applied to tissue samples causes RNA fragmentation and chemical modification to occur (2, 18, 20, 44, 47, 48, 54, 55), however, despite this, Companies such as Roche Diagnostics Corp. (Indianapolis, IN, USA),

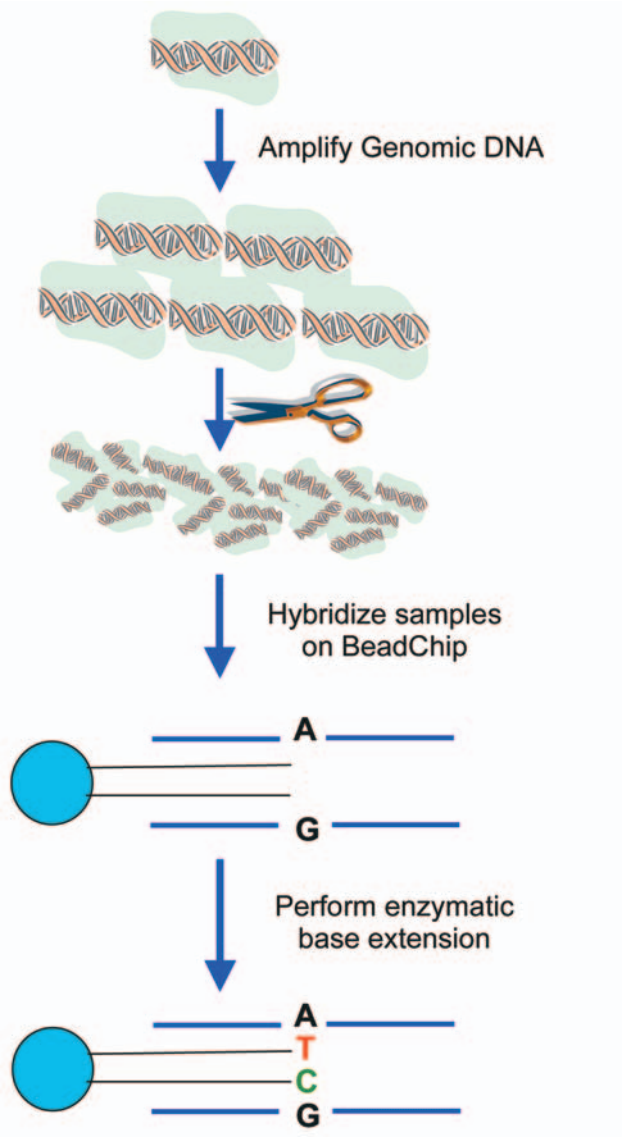

Figure 2. *The Infinium Whole-Genome Genotyping assay (WGA). The WGA can be used to interrogate large numbers of SNPs, currently up to 1 million per WGA. Genomic DNA must first be amplified followed by fragmentation. The DNA is then purified and applied to BeadChips during an overnight incubation. The amplified fragments anneal to 50-mers which are locus-specific and that are covalently linked to 1 million different bead types. Each bead type corresponds to each allele per SNP locus. Enzymatic base extension followed by fluorescent staining results in differential bead intensities which are detected using a BeadArray reader and the data is subsequently analyzed in order to determine the SNP profile of the particular DNA sample.*

Qiagen Inc. (Valencia, CA, USA), Ambion (Austin, TX, USA), Stratagene (La Jolla, CA, USA) and SuperArray (Frederick, MD, USA) have developed kits for the extraction of total RNA from FFPE tissue samples that can be used to perform qRT-PCR- (quantitative Reverse Transcriptase-Polymerase Chain Reaction). The ability to tap into this wealth of available FFPE tissue samples in conjunction with a molecular profiling system would be of great benefit to help identify bona fide prognostic and predictive genes.

The company, Illumina, has been specifically designed the DASL (cDNA-mediated Annealing, Selection, extension and Ligation) assay as a gene expression profiling system to be used with degraded RNAs such as those derived from FFPE tumor samples. The procedure employs random priming for cDNA synthesis and, therefore, probes are designed such that they can target any unique region of the gene without limiting the selection of the optimal probe to the 3' ends of transcripts. In addition, due to the small size of the targeted gene sequence ($\sim$50 nucleotides), RNAs that are otherwise too degraded for conventional microarray analysis can be readily detected.

The DASL assay requires the use of specifically designed complex oligonucleotides. The 5' oligonucleotides consist of two parts: the gene specific sequence and a universal PCR primer sequence. The 3'-oligos consist of three parts: the gene specific sequence, a unique address sequence which is complementary to one of 1536 capture sequences on the array and a universal PCR primer sequence at the 3' end. A single address sequence is uniquely associated with a single target site. This address sequence allows the PCR-amplified products to hybridize to a universal beadarray bearing the complementary address sequences (Figure 3). Oligonucleotides targeting 512 cancer-related genes can be used at a density of three non-overlapping probe sets per gene which results in a 1536-plex measurement for each sample (8, 16).

A 231 gene cancer panel was used in the DASL assay in order to profile both breast and colon cancer FFPE tumor samples. Cluster analysis was able to separate breast from colon tissue types and subsequently, divide each tissue sample set into cancer *versus* normal (8). By designing a breast cancer specific panel for the DASL platform, therefore, has the potential to enable the identification of differential signature gene sets that can be used for diagnosis, prognosis and/or monitoring of breast cancer that will be more effective than current traditional clinical/pathological factors or molecular-based signatures.

Epigenetics and Breast Cancer

It is becoming more evident that not only are genetic changes important in cancer progression and tumor growth but also epigenetic changes which involve subtle molecular modifications that can result in the activation and or suppression of genes leading to abnormalities resulting in tumorigenesis (15). Epigenetic changes have been discovered in many cancers including breast, colon, prostate and blood. Two important molecular mechanisms of epigenetic modification include aberrant methylation of the cytosine residue of DNA that results in either gene suppression or activation and acetylation of histones which results in gene activation. The epigenotype is affected by such factors as environment, age, epigenotype of the parents and the genotype at the loci that regulate DNA methylation and chromatin.

Unlike genetic mutations, epigenetic modifications are reversible and have opened up the possibility of developing drugs that can alter the epigenetic profile of a tumor. To date, two drugs that inhibit methylation, azacitidine (Vidaza) and decitabine (Dacogen), and are both used in the treatment of myelodysplastic syndrome, a blood disorder that can lead to leukemia, have been approved by the FDA in 2004 and 2006, respectively. Vorinostat (Zolinza) is the first oral histone deacetylase (HDAC) inhibitor to be approved by the FDA as an anti-cancer agent for the treatment of skin manifestations of cutaneous T-cell lymphomaTCL) (1).

In breast cancer, low expression of tumor suppressor genes is linked to tumorigenesis and there is evidence that epigenetic factors are involved in this process. In breast cancer cell lines for example, the state of methylation of certain genes, tumor suppressors or oncogenes, shed light on the important role that aberrant methylation may play in breast cancer progression (27). In one study, methylation data was obtained from 86 primary breast cancers and revealed that aggressive tumors tend to exhibit higher levels of promoter hypermethylation (60). Tumor suppressor genes such as *BRCA1* (breast cancer 1, early onset) and SFN (stratifin; 14-3-3sigma) or genes involved in cell adhesion such as *CADM1* (cell adhesion molecule 1; TSLC1) and *EPB41L3* (erythrocyte membrane protein band 4.1-like 3; DAL-1), are inactivated by methylation (26, 30). DNA methylation was also studied with regard to patients whose tumors were resistant to endocrine therapy with recurrent breast cancer. The DNA methylation status was evaluated for 117 candidate genes in a cohort of 200 steroid hormone receptor-positive tumors from patients who received the antiestrogen tamoxifen as first-line treatment for recurrent breast cancer. The study found that 10 genes were significantly associated with clinical outcome of tamoxifen therapy. Of interest, hypermethylation of the promoter region of one gene in particular, *PSAT1* (phosphoserine aminotransferase), was associated with a favorable clinical outcome and therefore, an indicator of response to tamoxifen in ER-positive patients with recurrent breast cancer (37).

Methylation microarrays. The application of methylation microarrays to the study of cancer in general and breast cancer in particular is very recent. In one such study,

published at the end of 2006, changes in DNA methylation were assessed by differential methylation hybridization, a high-throughput promoter CpG island microarray analysis. The authors investigated promoter DNA methylation profiles associated with acquired resistance to fulvestrant *versus* tamoxifen. The results showed that growth-promoting genes were more frequently activated due to promoter hypomethylation compared with gene inactivation by promoter hypermethylation in antiestrogen-resistant cells (17). The ability to scan the entire genome for hyper- and hypo-methylated sites could allow for the identification of therapeutic targets and strategies for resensitization.

Illumina has developed an array-based platform that combines high sample throughput with single-site CpG resolution for DNA methylation profiling (7). The assay involves treating the genomic DNA with bisulfite to convert unmethylated cytosines to uracil leaving methylcytosines unchanged which results in non-complementary top and bottom strands. The allele specific query oligonucleotides employing two different dyes are used to interrogate methylated and non-methylated sites of 1536 different CpG sites simultaneously. The DNA methylation assay scheme is detailed in Figure 4. Illumina has created their standard Cancer Methylation Panel, that spans 1,505 CpG loci selected from 807 genes with 71% of those selected genes containing at least two CpG sites. The genes selected for the Cancer panel fall into various classes, including tumor suppressor genes, oncogenes, genes involved in DNA repair, cell cycle control, differentiation, apoptosis, X-linked and imprinted genes. Custom selected CpG sites will also be feasible using this platform, which will allow any gene to be evaluated in the assay.

Conclusion

A number of factors, including genomic and epigenomic, contribute to the heterogeneous nature of breast cancer and consequently, newly developed targeted anti-cancer drugs will only be effective in a subset of patients and perhaps at only a certain stage of their disease. Therefore, the uniqueness of each individual patient's tumor in the context of their unique genetic background needs to be addressed in personalized medicine. The development over the past few years of genome-wide array technologies will allow scientists involved in translational research to uncover novel genes, associations and mechanisms related to breast cancer progression and development as never before. Genomic instability to varying degrees is found in all breast tumors. Array-CGH now has the resolution to allow researchers to more rapidly identify genes for which genomic CNAs have been found and relate them to different breast cancer phenotypes. Million SNP arrays will help uncover novel variant genes associated with breast cancer. Molecular profiling of FFPE tumor samples, the most abundant source of well-annotated archival breast cancer clinical specimens, will help identify novel prognostic as well as predictive gene sets. Methylation arrays will contribute to our understanding of the role that epigenetics plays in breast cancer. These technologies will certainly help towards unraveling the uniqueness of each tumor and contribute to our understanding of breast cancer and the development of novel targeted therapeutics with the goal of individualizing treatment to maximize benefit to the patient. The layers of complexity, however, are great, and will ultimately require a systems biology approach that takes into account all genomic, epigenomic, proteomic and environmental factors in order to understand how all of the parts fit together in order to unravel the unique properties of each and every breast tumor.

References

1 Vorinostat (Zolinza) for cutaneous T-Cell lymphoma. Med Lett Drugs Ther *49*: 23-24, 2007.

2 Abrahamsen HN, Steiniche T, Nexo E, Hamilton-Dutoit SJ and Sorensen BS: Towards quantitative mRNA analysis in paraffin-embedded tissues using real-time reverse transcriptase-polymerase chain reaction: a methodological study on lymph nodes from melanoma patients. J Mol Diagn *5*: 34-41, 2003.

3 Abramovitz M and Leyland-Jones B: A systems approach to clinical oncology: focus on breast cancer. Proteome Sci *4*: 5, 2006.

4 Al Kuraya K, Schraml P, Torhorst J, Tapia C, Zaharieva B, Novotny H, Spichtin H, Maurer R, Mirlacher M, Kochli O, Zuber M, Dieterich H, Mross F, Wilber K, Simon R and Sauter G: Prognostic relevance of gene amplifications and coamplifications in breast cancer. Cancer Res *64*: 8534-8540, 2004.

5 Amundadottir LT, Sulem P, Gudmundsson J, Helgason A, Baker A, Agnarsson BA, Sigurdsson A, Benediktsdottir KR, Cazier JB, Sainz J, Jakobsdottir M, Kostic J, Magnusdottir DN, Ghosh S, Agnarsson K, Birgisdottir B, Le Roux L, Olafsdottir A, Blondal T, Andresdottir M, Gretarsdottir OS, Bergthorsson JT, Gudbjartsson D, Gylfason A, Thorleifsson G, Manolescu A, Kristjansson K, Geirsson G, Isaksson H, Douglas J, Johansson JE, Balter K, Wiklund F, Montie JE, Yu X, Suarez BK, Ober C, Cooney KA, Gronberg H, Catalona WJ, Einarsson GV, Barkardottir RB, Gulcher JR, Kong A, Thorsteinsdottir U and Stefansson K: A common variant associated with prostate cancer in European and African populations. Nat Genet *38*: 652-658, 2006.

6 Bergamaschi A, Kim YH, Wang P, Sorlie T, Hernandez-Boussard T, Lonning PE, Tibshirani R, Borresen-Dale AL and Pollack JR: Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. Genes Chromosomes Cancer *45*: 1033-1040, 2006.

7 Bibikova M, Lin Z, Zhou L, Chudin E, Garcia EW, Wu B, Doucet D, Thomas NJ, Wang Y, Vollmer E, Goldmann T, Seifart C, Jiang W, Barker DL, Chee MS, Floros J and Fan JB: High-throughput DNA methylation profiling using universal bead arrays. Genome Res *16*: 383-393, 2006.

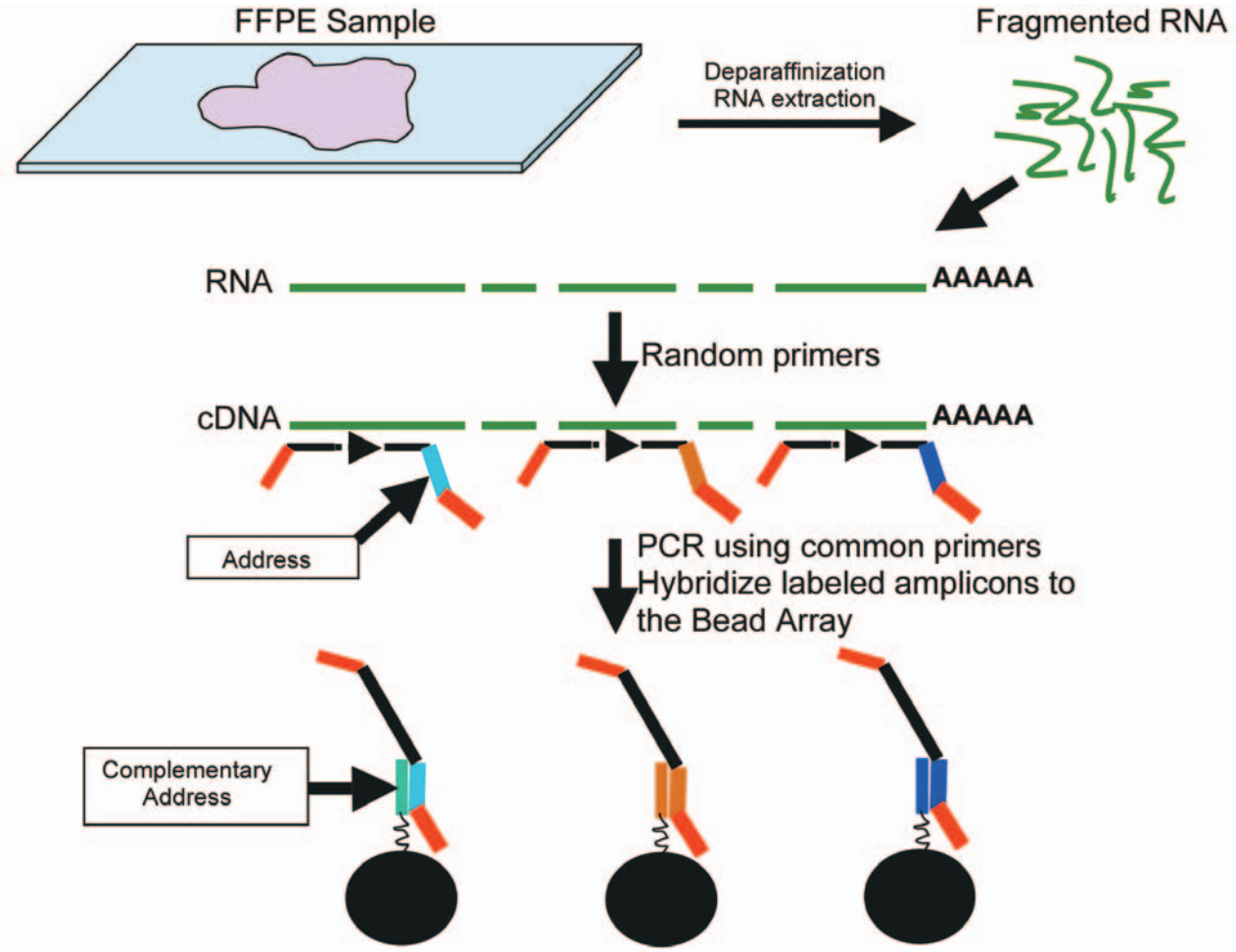

Figure 3. *The DASL assay. FFPE tumor tissue samples, 5-10 micron shaved sections or from microscope slides, are first subjected to deparaffinization and then RNA extraction. The extracted RNA is fragmented with an average size of ~200 nucleotides. Biotinylated cDNA is made from the fragmented RNA using random primers (9-mers). Complex oligonucleotides that contain a gene specific sequence and a unique address sequence are used to selectively bind to the cDNA. Extension and then ligation follows resulting in cDNA products which are subsequently amplified using fluorescently-labeled common primers. Depicted are three gene-specific oligonucleotide probe sets hybridized to cDNAs from a single gene. The amplified products contain 1,536 unique addresses (3 addresses per gene, therefore 512 genes) and are hybridized to the 50,000 bead-array (each unique complementary address bead at ~30-fold redundancy). The address sequence directs hybridization to their complements on the universal bead arrays.*

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

9 Bogaerts J, Cardoso F, Buyse M, Braga S, Loi S, Harrison JA, Bines J, Mook S, Decker N, Ravdin P, Therasse P, Rutgers E, van't Veer LJ and Piccart M: Gene signature evaluation as a prognostic tool: challenges in the design of the MINDACT trial. Nat Clin Pract Oncol *3*: 540-551, 2006.

10 Buyse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, d'Assignies MS, Bergh J, Lidereau R, Ellis P, Harris A, Bogaerts J, Therasse P, Floore A, Amakrane M, Piette F, Rutgers E, Sotiriou C, Cardoso F and Piccart MJ: Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst *98*: 1183-1192, 2006.

11 Chen WY, Bertone-Johnson ER, Hunter DJ, Willett WC and Hankinson SE: Associations between polymorphisms in the vitamin D receptor and breast cancer risk. Cancer Epidemiol Biomarkers Prev *14*: 2335-2339, 2005.

12 Collins C, Rommens JM, Kowbel D, Godfrey T, Tanner M, Hwang SI, Polikoff D, Nonet G, Cochran J, Myambo K, Jay KE, Froula J, Cloutier T, Kuo WL, Yaswen P, Dairkee S, Giovanola J, Hutchinson GB, Isola J, Kallioniemi OP, Palazzolo M, Martin C, Ericsson C, Pinkel D, Albertson D, Li WB and Gray JW: Positional cloning of ZNF217 and NABC1: genes amplified at 20q13.2 and overexpressed in breast carcinoma. Proc Natl Acad Sci USA *95*: 8703-8708, 1998.

13 Dawood S and Leyland-Jones B: Pharmacology and pharmacogenetics of chemotherapeutic agents. *In*: Breast Cancer: Translational Therapeutic Strategies. Lyman G, Burnstein H (eds.), pp. 45-56, 2007.

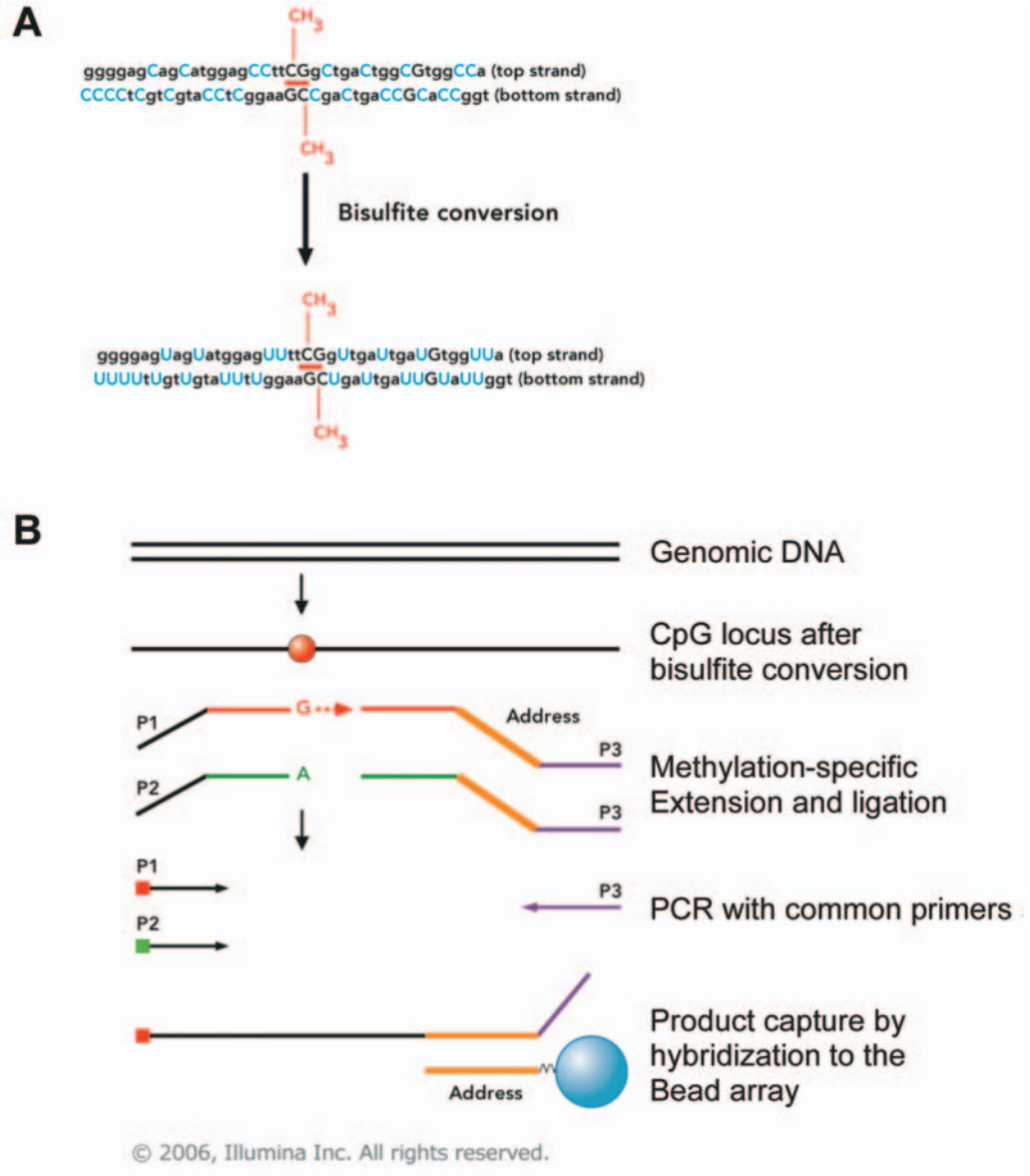

Figure 4. *The DNA methylation assay. (A) In order to perform the DNA methylation assay, genomic DNA must first undergo bisulfite conversion. Cytosines are converted to uracils but methyl-cytosine is unreactive. (B) In order to take advantage of this differential reactivity to bisulfite treatment, two pairs of complex oligonucleotide probe sets are designed for each CpG site. The first probe pair is made up of an allele-specific oligonucleotide (ASO) and locus-specific oligonucleotide (LSO) to interrogate the methylated state of the CpG site and a corresponding ASO-LSO pair for the unmethylated state. The ASO is used to determine if a site is methylated or not and also incorporates a universal PCR primer sequence P1 or P2. P1 and P2 are fluorescently labeled, each with a different dye, and associated with the "T" (unmethylated) allele or the "C" (methylated) allele, respectively. The LSO is made up of three parts, a CpG locus-specific sequence, an address sequence in the middle corresponding to a complementary address sequence on the bead-array, and a universal PCR priming site (P3) at the 3'-end. The pooled assay oligonucleotides are first annealed to bisulfite-converted genomic DNA. An allele-specific primer extension step is then carried out using ASOs which are extended only if their 3'-base is complementary to their cognate CpG site in the genomic DNA template. Allele-specific extension is followed by ligation of the extended ASOs to their corresponding LSOs, to create PCR templates. Common primers P1, P2, and P3 are then used to amplify the ligated products which are subsequently hybridized to a microarray bearing the complementary address sequences.*

14 De V, I, Hankinson SE, Colditz GA and Hunter DJ: A functional polymorphism in the progesterone receptor gene is associated with an increase in breast cancer risk. Cancer Res *63*: 5236-5238, 2003.
15 Esteller M: Cancer epigenomics: DNA methylomes and histone-modification maps. Nat Rev Genet *8*: 286-298, 2007.
16 Fan JB, Yeakley JM, Bibikova M, Chudin E, Wickham E, Chen J, Doucet D, Rigault P, Zhang B, Shen R, McBride C, Li HR, Fu XD, Oliphant A, Barker DL and Chee MS: A versatile assay for high-throughput gene expression profiling on universal array matrices. Genome Res *14*: 878-885, 2004.

17 Fan M, Yan PS, Hartman-Frey C, Chen L, Paik H, Oyer SL, Salisbury JD, Cheng AS, Li L, Abbosh PH, Huang TH and Nephew KP: Diverse gene expression and DNA methylation profiles correlate with differential adaptation of breast cancer cells to the antiestrogens tamoxifen and fulvestrant. Cancer Res *66*: 11954-11966, 2006.

18 Finke J, Fritzen R, Ternes P, Lange W and Dolken G: An improved strategy and a useful housekeeping gene for RNA analysis from formalin-fixed, paraffin-embedded tissues by PCR. Biotechniques *14*: 448-453, 1993.

19 Fridlyand J, Snijders AM, Ylstra B, Li H, Olshen A, Segraves R, Dairkee S, Tokuyasu T, Ljung BM, Jain AN, McLennan J, Ziegler J, Chin K, Devries S, Feiler H, Gray JW, Waldman F, Pinkel D and Albertson DG: Breast tumor copy number aberration phenotypes and genomic instability. BMC Cancer *6*: 96, 2006.

20 Godfrey TE, Kim SH, Chavira M, Ruff DW, Warren RS, Gray JW and Jensen RH: Quantitative mRNA expression analysis from formalin-fixed, paraffin-embedded tissues using 5' nuclease quantitative reverse transcription-polymerase chain reaction. J Mol Diagn *2*: 84-91, 2000.

21 Gunderson KL, Steemers FJ, Lee G, Mendoza LG and Chee MS: A genome-wide scalable SNP genotyping assay using microarray technology. Nat Genet *37*: 549-554, 2005.

22 Haiman CA, Brown M, Hankinson SE, Spiegelman D, Colditz GA, Willett WC, Kantoff PW and Hunter DJ: The androgen receptor CAG repeat polymorphism and risk of breast cancer in the Nurses' Health Study. Cancer Res *62*: 1045-1049, 2002.

23 Haiman CA, Hankinson SE, Spiegelman D, De V, I, Colditz GA, Willett WC, Speizer FE and Hunter DJ: A tetranucleotide repeat polymorphism in CYP19 and breast cancer risk. Int J Cancer *87*: 204-210, 2000.

24 Han J, Hankinson SE, Zhang SM, De VI and Hunter DJ: Interaction between genetic variations in DNA repair genes and plasma folate on breast cancer risk. Cancer Epidemiol Biomarkers Prev *13*: 520-524, 2004.

25 Han W, Han MR, Kang JJ, Bae JY, Lee JH, Bae YJ, Lee JE, Shin HJ, Hwang KT, Hwang SE, Kim SW and Noh DY: Genomic alterations identified by array comparative genomic hybridization as prognostic markers in tamoxifen-treated estrogen receptor-positive breast cancer. BMC Cancer *6*: 92, 2006.

26 Heller G, Geradts J, Ziegler B, Newsham I, Filipits M, Markis-Ritzinger EM, Kandioler D, Berger W, Stiglbauer W, Depisch D, Pirker R, Zielinski CC and Zochbauer-Muller S: Down-regulation of TSLC1 and DAL-1 expression occurs frequently in breast cancer. Breast Cancer Res Treat, 2007.

27 Huang TH, Perry MR and Laux DE: Methylation profiling of CpG islands in human breast cancer cells. Hum Mol Genet *8*: 459-470, 1999.

28 Ishibe N, Hankinson SE, Colditz GA, Spiegelman D, Willett WC, Speizer FE, Kelsey KT and Hunter DJ: Cigarette smoking, cytochrome P450 1A1 polymorphisms, and breast cancer risk in the Nurses' Health Study. Cancer Res *58*: 667-671, 1998.

29 Jemal A, Siegel R, Ward E, Murray T, Xu J and Thun MJ: Cancer statistics, 2007. CA Cancer J Clin *57*: 43-66, 2007.

30 Jing F, Zhang J, Tao J, Zhou Y, Jun L, Tang X, Wang Y and Hai H: Hypermethylation of tumor suppressor genes BRCA1, p16 and 14-3-3sigma in serum of sporadic breast cancer patients. Onkologie *30*: 14-19, 2007.

31 Johnson NA, Hamoudi RA, Ichimura K, Liu L, Pearson DM, Collins VP and Du MQ: Application of array CGH on archival formalin-fixed paraffin-embedded tissues including small numbers of microdissected cells. Lab Invest *86*: 968-978, 2006.

32 Joosse SA, van Beers EH and Nederlof PM: Automated array-CGH optimized for archival formalin-fixed, paraffin-embedded tumor material. BMC Cancer *7*: 43, 2007.

33 Kaklamani VG, Hou N, Bian Y, Reich J, Offit K, Michel LS, Rubinstein WS, Rademaker A and Pasche B: TGFBR1*6A and cancer risk: a meta-analysis of seven case-control studies. J Clin Oncol *21*: 3236-3243, 2003.

34 Kim C, Bryant J, Horne Z, Geyer CE, Wickerham DL, Wolmark N and Paik S: Trastuzumab sensitivity of breast cancer with co-amplification of HER2 and cMYC suggests pro-apoptotic function of dysregulated cMYC *in vivo*. San Antonio Breast Cancer Symposium 46. 2005. Ref Type: Abstract

35 Little SE, Vuononvirta R, Reis-Filho JS, Natrajan R, Iravani M, Fenwick K, Mackay A, Ashworth A, Pritchard-Jones K and Jones C: Array CGH using whole genome amplification of fresh-frozen and formalin-fixed, paraffin-embedded tumor DNA. Genomics *87*: 298-306, 2006.

36 Loo LW, Grove DI, Williams EM, Neal CL, Cousens LA, Schubert EL, Holcomb IN, Massa HF, Glogovac J, Li CI, Malone KE, Daling JR, Delrow JJ, Trask BJ, Hsu L and Porter PL: Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes. Cancer Res *64*: 8541-8549, 2004.

37 Martens JW, Nimmrich I, Koenig T, Look MP, Harbeck N, Model F, Kluth A, Bolt-de Vries J, Sieuwerts AM, Portengen H, Meijer-Van Gelder ME, Piepenbrock C, Olek A, Hofler H, Kiechle M, Klijn JG, Schmitt M, Maier S and Foekens JA: Association of DNA methylation of phosphoserine aminotransferase with response to endocrine therapy in patients with recurrent breast cancer. Cancer Res *65*: 4101-4117, 2005.

38 Nagata Y, Lan KH, Zhou X, Tan M, Esteva FJ, Sahin AA, Klos KS, Li P, Monia BP, Nguyen NT, Hortobagyi GN, Hung MC and Yu D: PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients. Cancer Cell *6*: 117-127, 2004.

39 Park K, Kwak K, Kim J, Lim S and Han S: c-myc amplification is associated with HER2 amplification and closely linked with cell proliferation in tissue microarray of nonselected breast cancers. Hum Pathol *36*: 634-639, 2005.

40 Perou CM, Jeffrey SS, van de RM, Rees CA, Eisen MB, Ross DT, Pergamenschikov A, Williams CF, Zhu SX, Lee JC, Lashkari D, Shalon D, Brown PO and Botstein D: Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proc Natl Acad Sci USA *96*: 9212-9217, 1999.

41 Perou CM, Sorlie T, Eisen MB, van de RM, 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 *406*: 747-752, 2000.

42 Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW and Albertson DG: High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet *20*: 207-211, 1998.

43 Press MF, Bernstein L, Sauter G, Zhou JY, Eiermann W, Pienkowski T, Crown J, Robert N, Bee V, Taupin H, Villalobos, I, Seelig S, Pegram M and Slamon DJ: Topoisomerase II-alpha gene amplification as a predictor of responsiveness to anthracycline-containing chemotherapy in the Cancer International Research Group 006 clinical trial of trastuzumab (herceptin) in the adjuvant setting. San Antonio Breast Cancer Symposium, 2005. Ref Type: Abstract
44 Reichmuth C, Markus MA, Hillemanns M, Atkinson MJ, Unni KK, Saretzki G and Hofler H: The diagnostic potential of the chromosome translocation t(2;13) in rhabdomyosarcoma: a Pcr study of fresh-frozen and paraffin-embedded tumor samples. J Pathol *180*: 50-57, 1996.
45 Richard F, Pacyna-Gengelbach M, Schluns K, Fleige B, Winzer KJ, Szymas J, Dietel M, Petersen I and Schwendel A: Patterns of chromosomal imbalances in invasive breast cancer. Int J Cancer *89*: 305-310, 2000.
46 Ross JS and Fletcher JA: The HER-2/neu oncogene in breast cancer: prognostic factor, predictive factor, and target for therapy. Stem Cells *16*: 413-428, 1998.
47 Rupp GM and Locker J: Purification and analysis of RNA from paraffin-embedded tissues. Biotechniques *6*: 56-60, 1988.
48 Sheils OM and Sweeney EC: TSH receptor status of thyroid neoplasms – TaqMan RT-PCR analysis of archival material. J Pathol *188*: 87-92, 1999.
49 Shen WH, Balajee AS, Wang J, Wu H, Eng C, Pandolfi PP and Yin Y: Essential role for nuclear PTEN in maintaining chromosomal integrity. Cell *128*: 157-170, 2007.
50 Shimada M, Imura J, Kozaki T, Fujimori T, Asakawa S, Shimizu N and Kawaguchi R: Detection of Her2/neu, c-MYC and ZNF217 gene amplification during breast cancer progression using fluorescence in situ hybridization. Oncol Rep *13*: 633-641, 2005.
51 Slamon DJ, Eiermann W, Pienkowski T, Martin M, Pawlicki M, Chan M, Smylie M, Liu M, Falkson C, Pinter T, Fornander T, Shiftan T, Valero V, Mackey J, Tabah-Fisch I, Buyse M, Lindsay MA, Riva A, Bee V, Pegram MD, Press M and Crown J: Phase III randomized trial comparing doxorubicin and cyclophosphamide followed by docetaxel (AC-T) with doxorubicin and cyclophosphamide followed by docetaxel and trastuzumab (AC-TH) with docetaxel, carboplatin and trastuzumab (TCH) in HER2 positive early breast cancer patients: BGIRG 006 study. SABCS, 2005. Ref Type: Abstract.
52 Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de RM, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein LP and Borresen-Dale AL: Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA *98*: 10869-10874, 2001.
53 Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, Martiat P, Fox SB, Harris AL and Liu ET: Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci USA *100*: 10393-10398, 2003.
54 Specht K, Richter T, Muller U, Walch A, Werner M and Hofler H: Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue. Am J Pathol *158*: 419-429, 2001.
55 Stanta G and Bonin S: RNA quantitative analysis from fixed and paraffin-embedded tissues: membrane hybridization and capillary electrophoresis. Biotechniques *24*: 271-276, 1998.
56 Tanner MM, Tirkkonen M, Kallioniemi A, Holli K, Collins C, Kowbel D, Gray JW, Kallioniemi OP and Isola J: Amplification of chromosomal region 20q13 in invasive breast cancer: prognostic implications. Clin Cancer Res *1*: 1455-1461, 1995.
57 Tokunaga E, Oki E, Kimura Y, Yamanaka T, Egashira A, Nishida K, Koga T, Morita M, Kakeji Y and Maehara Y: Coexistence of the loss of heterozygosity at the PTEN locus and HER2 overexpression enhances the Akt activity thus leading to a negative progesterone receptor expression in breast carcinoma. Breast Cancer Res Treat *101*: 249-257, 2007.
58 van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, 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 and Bernards R: A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med *347*: 1999-2009, 2002.
59 van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der KK, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R and Friend SH: Gene expression profiling predicts clinical outcome of breast cancer. Nature *415*: 530-536, 2002.
60 Wang Z, Yan P, Potter D, Eng C, Huang TH and Lin S: Heritable clustering and pathway discovery in breast cancer integrating epigenetic and phenotypic data. BMC Bioinformatics *8*: 38, 2007.
61 Yeager M, Orr N, Hayes RB, Jacobs KB, Kraft P, Wacholder S, Minichiello MJ, Fearnhead P, Yu K, Chatterjee N, Wang Z, Welch R, Staats BJ, Calle EE, Feigelson HS, Thun MJ, Rodriguez C, Albanes D, Virtamo J, Weinstein S, Schumacher FR, Giovannucci E, Willett WC, Cancel-Tassin G, Cussenot O, Valeri A, Andriole GL, Gelmann EP, Tucker M, Gerhard DS, Fraumeni JF Jr, Hoover R, Hunter DJ, Chanock SJ and Thomas G: Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet, 2007.
62 Yeon CH and Pegram MD: Anti-erbB-2 antibody trastuzumab in the treatment of HER2-amplified breast cancer. Invest New Drugs *23*: 391-409, 2005.
63 Yoshimoto M, Cutz JC, Nuin PA, Joshua AM, Bayani J, Evans AJ, Zielenska M and Squire JA: Interphase FISH analysis of PTEN in histologic sections shows genomic deletions in 68% of primary prostate cancer and 23% of high-grade prostatic intra-epithelial neoplasias. Cancer Genet Cytogenet *169*: 128-137, 2006.

Received April 12, 2007
Accepted April 18, 2007