International Journal of Education and Science Research Review
E-ISSN 2348-6457
July –August - 2021, Volume-8, Issue-4 P-ISSN 2349-1817
www.ijesrr.org Email- editor@ijesrr.org

VALIDATION OF PUTATIVE IDENTIFIED GENES IN BREAST CANCER TISSUE AS BIOMARKERS FOR EARLY DETECTION OF LOBULAR AND DUCTAL BREAST CANCER

Tamanna Walia

Research scholar, Kalinga University Naya Raipur Chhattisgarh, India

Dr.Sushma Dubey

Prof.&HOD, Department of Biotechnology, Kalinga University Naya Raipur Chhattisgarh, India

ABSTARCT

Breast cancer, a prevalent global health concern, comprises distinct histological subtypes, notably lobular and ductal, each posing unique challenges for early detection and targeted treatment. Despite advancements in genomic research identifying putative biomarkers, their clinical translation remains a critical gap. This study focuses on the validation of potential biomarkers in breast cancer tissue to facilitate early detection, with a specific emphasis on discriminating between lobular and ductal subtypes. Utilizing molecular profiling techniques, we aim to validate and refine the identified genes associated with each subtype, establishing their reliability as biomarkers. The validation process involves the analysis of breast cancer tissue samples, correlating molecular signatures with clinical outcomes. This comprehensive approach seeks to bridge the translational gap between genomics and clinical utility, fostering a deeper understanding of the molecular intricacies specific to lobular and ductal breast cancers. The anticipated outcomes of this research hold the potential to revolutionize breast cancer diagnostics, offering more precise and personalized approaches to early detection and intervention. By elucidating the molecular underpinnings of each subtype, we aspire to contribute to the development of targeted therapies, ultimately improving patient prognosis and reducing the overall burden of breast cancer. This study represents a critical step towards realizing the promise of genomic discoveries in the advancement of breast cancer management.

Keywords:*Validation, Breast Cancer, Biomarkers*

INTRODUCTION

Introduction

Breast cancer remains a formidable global health challenge, constituting a significant cause of morbidity and mortality among women. In the intricate landscape of breast cancer, tw o predominant histological subtypes, lobular and ductal, present distinct clinical and molecular characteristics. The urgency of early detection and accurate classification of these subtypes is paramount, as it directly influences treatment strategies and patient outcomes. Despite advancements in diagnostic technologies, identifying reliable biomarkers for the precise detection of lobular and ductal breast cancers at an early stage remains an elusive goal.

In recent years, genomic research has opened new avenues for understanding the intricate molecular underpinnings of breast cancer. Numerous putative genes have been proposed as potential biomarkers, holding promise for enhancing early detection and, consequently, improving patient prognosis. However, the translation of these putative biomarkers from research findings to clinical utility requires rigorous validation in clinically relevant settings, such as breast cancer tissue samples.

This study focuses on the validation of putative identified genes as biomarkers for early detection, with a specific emphasis on discriminating between lobular and ductal breast cancers. By scrutinizing the molecular signatures associated with these distinct subtypes, we aim to contribute to the refinement of diagnostic tools and the development of targeted therapies. The validation of these biomarkers is a critical step toward personalized medicine in breast cancer, offering the potential for more accurate diagnoses and tailored treatment plans.

Through comprehensive validation efforts, including molecular profiling of breast cancer tissues and correlation with clinical outcomes, this research endeavors to bridge the gap between genomic discoveries and clinical application. The outcomes of this study may not only enhance our understanding of the molecular landscape of lobular and ductal breast cancers but also pave the way for the development of novel and effective strategies for early detection and intervention. In doing so, we aspire to make significant strides in improving the overall management of breast cancer and ultimately reducing the burden of this devastating disease on women's health.

OBJECTIVE OF THE STUDY

1. To systematically validate and assess the potential of specific genes identified in prior research as biomarkers for early detection in breast cancer, with a particular emphasis on distinguishing between lobular and ductal breast cancer subtypes.

2. Examine the expression levels of identified genes in both normal breast tissue and breast cancer tissue samples, with a focus on understanding differential expression between lobular and ductal carcinomas.

3. Investigate the specificity of the identified biomarkers in distinguishing between lobular and ductal breast cancer subtypes, aiming to provide insights into the potential clinical relevance of these biomarkers in subtype-specific detection.

RESEARCH METHODOLOGY

Blood samples

For the purpose of obtaining a permission from the Regional Ethical Committee of Norway, blood samples were acquired from donors after receiving their informed consent. During the analysis, it was ensured that each donation remained anonymous. Before it was known whether the abnormality that was discovered during the initial screening was benign or malignant, blood was collected from women who had a suspicious initial mammogram. In each and every instance, the blood samples were collected between the hours of eight in the morning and four in the afternoon. Using either vacutainer tubes containing ethylenediaminetetraacetic acid as an anticoagulant or directly placing the blood in PAXgene tubes, a volume of 10 millilitres of blood was extracted from each individual lady by trained workers.

Preparation of cDNA arrays

A total of one thousand four hundred thirty-five cDNA clones were selected at random from a plasmid library that was created using the entire blood of five hundred and fifty healthy persons. According to the results of the sequencing analysis of more than 500 cDNAs, it was estimated that around twenty percent of the randomly selected clones had duplicates. The process of amplification of inserts involved the cultivation of bacterial clones in microtiter plates that contained 150 µl of Luria Broth media containing 50 µg/ml of carbenicillin. These plates were then incubated overnight at 37°C with agitation. In order to lyse the cells, 5 µl of each culture was diluted with 50 µl of distilled water and stored at 95°C for a period of 12 minutes. Out of this mixture, a PCR reaction was performed using 40 µmol of 5' – and 3' – sequencing primers in the presence of 1.5 mM MgCl2. The reactions were carried out with a volume of 2 µl. There were PCR reactions carried out using the cycling methodology that is described below.

The assay's background level, consistency, and sensitivity were all evaluated using the printed arrays, which also contained controls for the evaluation. In addition to the 1,435 cDNAs, these were also spotted at multiple positions. Furthermore, they included controls such as PCR mix (without any insert), controls of the Spot Report 10-array validation system, and cDNAs that corresponded to genes that are constitutively expressed. These genes include β-actin, γ-actin, glyceraldehyde-3-phosphate dehydrogenase, human ornithine decarboxylase, and cyclophilin.

RNA extraction, probe synthesis, and hybridization

The total RNA was purified in accordance with the instructions provided by the supplier after the blood that had been collected in EDTA tubes was thawed at 37 degrees Celsius and then transferred to PAX tubes. Following the procedure described above, total RNA was extracted from blood that was collected directly in PAX tubes. There was no transfer of the RNA to any additional tubes. A DNA-free kit was utilised in order to perform a DNase I treatment on the extracted RNA in order to eliminate any contaminating DNA. Following agarose-gel electrophoresis, the integrity of the 28S and 18S ribosomal bands was visually examined in order to evaluate the quality of the RNA.

For the purpose of this investigation, only samples from which high-quality RNA was extracted were utilized. According to our observations, blood taken in EDTA tubes frequently produced RNA of low quality, but blood collected in PAX tubes nearly invariably produced RNA of high quality. This was done by measuring the absorbance at 260 nm and 280 nm, which allowed for the determination of the concentration and purity of the extracted RNA. To extract microRNA (mRNA) from total RNA, Dyna beads were utilized in line with the directions provided by the service provider.

The studies involving hybridization and labeling were carried out in sixteen separate batches. Within each batch, the number of samples that were analyzed ranged anywhere from six to nine. Only the arrays that were produced during the same print run were utilised in each batch. This was done in order to reduce the amount of noise that was caused by variations in printing from batch to batch.

Quantification of hybridization signals

Phosphoscreens with a super resolution were used to expose the hybridised membranes, and a picture file was created with the help of the Phosphor Imager. The Phoretic programme (Nonlinear Dynamics, Newcastle upon

Tyne, United Kingdom) was utilised in order to carry out the tasks of identifying and quantifying the hybridization signals, as well as subtracting the values of the local background. In order to perform background subtraction, the intensity of the signals that were evaluated in each spot was subtracted from the median of the line of pixels that surrounded each spot outline.

DATA ANALYSIS

One and a quarter percent of the lowest signals and one and a half percent of the highest signals were removed from each membrane using the background-subtracted data for 1,435 genes. The values of 67 cDNAs in total were deleted from all membranes, and the expression data for just 1,368 genes were further analyzed. This was done since the cDNAs that had signals that fell within this range differed from membrane to membrane. After performing a cube-root transformation, the data were normalized by first dividing the value of each location by the mean of the signals in each array, and then performing the transformation. There is a distinct batch effect in the data that has been normalized using the cube root, as seen in the supplementary figure 1 (similar effects were also obvious in the raw data). For the purpose of adjusting for the batch effects, a straightforward one-way analysis of variance was carried out. Through the use of the ANOVA correction, the systematic batch effects were eliminated, as seen in the supplementary figure 1. After the batch-adjusted data were processed, the nearest-shrunken-centroid approach was utilized to do the analysis.

The ideal shrinkage threshold is determined by the utilization of this method, which employs the typical 'external' cross-validation technique. After that, the ideal threshold is utilized in conjunction with the entire training set in order to design the centroid. The conclusion that can be drawn from this is that the estimate of cross-validation error that is derived is nearly unbiased for the actual test-error rate, regardless of the value of the threshold.

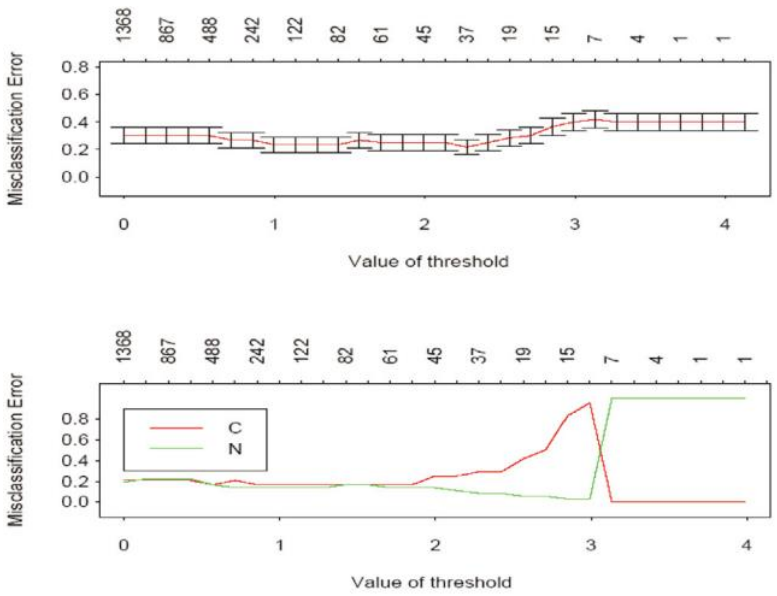

Figure 1. The rate of incorrect categorization as a function of the threshold The percentage of incorrect classifications as a function of the threshold value and the total number of genes involved

RESULTS

Table 1 presents the results of our analysis of gene-expression patterns in sixty blood samples acquired from fifty-six distinct women. There were sixteen different batches of trials carried out. The purpose of this study was to explore the repeatability of the results. A total of 102 experimental samples were analyzed, with 13 samples coming from women who had breast cancer and 23 samples coming from women who did not have breast cancer. The samples were analyzed in various batches using aliquots from the same mRNA pool.

Preprocessing was performed on the expression data that was generated, and then the nearest-shrunken-centroid technique was used to analyze the data. It was determined that the best amount of shrinkage threshold could be found by the use of a typical leave-one-out cross-validation technique. Due to the fact that we had sixty distinct blood samples and that some of the tests were repeated more than once, the data were separated into sixty non- overlapping subsets for the purpose of cross-validation. Each of these subsets represented a different blood sample and contained all of the repetitions that were included in the data set. Only when the majority of members in the appropriate cross-validation segment were correctly categorized was a sample considered to have been correctly classified on the basis of the classification. Following the application of a threshold value of 2.28, the total misclassification error was found to be at its lowest, resulting in a selection of 37 genes (Fig. 1). At this point, ten of the 57 samples were incorrectly categorized, and three samples were deemed to be non decisions. This was due to the fact that there was no majority for either the breast-cancer or non-breast-cancer class (Table 2). Table 1 is a comprehensive listing of the results of the forecast.

When it came to samples taken from women who were in the early stages of breast cancer, specifically stage 0 and stage I, the prediction was quite accurate. There was one sample that did not make a decision out of the fourteen samples that represented the early phases, and eleven out of thirteen samples were accurately predicted. The predictions were accurate for five out of seven samples from stage II and one out of two samples from stage.

Table 1. Proteins that are poorly controlled by the IDC and are implicated in ECM/TME remodelling and EMT

Dysregulated Proteins	Genes	Proteomics -Based Methods	Functions	Associated Roles in Cancer
Tenascin	TNC	LC–MS/MS	ECM protein	partial EMT marker; cell adhesion, tissue remodeling, transduction of cellular signaling pathways
Collagen isoforms	COL1A1, COL1A2, COL14A 1	LC–MS/MS; MALDI-FT-ICR MSI;	TME/ECM protein	cancer fibrosis, EMT

		HRAM, nanoLC-ESI–MS/MS		
Fibronectin	FN1	LC–MS/MS	component of the mammary mesenchymal compartment of breast tumor	cell invasion, metastasis, tumor progression, EMT
Periostin	POSTN/OSF-2	FFPE, LCM, IHC, RT-PCR; LC–MS/MS	secreted ECM cell adhesion glycoprotein	EMT, proliferation, adhesion, migration
Thrombospondins	THBS1/TSP1, THBS2/TSP2	LC–MS/MS	ECM proteins	cell adhesion, invasion, migration, proliferation, apoptosis, tumor immunity
Decorin	DCN	LC–MS/MS	small leucine-rich ECM proteoglycan	overexpression decreases migration, invasion, stemness and tumor growth and metastasis
Lumican	LUM	LC–MSE, MALDI-MS/MS	small leucine-rich ECM proteoglycan	EMT regulator
Mimecan/osteoglycin	OGN	LC–MS/MS	small leucine-rich ECM proteoglycan	inhibits BC cell proliferation and reverses EMT via repressing PI3K/AKT/mTOR pathway

Matrix metalloproteinases	MMP-2, MMP-9	2-DE, MALDI-ToF MS	Zn-dependent endopeptidases	ECM remodeling, tumor initiation, progression, metastasis

Table 2. Investigation of protein isoforms in IDC using proteomics as the source of data

Protein Isoforms	Biological Samples	Other Conventional Analytical and Coupled Methods	MS-Based Proteomics	Results	Functions
Isoforms of the folate receptor (FRα and FRβ); the possibility of an isoform-based diagnostic in breast cancer	BC cells lines and IDC tissue	WB, IHC	LC-ESI–MS/MS	quantitative analysis of FR isoforms that is both simultaneous and accurate: BC cells and tissue samples have an elevated expression of FRα, while TAMs exhibit a high abundance of FRβ.	unidirectional folate transport into cells
The ratios of progesterone receptor A and B, as well as PRA/PRB, with the course of breast cancer	BC cell line model	SDS-PAGE	HPLC–MS/MS	alterations in the proteome of breast cancer that are unique to isoforms; high PRA/PRB ratios in breast cancer that are related with resistance to	cell metabolism, cell cycle, apoptosis

				treatment and a bad prognosis	
Haptoglobin and α1-AT precursor isoforms	serum	2-DE; FFPE tissue sections-IHC	MALDI-MS	DEPs; identification of novel serum biomarkers in IDC patients compared with healthy women	possible role in tumor growth
Alternative splicing of ceramide synthase 2 (AS CERS2)	BC cell lines, IDC and adjacent normal tissue	RT-PCR, WB, SDS-PAGE, IHC	LC–MS/MS	higher expression of AS CERS2 in luminal B IDC	A dysregulated sphingolipid pathway, the beginning of cancer, cell proliferation and migration, cell survival, and apoptosis are all components of cancer.

The misclassification error was also calculated by taking an average of the class probability for each sample across all sixty cross-validation segments. This was in contrast to our previous method, which determined that a sample was only considered to be correctly classified if the majority of members in the cross-validation segment in question were correctly classified. As a result, each segment indicated an average class probability for each sample, and we projected that each sample would belong to the class that had the greatest average probability.

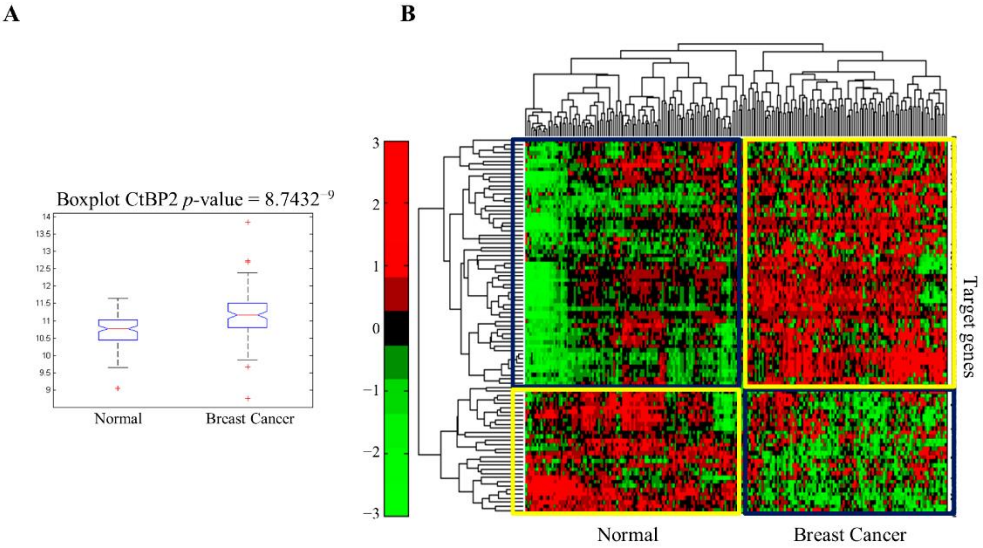

Figure 2. An analysis of human breast cancer using transcriptomics techniques

The ability to reach a conclusion regarding class membership that was accepted by all members of the group was the primary motivation behind the adoption of this strategy. Using the average-class strategy, the least error rate was attained at a threshold value of 2.42. This approach comprised a subset of just 25 genes, which resulted in an additional reduction of 12 genes (Fig. 2) (Additional file 4). Additionally, out of the sixty samples, ten were incorrectly categorised, seven of which were breast cancer samples and three of which were samples that did not include breast cancer. This is a somewhat better result than the one obtained with 37 genes, where there were three nondecisions (Fig. 3).

Table 3. Antigens used for generation of polyclonal antibodies

Antigen	Full name	NCBI reference	DNA fragment, bp
LGALS3BP	The lectin, galactosidase-binding, soluble, and three-binding proteins that are found in Homo sapiens	NM_005567.2	1483-1686
RAD50	S. cerevisiae cells that are similar to Homo sapiens RAD50	NM_005732.2	2552-3374
FAM50A	Homo sapiens family with sequence similarity 50, member A	NM_004699.1	76-1095
RBPJ	Recombination signal binding protein for the immunoglobulin kappa J region of the Homo sapiens genome	NM_203284.1	492-1850
PABCP4	The inducible version of the poly(A) binding protein found in the cytoplasm of Homo sapiens:	NM_003819.2	927-3052
LRRFIP1	Protein 1 of the Homo sapiens species that interacts with leucine rich repeats (in FLII)	NM_004735.2	417-1804

The shrunken t-statistic scores of the 37 predictive genes that were chosen for the purpose of comparing the breast-cancer class to the no breast-cancer class are presented in Table 3. Additionally, the genes in the public databases to which they exhibit sequence similarity as well as their suspected biological function are also included in this table. Figure 2 illustrates the relative expression of 12 predicted genes that have received the highest ratings. When compared to the samples taken from breast cancer patients, the majority of the prognostic genes (29 out of 37) had a lower level of expression, representing a positive score. By partly sequencing the

spotted cDNA clones that corresponded to the predicted genes and searching for gene similarities in public databases, the identification of the predictive genes was discovered.

Table 4. The currently available biomarkers that show promise for the diagnosis of breast cancer

Name of biomarker	Technology used for discovery	Type
RS/DJ-1	Humoral response	autoantibody
p53	Humoral response	autoantibody
HSP60	Humoral response	autoantibody
HSP90	Humoral response	autoantibody
Mucin-related	Humoral response	autoantibody
CA 15-3	Serum profiling	serum protein
RS/DJ-1	Serum profiling	serum protein
HER-2/neu	Serum profiling	serum protein
α-2-HS-glycoprotein	Nipple aspirate fluid profiling	Ductal protein
Lipophilin B	Nipple aspirate fluid profiling	Ductal protein
beta-globin	Nipple aspirate fluid profiling	Ductal protein
Hemopexin	Nipple aspirate fluid profiling	Ductal protein
Vitamin D-binding protein	Nipple Aspirate Fluid Profiling	Ductal protein

There are eight predictive genes that have been shown to have higher expression in breast cancer patient samples. Two of these genes encode the histone replacement protein H3.3, which is believed to be involved in the process of chromatin remodelling. Additionally, there are six genes that encode proteins that may be concerned with defense-related tasks. Ferritin and calgranulin B were found to be encoded by four genes that had enhanced expression. The intracellular storage and sequestration of iron are both processes that ferritin is involved in. In response to an oxidant challenge, it has been demonstrated that an increase in the expression of ferritin can minimise the formation of reactive oxygen species in HeLa cells. It is possible that cal granulin B has a function in the defence of the host since it is produced by blood cells both during an infection and during an inflammatory response.

It has been suggested that interferon-induced transmembrane protein 2 plays a role in the immune response. On the other hand, it is believed that human granule proteoglycan peptide core forms stable complexes with proteases and other granule-localized proteins in order to prevent the intragranular autolysis of these proteins

and to facilitate their concerted action extracellularly. Furthermore, it is interesting to note that the majority of the predictive genes that were discovered in this research belonged to the family of genes that displayed changed expression in neutrophils following stimulation by both nonvirulent and virulent bacterial stimuli.

CONCLUSION

MS-based technology should be the method of choice in current surgical oncology since it has the capability to perform sensitive, quick, and accurate "proteome point sampling" and "proteome point characterization" in biological tissues for the purpose of breast cancer profiling and for the identification of cancer kinds or subtypes. MS-based technology is also used for the molecular intraoperative characterization of healthy and tumour tissue in just a few seconds. This is accomplished through the utilisation of contemporary sampling techniques, such as the handled MasSpec Pen or the PIRL-DESI MSI, which are capable of performing in vivo proteomics-based analyses or involve minimal tissue removal. For the purpose of enhancing the identification of low-molecular-weight (LMW) or low-abundance proteins and protein fragments that are present in bodily fluids in extremely low concentrations and are "invisible" to shotgun proteomics, sample preparation techniques may be engineered to capture and enrich this portion of the proteome. Despite the fact that top-down proteomics has the potential to reveal significant differences between ductal non-invasive and invasive breast cancer tissues, as well as significant differentially expressed intact proteoforms with a biomarker value, only a small number of studies have utilized proteomics-based analyses of IDC-associated proteoforms in breast primary tumours or liquid biopsies. These methods must to be analysed and taken into consideration in order to achieve quick and sensitive ex vivo and/or in vivo MS profiling, as well as to accurately differentiate and delimitate the various kinds of tissue in intradermal cellular coagulation (IDC).

REFERENCES

1. A. Jemal, R. Siegel, J. Xu, and E. Ward, “Cancer statistics, 2010,” *CA Cancer Journal for Clinicians*, vol. 60, no. 5, pp. 277–300, 2010.

2. Early Breast Cancer Trialists' Collaborative Group (EBCTCG), “Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials,” *The Lancet*, vol. 365, pp. 1687–1717, 2005.

3. J. M. Nabholtz, A. Buzdar, M. Pollak et al., “Anastrozole is superior to tamoxifen as first-line therapy for advanced breast cancer in postmenopausal women: results of a North American multicenter randomized trial,” *Journal of Clinical Oncology*, vol. 18, no. 22, pp. 3758–3767, 2000.

4. J. Bonneterre, B. Thurlimann, J. F. Robertson et al., “Anastrozole versus tamoxifen as first-line therapy for advanced breast cancer in 668 postmenopausal women: results of the tamoxifen or arimidex randomized group efficacy and tolerability study,” *Journal of Clinical Oncology*, vol. 18, no. 22, pp. 3748–3757, 2000.

5. J. Bonneterre, A. Buzdar, J. M. Nabholtz et al., “Anastrozole is superior to tamoxifen as first-line therapy in hormone receptor positive advanced breast carcinoma,” *Cancer*, vol. 92, no. 9, pp. 2247–2258, 2001.

6. C. A. Hudis, "Trastuzumab—Mechanism of action and use in clinical practice," *New England Journal of Medicine*, vol. 357, no. 1, pp. 39–51, 2007.

7. D. Bilancia, G. Rosati, A. Dinota, D. Germano, R. Romano, and L. Manzione, "Lapatinib in breast cancer," *Annals of Oncology*, vol. 18, supplement 6, pp. vi26–vi30, 2007.

8. G. M. Higa and J. Abraham, "Lapatinib in the treatment of breast cancer," *Expert Review of Anticancer Therapy*, vol. 7, no. 9, pp. 1183–1192, 2007.

9. R. W. Carlson, B. O. Anderson, H. J. Burstein et al., "Invasive breast cancer: clinical practice guidelines in oncology," *Journal of the National Comprehensive Cancer Network*, vol. 5, no. 3, pp. 246–312, 2007.

10. B. J. Yoder, E. J. Wilkinson, and N. A. Massoll, "Molecular and morphologic distinctions between infiltrating ductal and lobular carcinoma of the breast," *Breast Journal*, vol. 13, no. 2, pp. 172–179, 2007.

11. F. A. Tavassoli, "Invasive lobular carcinoma," in *Pathology of the Breast*, F. A. Tavassoli, Ed., pp. 426–36, Appleton & Lange, Stamford, Conn, USA, 2nd edition, 1999.

12. "Invasive lobular carcinoma," in *Rosen's Breast Pathology*, P. P. Rosen, Ed., pp. 627–52, Lippincott Williams & Wilkins, Philadelphia, Pa, USA, 2nd edition, 2001.

13. J. S. Reis-Filho, P. T. Simpson, T. Gale, and S. R. Lakhani, "The molecular genetics of breast cancer: the contribution of comparative genomic hybridization," *Pathology Research and Practice*, vol. 201, no. 11, pp. 713–725, 2005.

14. S. E. Shackney and J. F. Silverman, "Molecular evolutionary patterns in breast cancer," *Advances in Anatomic Pathology*, vol. 10, no. 5, pp. 278–290, 2003.

15. P. T. Simpson, J. S. Reis-Filho, T. Gale, and S. R. Lakhani, "Molecular evolution of breat cancer," *Journal of Pathology*, vol. 205, no. 2, pp. 248–254, 2005.

16. D. E. Stange, B. Radlwimmer, F. Schubert et al., "High-resolution genomic profiling reveals association of chromosomal aberrations on 1q and 16p with histologic and genetic subgroups of invasive breast cancer," *Clinical Cancer Research*, vol. 12, no. 2, pp. 345–352, 2006.

17. D. G. Albertson, C. Collins, F. McCormick, and J. W. Gray, "Chromosome aberrations in solid tumors," *Nature Genetics*, vol. 34, no. 4, pp. 369–376, 2003.

18. D. G. Albertson and D. Pinkel, "Genomic microarrays in human genetic disease and cancer," *Human Molecular Genetics*, vol. 12, no. 2, pp. R145–R152, 2003.