MOLECULAR ANALYSIS OF PRE- AND POSTOPERATIVE BIOPSIES IN BREAST CANCER PROGRESSION

Margit Leonie Hesla Riis

UiO : Institute of Clinical Medicine
University of Oslo

University of Oslo
Faculty of Medicine
Institute of Clinical Medicine
Division of Surgery

and

Akershus University Hospital

Akershus University Hospital
Department of Surgery

Series of dissertations submitted to the
Faculty of Medicine, University of Oslo
No. 1532

ISBN 978-82-8264-512-6

Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika publishing.
The thesis is produced by Akademika publishing merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.

Table of contents

Acknowledgements

I would like to thank my supervisors Vessela Kristensen and Ida Bukholm for making this study possible. They have contributed greatly to the planning and the processing of the study. The study could not have proceeded without my good friend and bioengineer Torben Luders. He is the main driving force behind all laboratory work, most of the statistical work and the technical details of the manuscripts. He has introduced me to the world of molecular genetics and he has been patient with me in laboratory work. I would like to thank my collegue surgeons at Akershus University Hospital and the plastic surgeons at The Collosseum Clinic in Oslo for contributing to the biopsies taken. I am also greatful to the radiologist at Akershus University Hospital which also contributed to the tissue collection.

I am truly greatful to my current leader, Dr Ellen Schlichting for giving me the opportunity to complete my surgical speciality of Breast and Endocrine Surgery parallell to writing my thesis. She has given me good advice and good support both concerning the surgical speciality and the PhD studies.

I also want to thank all the women participating in the study. They give their time, tissue and information in order to prevent breath cancer deaths in future generations.

I greatly acknowledge the University of Oslo for admitting me into the PhD program and the South-Eastern Regional Health Aurthorithy for the grants and financial support of the project. I would also like to thank the opponents for taking time to read, evaluate and discuss my thesis.

Last but not least I would like to thank my dear family. My husband Rolf introduced me to the complex world of surgery. He has thouhgt me most of what I know within surgical technique. He is my best friend, my partner in life and a great father to my three beautiful children. He

has been greatly supportive during this period of my career when I have been working on my thesis. My daughter Leonie and my two sons, Eirik and Marius has inspired me through days and night. Seeing their beatiful faces and hearing their laughs keeps me going. My family is all that really matters. Thank you for being there for me.

AIMS OF THE STUDY

- To evaluate the gene expression profiles in preoperative tissue (core needle biopsies) while the tumor is still in its normal milieu to postoperative tissue from the same tumor obtained during surgery (Paper I).
- To find a gene expression profile which will best characterize the two most common groups of tumors: T1 and T2 tumor categories, which are the most important factors of treatment decision (Paper II).
- To use this profile for more refined selection of patients who would benefit from adjuvant treatment, preventing possible severe side effects, and unnecessary health expenses.
- Since the results from paper II suggested progression related activation of self renewal pathways, we aimed at characterising the expression of key members of this pathway: Bmi and Mel18 at both mRNA and protein level.

List of Papers

Paper I

Molecular profiles of pre- and postoperative breast cancer tumours reveal differentially expressed genes

Margit L. H. Riis, Torben Lüders, Elke K. Markert, Vilde D. Haakensen, Anne-Jorun Nesbakken, Vessela N. Kristensen, and Ida R. K. Bukholm

Published in ISRN Oncology 2012

Paper II

Gene Expression Profile Analysis of T1 and T2 Breast Cancer Reveals Different Activation Pathways

Margit L. H. Riis, Xi Zhao, Fateme Kaveh,Hilde S. Vollan, Anne-Jorunn Nesbakken, Hiroko K. Solvang, Torben Lüders, Ida R. K. Bukholm, and Vessela N. Kristensen

Published in ISRN Oncology 2013

Paper III

Expression of BMI-1 and Mel-18 in breast tissue – a diagnostic marker in patients with breast cancer

Margit Riis, Torben Lüders, Anne-Jorunn Nesbakken, Hilde Vollan, Vessela Kristensen, Ida Bukholm

Published in BMC Cancer 2010

1. Introduction

1.1 Epidemiology - Incidence

Breast cancer is by far the most frequent cancer among women with an estimated 1.38 million new cancer cases diagnosed in 2008 (23% of all cancers), and ranks second overall (10.9% of all cancers). It is now the most common cancer both in developed and developing regions with around 690 000 new cases estimated in each region (population ratio 1:4). Incidence rates vary from 19.3 per 100,000 women in Eastern Africa to 89.7 per 100,000 women in Western Europe, and are high (greater than 80 per 100,000) in developed regions of the world (except Japan) and low (less than 40 per 100,000) in most of the developing regions(1). The age-adjusted incidence worldwide is shown in figure 1.

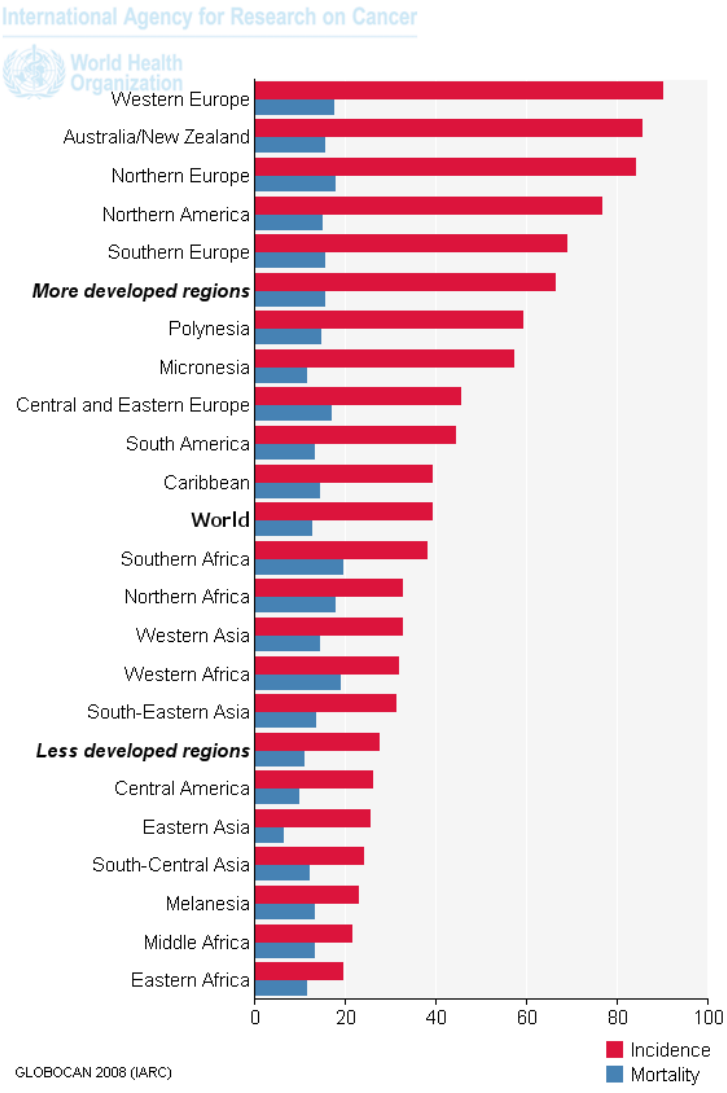


Figure 1 Age adjusted incidence worldwide(1)

The incidence of breast cancer has been rising dramatically since 1982, based on data collected by SEER(2). Mammographic screening has led to the finding of breast cancer at a preclinical stage. Naturally incidence has increased after the introduction of mammographic

screening. Effects of routine mammography on breast cancer mortality is highly dependent on study design(3, 4). In 2009, 27 520 new cases of cancer were recorded in Norway, for which 14 792 occurred among men and 12 728 among women(5). Breast cancer remains the most frequent neoplasm in women in Norway, as in the rest of the world, with 2745 new cases in 2009, followed by colorectal and lung cancer(5). The trend in incidence in Norway is shown in figure 2(5).

A B

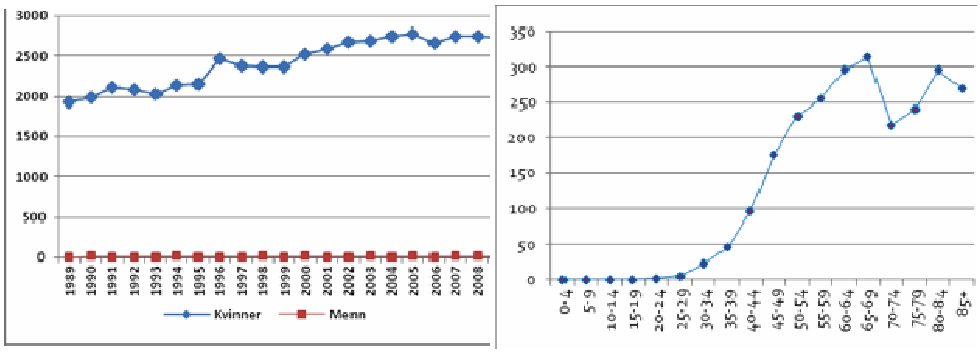


Figure 2. Number of new cases of breast cancer in Norway per year 1989 – 2009 (A) and age-adjusted incidence by age at diagnosis 2005 - 2009(5)

1.2 Mortality

Trends in breast cancer mortality are of major public interest, but their interpretation is complex because they reflect the combined effects of trends in underlying risk of breast cancer, changes in screening practices, and effectiveness of treatment. There is a marked difference in mortality rates in black women in all age groups compared to whites. The increase is significant and there is no evidence of recent decline, as has been seen for the white women(6). Mortality rates by race taken from SEER(2) is shown in table 1. The

reasons for these differences in mortality between different ethinic groups are not clear, but probably may due to differences in molecular biological signature of these tumors. Age-standardised mortality rates of breast cancer in Norway is 12.8%. Only lungcancer has a higher mortality rate (16.8%)(5).

Table 1 The age-adjusted death rate was 23.0 per 100,000 women per year. These rates are based on patients who died in 2005-2009 in the US.

Death Rates by Race	
Race/Ethnicity	**Female**
All Races	23.0 per 100,000 women
White	22.4 per 100,000 women
Black	31.6 per 100,000 women
Asian/Pacific Islander	11.9 per 100,000 women
American Indian/Alaska Native [a]	16.6 per 100,000 women
Hispanic [b]	14.9 per 100,000 women

1.3 Survival

5-year survival has improved significantly over the years. Prognosis of breast cancer is naturally dependent on stage of the disease. In patients where the disease is localized to the breast, 5-year relative survival is 94.1% as opposed to 16.9% where distant metastasis is diagnosed(7). Figure 3 shows 5-year survival in 1956 – 2000 in Norway, all stages, and figure 4 shows five year survival divided by stage of disease, also in Norway(7).

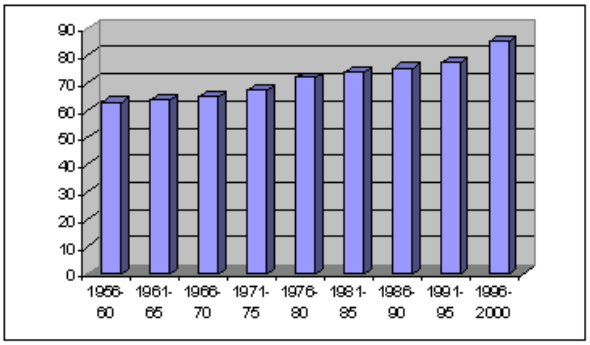


Figure 3. 5-year survival of breast cancer has changed over the years. This figure show 5-year survival in 1956 – 2000, all stages(7) in Norway.

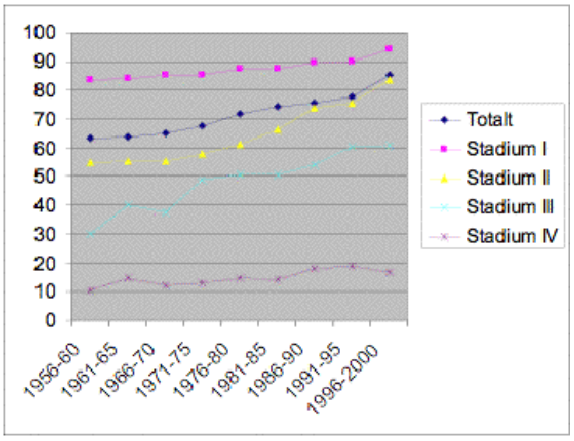


Figure 4. 5-year survival divided by stage of disease(7) in Norway.

2. Breast Anatomy and Development

2.1 Embryology

Rudimentary mammary glands develop in the embryo and grow after birth in a parallell manner to body growth. During the third trimester of pregnancy placental sex hormones enter the fetal circulation and induce canalization of the branched epithelial tissues. During early childhood, the end vesicles become further canalized and develop into ductal structures by additional growth and branching(8). The normal mammary glands of humans consist of a branching network of ducts terminating in end buds prior to puberty, thereafter in increasing numbers of alveolar lobules, and finally in secretory alveoli during pregnancy and lactation(8). These changes are under the influence of circulating hormones, especially estrogen. In experimental studies estrogen alone induces a pronounced ductular increase, whereas progesterone alone does not. The two hormones together produce full ductular-lobular-alveolar development of mammary tissues(9). Prolactin produces the alveoli that secrete milk products. Other circulating hormones like cortisol and insulin in addition to local trophic agents are involved and required in growth control(10) . The development of the ductal system in the normal breast is well studied, and briefly one can say that alveolar buds cluster around a terminal duct and form type I lobules. Transition to type II and Type III lobules gradually results from continued sprouting of new alveolar buds(8). Concerning tumor growth there are hormone-dependent tumors, which comprise 25-30% of all human carcinomas(9, 11). The remaining 70-75% are hormone-independent, meaning they no longer require circulating levels of mammotrophic hormones in order to grow. They also grow without the local trophic agents. Figure 5 shows a diagrammatic representation of the lobular structures of the human breast(8)

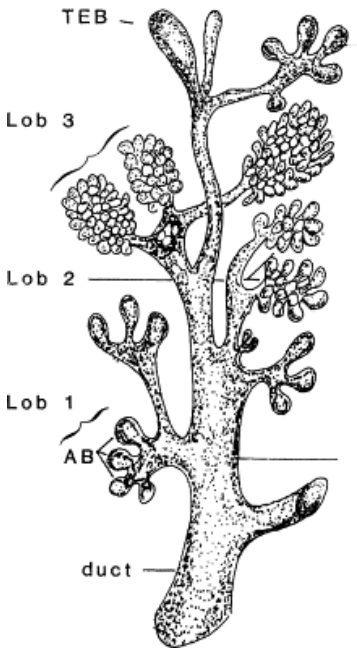


Figure 5. Diagrammatic representation of the lobular structures of the human breast(8)

2.2 Adult Breast

The normal mammary gland consists of two cellular compartments: the mesenchymal compartment of fatty stroma, which is permeated by blood vessels and nerves, and the epithelial compartment of ducts and lobules. These two compartments are separated by a basement membrane(12). During the development of malignant tumors, the natural bounderies and requirements for normal growth may be lost.

The cells of the epithelial compartment are as follows: 1) the epithelial cells which line ducts, 2) the alveolar epithelial cells which line alveoli and synthesize milk proteins, and 3) the myoepithelial cells which are placed between the two above mentioned cell layers and the

basement membrane(12). The epithelium is all stratified epithelium. Benign lesions of the breast may contain all these types of cells while malignant lesions contain only epithelial cells(13, 14).

Carcinomas of the breast are thought to originate from epithelial cells or progenitor epithelial cells of the terminal duct-lobular unit(15). Growing evidence indicates that stroma may play an important role in cancer initiation and progression(16-18). This is also shown by our scientist group in an ongoing study were we emphasize the difference in gene expression between normal breast epithelium of a breast cancer patient as opposed to normal breast epithelium of a healthy woman (unpublished).

3. Pathology of Breast cancer

3.1 Noninvasive Breast Cancer

There are two types of noninvasive breast cancers, lobular carcinoma in situ(LCIS) and ductal carcinoma in situ(DCIS). It is not considered a malignant lesion, rather a risk factor for development of breast cancer. The name is reflected upon histology. LCIS is recongnized by its conformity to the outline of the normal lobules, with expanded and filled acini. DCIS is a more heterogenous group devided into four broad categories: papillary, cribriform, solid and comedo. DCIS is recognized as discrete spaces surrounded by basement membrane, filled with malignant cells and usually with a recognizable basally located cell layer made up of presumably normal myoepithelial cells(19). It is important to notice that these premalignant lesions might coexist with invasive cancers, especially DCIS.

3.2 Invasive Breast Cancers

Invasive cancers are recognized by their lack of overall architecture, by the infiltration of cells hazardly into a variable amount of stroma, or by forming sheet of continuous and monotonous

cells without respect for form and function of a glandular organ(19). The two most frequent types of invasive breast cancer are lobular and ductal. Invasive lobular cancer tends to permeate the breast in a single file nature, which explains why it remains clinically occult, escaping detection on mammography and physical examination until it has become quite large. Ductal cancers tend to grow as a more coherent mass, which makes it easier to detect at an earlier stage. This is illustrated in figure 6. Invasive ductal carcinoma accounts for 50-70% of all invasive breast cancers(19). Invasive lobular carcinoma accounts for 10-15% of all breast cancers. In addition there is a mixed lobular and ductal form which is increasing in incidence. The remaining types are sometimes named according to the features they display, being tubular, mucoid or colloidal, medullary, cribriform, papillary, adenoid cystic or metaplastic(fig 7). The different histopathological groups have also been studied in a molecular aspect, with a special interest in the medullary carcinomas which has shown to often be of the basal-like subtype(20). Otherwise they are referred to as NOS, meaning nothing otherwise specified. Roughly the infiltrating ductal carcinomas have the worst prognosis. There are other malignant tumors primary to the breast, these being the phyllodes tumors and the angiosarcomas. The phyllodes tumors are composed of mixed connective tissue and epithelium. They can be benign or malign. The benign ones resemble fibroadenomas but with marked differences. The malignant ones are malignant stromal sarcomas which require wide exisional surgery similar to sarcomas on the trunk. Absolutely noteworthy is that these spread hematogenously, not through lymph nodes like all the other breast cancers. Therefore lymph node dissection is not required in such cases(19).

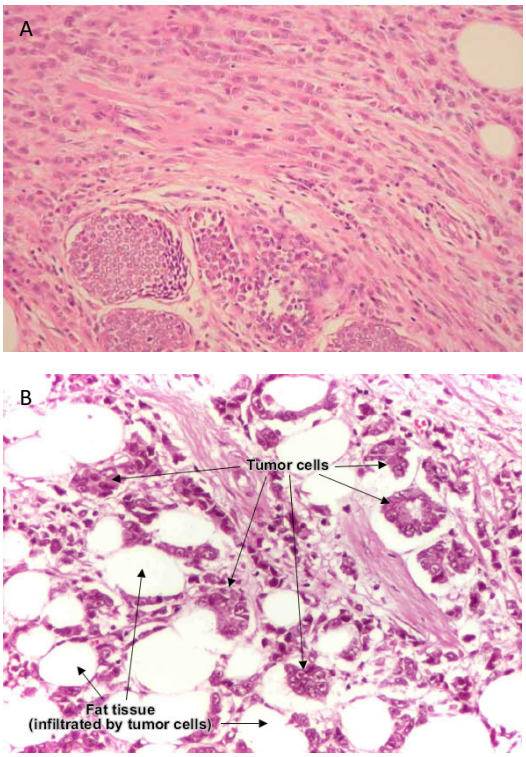


Figure 6 Histological appearance of the two major types of breast cancer(A)Typical streaming pattern of infiltrating lobular carcinoma. (B) Infiltrating ductal carcinoma with atypical tumor cells form ribbons, tubules or nests, broke the basement membrane of the duct and infiltrate the surrounding tissues

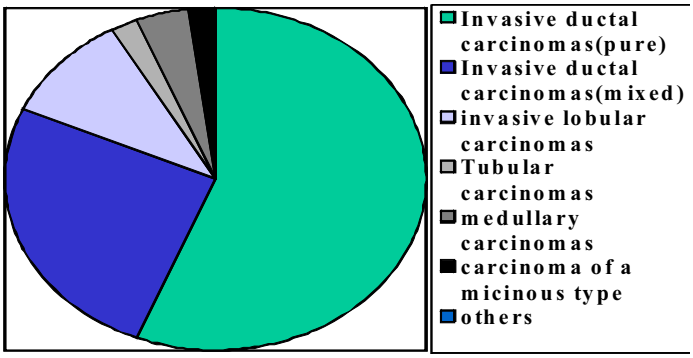


Figure 7 A more precise definition of the histopathological types of breast cancer

3.3 Staging of Breast Cancer

The most widely used system for staging primary breast cancer has evolved from classifiations proposed by the International Union Against Cancer (UICC) and the American Joint Committee on Cancer (AJCC). The system is based upon the description of the tumor (T), the status of the regional nodes (N), and the presence of distant metastases (M). Staging systems constantly change and reflect current trends in treatment, in changing outcomes after therapy, and new diagnostic technology(19). A detailed classification of breast cancer from the AJCC Cancer Staging Manual has been recently outlined (21).A more clinical categorization and of special importance to the surgeon is to separate the stages according to operability. This is illustrated in table 2(21).

Primary operable/curable breast cancer	
Stadium I	clinicalT1N0M0
Stadium II	clinicalT0-2N1M0
	clinicalT2N0M0

Primary inoperable/noncurable breast cancer	
Stadium II	clinicalT3N0M0
Stadium III	clinicalT0-2N2M0
	clinicalT3N1-2M0
	clinicalT4N0-2M0
	clinicalT0-4N3M0
Stadium IV	clinicalT1-4N0-3M1

Table 2 Staging of breast cancer separated by the operability of the tumor(21)

These histopathological characteristics in addition to prognostic markers like hormone receptor status, HER 2 and eventually Ki67 are robust well established prognostic and predictive markers. However there is a perception that the traditional factors presently available are insufficient to reflect the whole clinical and genetically heterogeneity of breast

cancer, and less than perfectly adapted to each patient particularly with the increasing range of treatment options. Recent advances in human genome research and high-throughput molecular technologies have begun to address and handle the molecular complexity of breast cancer and have contributed to the realization that the biological heterogeneity of the disease has implications for treatment.

4. **Breast Cancer Today and Guidelines for Treatment**

Breast cancer is the leading cause of cancer mortality in women(1, 22). The prognosis of breast cancer is dependent on stage at the diagnosis, tumors diagnosed at early stage having better prognosis. It is therefore important to detect breast tumors at as early stage as possible(23). Large meta-analysis have shown that recurrence is likely in 20-30% of young women with early-stage (lymph-node-negative) breast cancer who undergo only surgery and localized radiation therapy(24). Traditional approaches to breast cancer diagnosis and treatment rely upon histological appearance coupled with pathological parameters such as tumor grading, size, lymph node involvement and the presence of metastases (http://nbcg.no/nbcg.blaaboka.html). TNM (T=Tumor Size, N=Nodal status, M=Metastasis) is the classification system used by most pathologists, and for histological evaluation the Nottingham grading system (Nottingham modification of the Bloom-Richardson system)(25) is generally applied. TNM classification correlates well with survival(26). Lymph node status is known to be the strongest single prognostic indicator(27) but T was also strongly correlated with survival and in some instances it nearly balanced the effect of N status. Chemotherapy and/or hormonal therapy reduces the risk of distant metastases by approximately 30%(24, 28). This means 70-80% of patients receiving this treatment would have survived without it. Current predictive factors include hormone receptor status to predict response to Tamoxifen and HER2/neu receptor status to identify patients who will benefit from Herceptin. Can

molecular studies add to the clinical parameters of classification to enable us to identify breast cancers in early stages and to identify patients who will benefit from adjuvant therapy? Can personalized cancer medicine be enabled through analysis of gene expression patterns(29)? We are far from there yet, but this was the thought and driving force behind the following work.

4.1 Surgery

For early stages of breast cancer, surgical removal provides a reasonable chance for cure. The method of the surgical procedure has changed dramatically over the past century. We will shortly present the different methods(19).

Radical mastectomy – the breast and underlying pectoralis muscles are sacrificed and regional lymph nodes along the axillary vein to the costoclavicular ligament are removed.

Modern mastectomy – mastectomy refers to a complete removal of the mammary gland, including the nipple and the areola. Simple mastectomy is done in a plane just under the pectoral fascia and over the fibers of the muscles, and it is done by division of the gland from the axillary contents. Extension of the operation under the pectoralis major muscle and extending up to the axillary vein, removing the axillary lymph nodes, is called a modified radical mastectomy.

Breast conserving therapy (wide local exision and primary radiation therapy) – the removal of the tumor with a surrounding rim of grossly normal breast parenchyma. This requires postoperatively radiation treatment. The purpose of the radiation is to prevent local recurrence.

Sentinel node procedure Sentinel lymph node biopsy intraoperatively is performed to decide the extent of the axillary dissection. Most cases of sentinel node biopsy are performed for clinically uninvolved axilla. If the surgeon believes a palpable axillary node is only reactive and not metastatic, the sentinel node procedure can still be performed as long as the surgeon maintains a low threashold for default to conventional axillary lymph node dissection(30). Axillary dissection is done through a separate incision in most patients. Sentinal node is performed by injecting a radioactive substance into and around tumor before surgery. This is done in the nuclear medicine department. The radioactivity will be taken up by the lymphoid system and sentinel lymph node will reveal if the cancer has spread. Some surgeons also use a blue dye in addition to the radioactive colloid. This is to be able to visualize the sentinel node. If there are caner cells in the sentinel node, complete axillary dissection is most often needed. If not the operation can be terminated and the axilla left intact. Axillary node dissection gives good regional control in node negative patients following an adequate axillary dissection(31). The number of axillary nodes removed in node negative breast cancer predicts for regional recurrence, with less than six nodes removed associated with significantly higher regional nodal recurrence(31, 32). Sentinel-node biopsy is a safe and accurate method of screening the axillary nodes for metastasis in women with a small breast cancer(33).

There have been several studies comparing mastectomy with breast conserving therapy(34-36). They all conclude that breast conserving therapy has the same results considering survival rates as mastectomy provided that the surgery is thorough and with required margins.

4.2 Radiational Therapy

Radiational therapy can be given either pre- or postoperatively, or as the only local treatment. It is primarily used to get regional and local control over the disease. There are strict

guidelines for the use of radiational therapy in breast cancer, where age, T-status and N-status are considered. Intraoperative radiation therapy in breast cancer has been done and evaluated(37), but has not yet become standard procedure.

4.3 Adjuvant Systemic Treatment

Adjuvant systemic treatment is given to the patients to reduce the risk of recurrence. Prognostic indicators like axillary involvement and special tumor characteristica indicate increased risk for recurrence and should be considered and treated. Within the concept of adjuvant therapy both hormonal treatment and chemotherapy is considered. There is a detailed flow-chart used by Norwegian surgeons and oncologists which is given out by the Norwegian Breast Cancer Register. It includes age, TNM status, grade and hormonal status. These guidelines are similar to those used internationally.

5. Material and Method

5.1 Material

Women are submitted to our hospital both due to benign lesions, malignant lesions and precancerous lesions. This allows us to make a biobank with tissue specimens from women with a large variety of clinical presentation and different risks for breast cancer. The first group is the women with diagnosed breast cancer. This group has taken a preoperative biopsy, in addition to samples taken during surgery. At surgery we wanted to take a sample both from the tumor and from histologically normal tissue adjacent to the tumor. Some of these women did not have invasive lesions but precancerous lesions referred to as in situ lesions (DCIS or LCIS). The next group of women are benign lesions which are biopsied at the radiology

department. Both of these two groups are from women tending to Akershus University Hospital. The third group of women are healthy women having had reduction mammoplasties. These samples wer taken at the Collosseum Clinic, a private clinic in Oslo doing plastic surgery. All of the women have signed a written consensus to participate in our study.

In our first paper where we look at the gene expression between pre and postoperative tumors we have 13 pairs. All of these patients had invasive breast cancer and had been operated with ablatio mamma (surgical removal of the entire breast). Sentinel node diagnostics using 99m Tc-labelled colloids were performed in eleven patients as a part of the surgical procedure(38), while the remaining two underwent direct complete axillary dissection. No recurrence of disease has been observed so far for these patients, but the follow-up time is short.

In our second paper there are 46 patients in the discovery cohorts and these are all women operated at Akershus University Hospital. The validation cohort is a combination of different cohorts given the total amount of 947 breast tumor samples.

In the third paper a total of 79 tumors, of which 71 malignant tumors of the breast, 6 fibroadenomas, and 2 DCIS were studied for the expression of Bmi1 and Mel18 and compared to the reduction mammoplastic specimens of 12 healthy women. In addition there was available adjacent cancer free tissue for 23 of the malignant tumors.

5.2 Methods

5.2.1 Tissue Collection

Preoperative needle biopsies were obtained by an experienced radiologist using a 16 Gauge core needle device through a small skin incision in a sterile field. Three samples were processed for routine histological diagnosis while one sample for molecular analysis was put directly into RNAlater (Sigma Aldrich, St Louis, MO). The surgical samples were taken by the breast cancer surgeon upon removal of the breast, and preserved in RNAlater. Tissue samples from malignant tumors and normal counterpart were obtained in cases of breast cancer were it was possible. The RNAlater-stabilised tissue samples were stored at −80ºC. Tissue samples from non cancer controls (reduction mammoplasty) have been gathered at the Collosseum Clinic in Oslo and stabilized in the same way as the rest of the samples.

5.2.2 RNA Isolation

RNA was prepared using the method of Wei and Khan(39). Briefly, frozen tissue samples were homogenized in TriReagent (Ambion, Austin, TX) using a 5 mm steal bead in a Mixer Mill MM301 (Retsch, Haan, Germany) at 30 Hz for 2 min. After phase-separation with 0.2 vol chloroform, the aqueous phase (containing RNA) was mixed with 1.5 vol 100% ethanol and transferred to RNeasy Mini columns (Qiagen, Hilden, Germany). Further processing (including one-column DNase digestion) was per the manufacturer's protocol and the purified RNA was eluted with RNase-free water. RNA concentration was measured using NanoDrop ND-1000 UV-VIS Spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and the RNA quality analyzed on an 2100 Bioanalyzer (Agilent, Santa Clara, CA). The purified RNA was stored at −80ºC.

5.2.3 Microarray Analysis

mRNA amplification, labelling, and hybridization were done following the manufacturer's instructions. We have used both one-color and two color techniques. For details one refers to the individual papers. After hybridization for 17 hours the arrays were washed and scanned using an Agilent scanner and microarray data extracted with Agilent Feature Extraction software.

5.2.4. Statistical Analysis

Preprocessing of the microarray data was done in J-Express Pro (www.molmine.org) while between-array quantile normalization was done in BioConductor(40). The microarray data are submitted to The ArrayExpress Archive (http://www.ebi.ac.uk/microarray-as/ae/), accession number is given in the individual papers

Further advanced statistical analysis was performed on the microarray expression data depending on the specific hypothesis and the aim of the different individual papers.

Microarrays can measure the expression of thousands of genes to identify changes in expression between different biological states. A number of methods were developed to determine the significance of these changes while accounting for the large number of genes. One of the first and most widely used methods was described by Tusher et al(41) as Significance Analysis of Microarrays (SAM), that assignes a score to each gene on the basis of change in gene expression relative to the standard deviation of repeated measurements. For genes with scores greater than an adjustable threshold, SAM uses permutations of the repeated measurements to estimate the percentage of genes identified by chance, the false discovery rate (FDR). This statistical equipment was used in all of the papers in this thesis. To adjust for parameters known to have prognostic value, like lymph node status, grade, estrogen

and progesterone receptor status, and breast cancer subtype, a partial least squares regression analysis was performed with the pls package in R(42, 43). Between group comparission were done using Student`s two-sided, two-class t-test and ANOVA using the function aov in R..

Gene function classification of the significant genes were performed using DAVID(44, 45) and pathway analysis was done through the use of Ingenuity Pathway Analysis (IPA; Ingenuity® Systems, Redwood City, CA).

Survival analysis was a major topic of paper number two. The signatures were evaluated for prediction of Distant Metastasis Free Survival (DMFS). A total of 912 patients on the validation set (n = 947) had available DMFS status with median follow-up for 81 months. The Kaplan-Meier survival curves were plotted for the corresponding risk groups. The differences in survival probabilities associated with the risk groups were tested by a logrank test.

A *likelihood ratio test* was used to assess the significance of the overall effect in a univariate comparison of predictors. *Deviance* was used to check the goodness of the model fit. The marginal contribution by a single predictor in the univariate setting was evaluated using the *proportion of variation explained* in the outcome variable (PVE) (46), which is an indicator for the importance of covariates in the Cox model. The *Hazard Ratio* (HR) was used as an accuracy measure for the risk group prediction for different predictors. The *concordance index* (C-index) (47) was computed to assess the predictive discrimination ability of each of the predictors in the corresponding univariate Cox model. For a multivariate comparison of predictors, the relative importance of a covariate in a multivariate Cox model was measured by the partial PVE.

One goal of microarray studies is to find a subset of genes that are differentially expressed across experimental conditions. The key question is the strength of the claim that a given gene is differentially expressed. In other words, what is the power of the test to determine if a gene is differentially expressed to a specific degree. In our study we have calculated power based on the SAM results. Power=1-q (q is the highest FDR-value)

5.2.5 Quantitative RT-PCR

To confirm the results of the microarray experiment, qRT-PCR using TaqMan low density arrays (Applied Biosystems, Carlsbad, CA) were performed. Primer-probe of selected genes were used. 500 ng RNA was reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) per the manufacturer's instructions. Due to lack of material qRT-PCR was only performed for ten of the sample pairs. The samples were further processed using TaqMan® Gene Expression Master Mix (Applied Biosystems) and run on the 7900HT Real-Time PCR System (Applied Biosystems) as per the manufacturer's instructions. Relative changes in gene expression were analyzed using the ΔΔCt-method(48).

5.2.6 Histology

The slides were evaluated by an experienced pathologist (AJN) and graded according to the Nottingham grading system (Nottingham modification of the Bloom-Richardson system)(25)

5.2.7 Immunhistochemistry

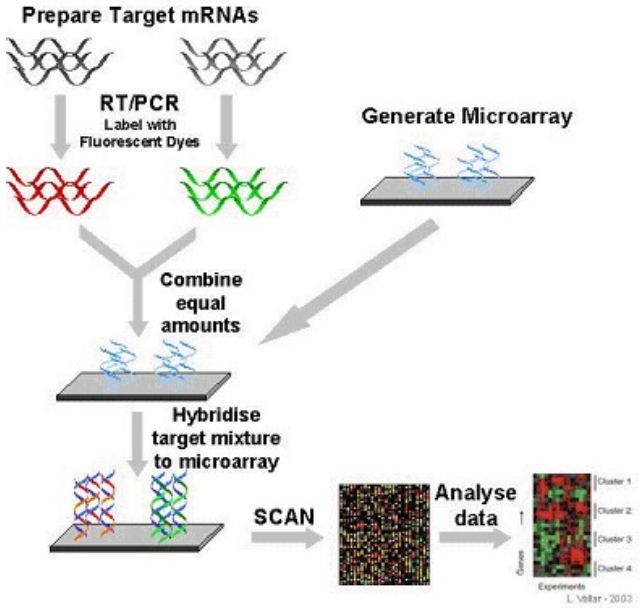
Immunohistochemistry was performed on 5 µm sections from formalin-fixed, paraffin-embedded tissue applied to coated slides. Deparaffinization, rehydration and epitope retrieval

were performed in a Dako PT link (Dako) at 97°C for 20minutes. Dako Autostainer Plus together with Envision™ Flex, high pH system (K8000, Dako) were used in the immunostaining procedure following the operating manual.

5.3 The Technology of Microarray

DNA microarray analysis is a fast and versatile approach to achieve high throughput explorations of genome structure, gene expression program, and of gene function at both cellular and organism levels. A DNA microarray-based analysis is a complex multistep process involving numerous specific equipments, and requiring a strong expertise in various areas, including molecular biology, image analysis, computing, and statistics.

The most common application of DNA/oligonucleotide microarray is gene expression analysis. In this technique, RNA isolated from two samples is labeled with two different fluorochromes (generally the green cyanine 3 and the red cyanine 5 (Cy3, Cy5)) before being hybridized to a microarray consisting of large number of cDNAs/oligonucleotides orderly arranged onto a glass microscope slide. After hybridisation under stringent conditions, a scanner records, after excitation of the two fluorochromes at given wavelengths, the intensity of the fluorescence emission signals that is proportional to transcript levels in the biological sample. The microarray data are analysed using specific softwares that enables clustering of genes with similar expression patterns, assuming that they share common biological functions. This is illustrated in figure 8.



www.microarray.lu/en/MICROARRAY_Overview.shtml

Figure 8 The concept of microarray

5.4 Objectives of Microarray Studies

1. Class Discovery

Class discovery involves the grouping together of specimens based on their expression profiles across the set of genes represented on the array or the grouping together of genes with regard to their expression profiles across the samples assayed(49). The classical example is how Perou et al(50) divided breast cancers into several subgroups using clustering analysis. Cluster analysis alogarythms are called unsupervised because the grouping is not driven by any phenotype external to the expression profile. A commonly used hierarchical clustering method starts by defining a distance between two breast tumors as a function of the difference in gene expression. One then regroups the two closest tumors and preceeds by regrouping tumors to obtain a cluster tree, which can be split into branches by selecting a cutoff distance.

This method of hierarchical alogarythmical clustering is described by Eisen et al in 1998(51). The application of microarray is illustrated in figure 9.

2. Class Comparison

Class comparison focuses on determining which genes are differently expressed among samples represented of predefined classes(49). In cancer studies, the classes often represent different categories of tumors, differing with regard to stage, size, grade, prognosis, response to therapy and similar conditions. An example of this is how van`t Veer et al in 2002(52) compared tumors from 34 breast cancer patients who developed a distant metastasis within 5 year after surgery and tumors from another group of 44 patients who did not develop metastases. A statistics measuring the difference in gene expression between the two types of tumors is applied. Genes are then ranked according to this statistic, starting with the most differentially expressed. They selected the 70 genes with the highest correlation with the distant metastasis status at 5 years. In this case they used supervised classification, meaning there were predefined classes. Class comparison was also used by Sotiriou et al(53) to evaluate genes whose expression was correlated specifically with grade. They studied 99 tumours from node-positive and node-negative, the majority receiving adjuvant treatment according to accepted practise guidelines at the time of diagnosis. They performed the comparison of the different clinicopathological characteristics, tumor grade 1 or 2 vs grade 3; tumor size $\leq$ 2 cm vs > 2 cm; age <50 years vs age $\geq$ 50 years; node negative vs node positive; and ER- vs ER+. Their results were concordant with earlier studies despite the difference in patient populations, treatment used, and technology platforms used. They found that ER status of the tumor was, indeed, the most important discriminator of expression sybtypes and that tumor grade was a distant second.

3. Class Prediction

In class prediction the emphasis is on developing a mathematical function that can predict which class a new specimen belongs to based on its expression profile(49). This is important for problems of diagnostic classification, prognostic prediction and treatment selection. Again van`t Veer et al(52) determined a prediction rule for prognosis based on the expression of the 70 genes from data on 78 node negative breast cancer patients and then evaluated on another 19 patients. Another example is Ma et al(54) who developed such a predictor for patients with estrogen receptor positive primary breast cancer who received tamoxifen monotherapy after local therapy. Here, as well as for class comparison, supervised analysis is used.

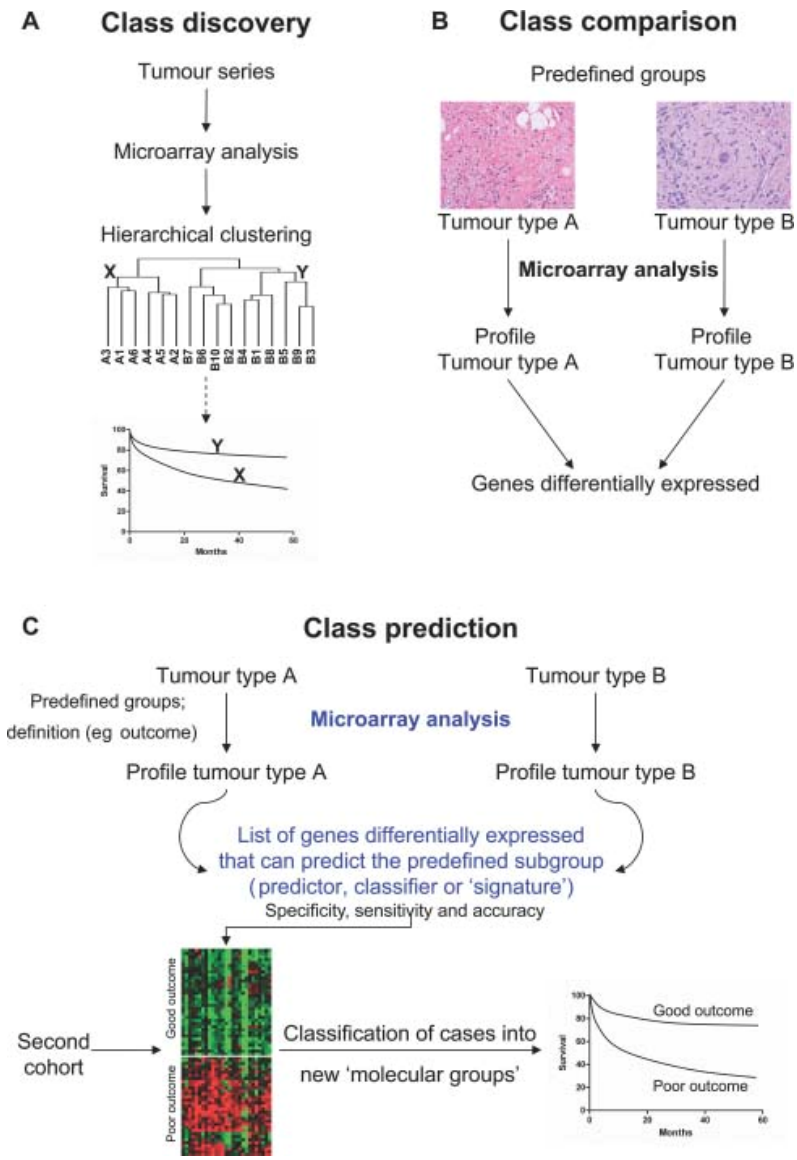


Figure 9 Work flow in microarray technology

The concept of microarray and the contribution of gene expression profiling to breast cancer classification, prognostication and prediction was summarized by Weigelt et al in 2010(55).

There has been skepticism to this concept, especially by clinicians. The strongest argument against using gene expression in the clinics is that gene expression signatures are highly unstable(56), that is the gene lists obtained for the same clinical types of patients by different groups differed widely and had only very few genes in common. This is exemplified by the prognostic gene expression signature of breast cancer using expression data of tumor samples of known clinical outcome (52). This study yielded the 70-gene signature for breast cancer prognosis. When another research group independently used a similar approach, a 76-gene signature was identified(57). These two signatures surprisingly had only three genes in common, hence some claimed the signatures to be unstable. Both these signatures have independently been validated in large group of patients (58, 59) with similar outcome despite the small number of overlapping genes. The concordance among gene-expression-based predictors for breast cancer has later been tested and verified(60). Even though different gene sets were used for prognostication in patients with breast cancer, four of the five tested showed significant agreement in the outcome predictions for individual patients. It is more likely that different signatures use different genes to monitor the same biological process. By combining information from multiple gene signatures, one would potentially increase the prediction power and bring out an overall picture of this disease. Zhao et al aimed to develop an analytical framework that allows us to utilize the combined strength from individual gene signatures(61). Such a framework and the resulting model will be broadly applicable for survival prediction across heterogeneous tumor groups capturing a broad spectrum of biological aspects. At the same time it is argued that combining different signatures may reach a limit without improving prognostication further(62).

6 Stem Cells and Important Phenomenon in Cancer Development – this Chapter is Related to Our Third Paper where the Genes Bmi1 and Mel18 are specifically Studied

6.1 Stem Cells

Stem cells have a long life and a large replicative potential, making them good candidates for the cells of origin of cancer. They are defined as cells with the capacity to self-renew and to generate daughter cells that can differentiate down several cell lineages to form all cell types that are found in the mature tissue(13). The field of stem cell biology has emerged initially from hematology (63, 64). Hematopoeietic stem cells have already been used in therapeutic settings being a vital element in bone marrow transplantation. Stem cells go through assymetrical cell divisions giving rise to an identical cell with the same properties as well as to another well defined differentiated cell. The cells that form the intermediates between stem cells and terminally differentiated cells are usually referred to as progenitor cells(65, 66). Stem cells can be divided into two functional classes. The first class represents stem cells that are responsible for tissue renewal, like in the bone marrow. They are responsible for replacing terminally differentiated cells as they mature or die or are shed from an epithelial surface. The second class comprises stem cells that are inactive until required in response to environmental factors, for instance to repair tissue damage(65).

In the adult mammary gland, no definitive identification has been made of an adult mammary epithelial stem cell. Some cells have been attributed stem cell activity which in the beginning of life was considered to be located in the terminal end buds and give rise to the different cells of the mammary gland(67). AS described in detail above the mammary gland in humans and other mammals is a dynamic organ that undergoes significant developmental changes during

pregnancy, lactation, and involution. It is considered that only stem cells have the replicative potential that would be needed to maintain this process. If such stem cells do exist they would be in the second group described above meaning they are quiescent until responding to physical cues(65). Another sign of mammary stem cell existence is the replacement of cells that are shed from the epithelium into the lumen during routine cell turnover. These cells would then fall into the first group described above being continually active stem cells(65). The existence of mammary stem cells is being discussed but direct evidence of specialized mammary stem cells has come from several studies(68, 69).

A number of investigators suggest that stem cells may represent important targets for transformational events(70). Adult stem cells are slow-dividing, long-lived cells that by their very nature are exposed to damaging agents for long periods of time. Therefore they may accumulate mutations that can result in their transformation. Normal stem cells and cancers share a number of important phenotypic properties which is very important in carcinogenesis. These are: 1) capacity for self-renewal; 2)ability to differentiate; 3) active telomerase and anti-apoptotic pathways; 4)increased membrane transporter activity; and 5) anchorage independence and ability to migrate(66)). There is now evidence for the existence of "tumor stem cell" in human leukemias(71), myelomas(72), brain tumors(73) , as well as in breast tumors(74). Using a model in which human breast cancer cells were grown in immunocompromised mice, Al-Hajj and collegues found that only a minority of breast cancer cells had the ability to form new tumors. They were able to distinguish the tumor initiating cells from those not being able to produce tumor based on cell surface marker expression (CD44+,CD24-Lineage-). These cells did not just have the ability to self-renew but were also able to give rise to cells with different characteristics or phenotypes that made up the bulk of the tumor. These observations support the concept of these cells being breast cancer stem

cells. A major factor that protects women from breast cancer is an early first full-term pregnancy. The mechanism behind this concept is that there is a burst of differentiation during pregnancy and therefore a depletion in stem cells. This means there are fewer stem cells with the potential of becoming breast cancer stem cells over a life period(75).

As we have seen there is great resemblance between normal stem cells and cancer. Therefore the pathways in stem cell biology is of great importance.

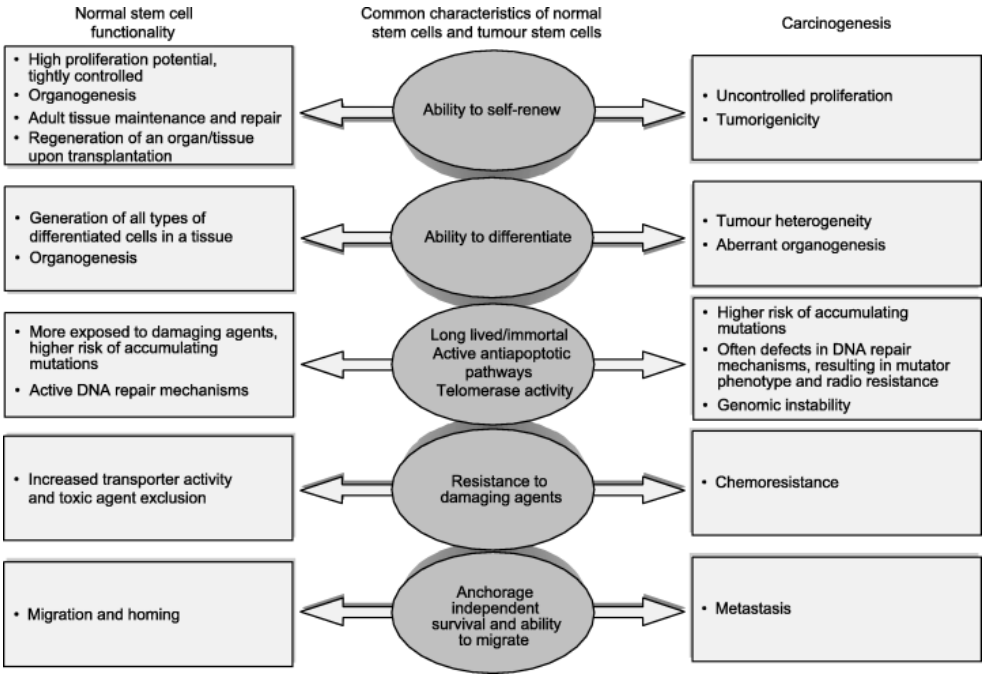


Figure 10. Similarities between normal stem cells and cancer cells and their impact on stem-cell functionality and carcinogenesis(66).

6.2 Self-renewal

As we have seen stem cells are long-lived targets for occurrence of mutations compared with differented cells that die within a short period of time. The accumulation of mutations neccessary for carcinogenensis is more likely to occur in stem cells that self-renew over the lifetime of an individual rather than in the mature cells that exit the cell cycle and/or undergo apoptosis after a brief period. This supports the new thinking of how tumors are likely to arise from stem cells(75).

As the name indicates the property of self-renewal is to give rise to a daughter cell that has the exact same qualities as the mother cell. Stem cells must keep a fine balance between self-renewal and differentiation, because when unregulated it could lead to cancer. The relation between stem cells and tumorigenesis is therefore close, and it is of great importance to understand the mechanism and pathways involved in self-renewal. Three important pathways are well studied, these being the Wnt signaling pathway(76, 77), the Hedgehog signaling pathway(77, 78), and the Notch signaling pathway(77) . As stage dependent activation of these pathways as well as their key members were among the main findings of this thesis, these will be described shortly.

6.3 The Wnt Signaling Pathway

Wnt proteins are secreted intercellular signaling molecules that act as a ligand to trigger a specific signal transduction pathway. There are at least 16 members of this ligand family in vertebrates(79). Briefly described, in the abscence of Wnt, there will be degradation of beta-catenin giving the Tcf/LEF complex(T-cell factor/lymphoid enhancing factor) the possibility

to associate with the transcriptional repressor Groucho(75, 80). The process includes the following proteins: the degradation complex consisting of axin, adenomatous polyposis coli(APC), glycogen synthase kinase(GSK3beta) and casein kinase I(CKI). Upon binding of Wnt to its receptor called Frizzled and co-receptor LRP(low-density lipoprotein receptor related protein), through the same cascade as above, beta-catenin will not be degraded and move into the nucleus where it can act as a co-activator of the Tcf/LEF family of transcription factors to regulate specific target genes. To summarize Wnt allows transcription of Wnt target genes through the stabilized form of beta-catenin (figure 11). Example of target genes are c-myc and cyclin D(75, 80) . There has been several studies on the role of Wnt signaling in self-renewal in different organs. By the use of transgenic mice it was shown that overexpression of Wnt ligands in mammary stroma or activated beta-cataenin in mammary epithelium leads to increased numbers of mammary stem cells(81) . Studies linking this process to mammary carcinogenensis include those showing that mammary stem cells and progenitors might be targets for oncogenesis by Wnt 1 signaling elements(82). Liu et al has drawn a schematic diagram of the Wnt sinaling pathway which is shown in figure 11(77). Here we have only briefly described the canonical Wnt/beta-catenin pathway.

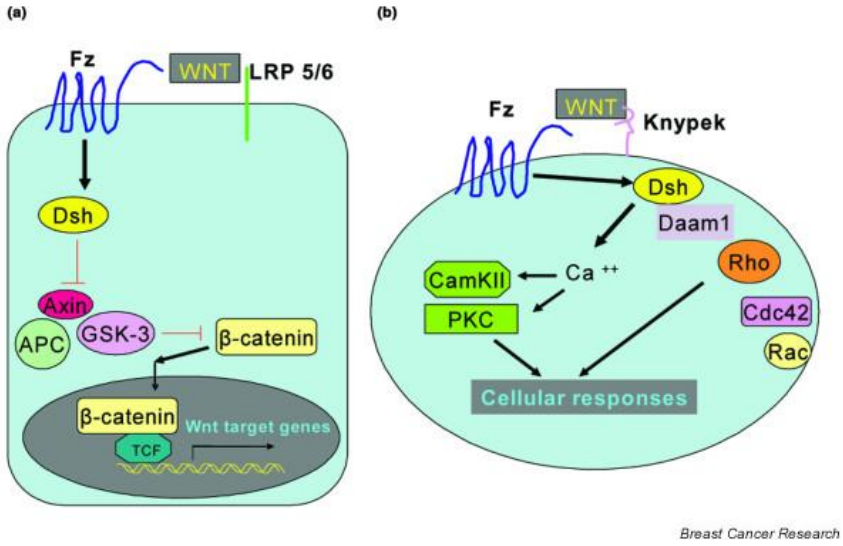


Figure 11(77) A schematic diagram for the up-stream regulation of the Wnt signaling pathway. (a) The canonical Wnt/β-catenin pathway. Canonical Wnt signaling requires the Frizzled (Fz) and low-density lipoprotein receptor-related proteins 5 and 6 (LRP5/6) co-receptors to activate Dishevelled (Dsh). Then Dsh inhibits the activity of the β-catenin destruction complex (adenomatous polyposis coli (APC), axin, and glycogen synthase kinase-3 (GSK-3)), which phosphorylates β-catenin in the absence of the ligands. β-Catenin is stabilized and translocated to the nucleus, where it recruits transactivators to high mobility group (HMG)-box DNA-binding proteins of the lymphoid enhancer factor/T cell factor (LEF/TCF) family. (b) The noncanonical Wnt signaling pathway. Noncanonical Wnt signaling requires Frizzled receptors and the proteoglycan co-receptor Knypek. In this pathway, Dsh localizes to the cell membrane through its DEP domain. A main branch downstream of Dsh involves the small GTPases of the Rho family. Dsh activation of Rho requires the bridging molecule Daam1. Dsh can also stimulate calcium flux and the activation of the calcium-sensitive kinases protein kinase C (PKC) and calmodulin-dependent protein kinase II (CanKII). At the end, the activation of this pathway induces the complex and dynamic cellular response.

6.4 The Hedgehog Signaling Pathway

Hedgehog proteins are intercellular signaling molecules, similar to Wnt proteins. There are three members of these proteins, Sonic Hedgehog(Shh), Desert Hedgehog(Dhh), and Indian Hedgehog(Ihh). These ligands bind to the hedgehog-interacting protein 1(Hip1) and Patched(Ptch) which are transmembrane receptors of these ligands. Patched and Smoothened(Smo) are two transmembrane proteins that form a receptor complex in the absence of ligands. Upon binding of Hh to Ptch inhibition of Smo is relieved . This causes a large protein complex to dissociate and release the transcription factor Gli. Gli is translocated into the nucleus where it regulates the expression of its target genes(75, 77). Target genes include those controlling cell proliferation like cyclin D, cyclin E, c-myc, components of the epidermal growth factor pathway, and angiogenesis components including platelet-derived growth factor and vascular endothelial growth factor. The protein complex which inhibits the action of Gli is not well defined, but it includes Fused(Fu) and Suppressor of Fused(SuFu)(83). Figure 12 shows a schematic diagram of the Hedgehog signaling pathway(77).

Liu et al in 2006(78) studied the role of hedgehog signaling and Bmi-1 in regulating self-renewal of normal and malignant human mammary stem cells. They showed that hedgehog signaling components Ptch1, Gli1, and Gli2 are highly expressed in normal human mammary stem/progenitor cells cultured as mammospheres and that these genes are down-regulated when cells are induced to differentiate. They also proposed that these effects are mediated by Bmi-1, studied in this thesis.

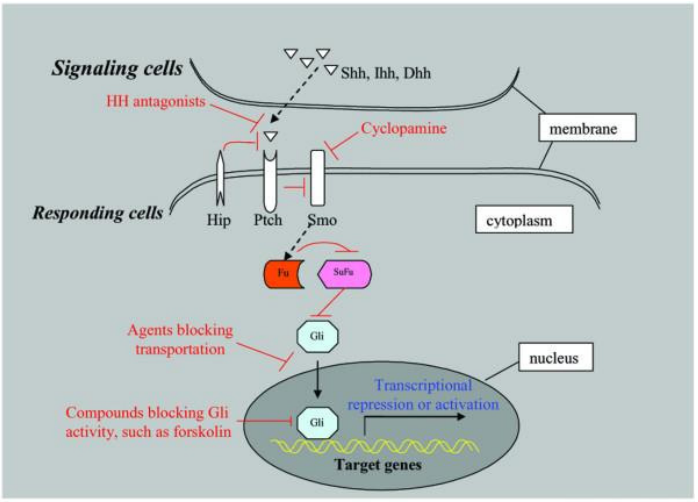


Figure 12(77) A schematic diagram for the hedgehog (HH) signaling pathway. Ligands, such as Sonic Hedgehog (Shh), Indian Hedgehog (Ihh), and Desert Hedgehog (Dhh), are secreted by signaling cells and bind the transmembrane receptor patched (Ptch) in HH responding cells. In the absence of ligands, Ptch binds to Smoothened (Smo) and blocks Smo's function, whereas this inhibition is relieved in the presence of ligands, and Smo initiates a signaling cascade that results in the release of transcription factors Glis from cytoplasmic proteins fused (Fu) and suppressor of fused (SuFu). In the inactive situation, SuFu prevents Glis from translocating to the nucleus; in the active situation, Fu inhibits SuFu and Glis are released. Gli proteins translocate into the nucleus and control target gene transcription. The red lines and the agents in red show the inhibitors of this pathway with potential therapeutic value.

6.5 Notch Signaling

The Notch proteins, of which there are four members in mammals, Notch 1 to Notch 4, are expressed in a variety of stem or early progenitor cells(77, 84). They interact with DSL ligands which are surface bound ligands. This in turn triggers a cascade which eventually

leads to Notch translocating into the nucleus where it acts upon its targets. Liu et al has also made a schematic diagram of this pathway which is shown in figure 13.

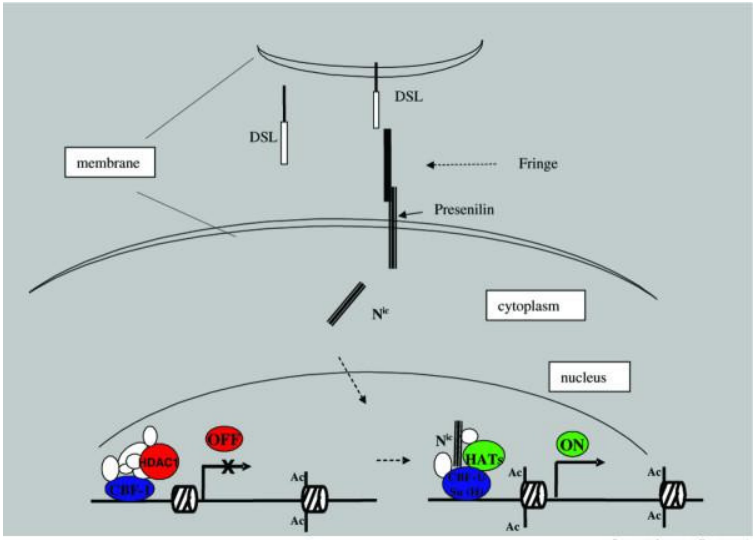


Figure 13(77) A schematic diagram for the Notch signaling pathway. Upon binding of the DSL ligand, Notch signaling is modulated by fringe, and Notch receptors are activated by serial cleavage events involving members of the ADAM (for 'a disintegrin and metalloproteinase') protease family, as well as an intramembrane cleavage regulated by γ-secretase (presenilin). This intramembrane cleavage is followed by translocation of the intracellular domain on Notch to the nucleus, where it acts on downstream targets. CBF, C promoter binding factor; HDAC, histone deacetylase; HAT, histone acetyltransferase.

6.6 Interaction between Self-renewal Pathways

Several studies have been conducted to follow the interaction between the different pathways of self-renewal. Liu et al has drawn a hypothetical interacting model after reviewing these studies(77). Figure 14 is taken from this review article.

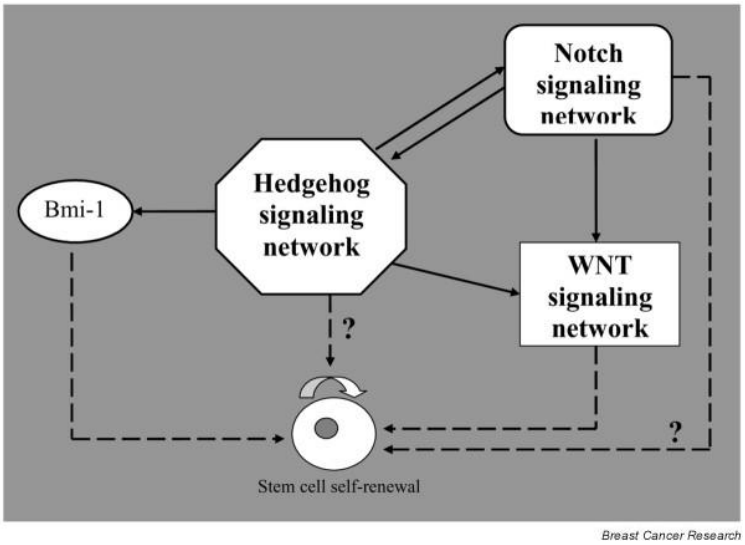


Figure 14(77) A hypothetic interacting model in the regulation of stem cell self-renewal by the Hedgehog signaling pathway, the Notch signaling pathway, the Wnt signaling pathway, and B lymphoma Mo-MLV insertion region 1 (Bmi-1). Interactions between the Hedgehog, Notch, and Wnt pathways and Bmi-1 are shown by solid arrows; interactions between stem cell self-renewal regulation by the pathways and Bmi-1 are shown by dashed arrows; the question marks represent the postulated interactions.

6.7 Bmi-1 and Self-renewal

This gene and its protein product is central in paper III. Here we will just briefly describe its role in self-renewal. Recently, Bmi-1 has been shown to be a key regulator of the self-renewal of both normal and leukemic stem cells(85, 86). Bmi-1 has also shown to be important in neuronal stem cell self-renewal(87) . The exact mechanisms by which Bmi-1 regulates stem cell self-renewal is unclear, but it has been suggested that it is by silencing of the P-16 gene(77, 86, 88) . We have studied the expression of Bmi1 both in genes and in proteins in different sources of breast tissue. This is evaluated in paper III(89). We found that there was a significant difference in the expression of Bmi1 in histologically normal breast tissue of

women operated for breast cancer as compared normal tissue in non cancer controls. This opens the possibility to study the expression of Bmi1 in breast cancer patients for prognostic purpose, and thereby eventually be able to offer more personized guided therapy, either surgically or medically.

7 Papers with Individual Main conclusions

7.1 Paper I

Gene expression studies on breast cancer have generally been performed on tissue obtained at the time of surgery. In this study we have compared the gene expression profiles in preoperative tissue (core needle biopsies) to postoperative tissue from the same tumor obtained during surgery. Thirteen patients were included of which eleven had undergone sentinel node diagnosis procedure before operation. Gene expression and miRNA microarray analysis were performed using total RNA from all the samples. Paired significance analysis of microarrays revealed 228 differentially expressed genes, including several early response stress-related genes such as members of the *fos* and *jun* families as well as genes of which the expression has previously been associated with cancer. The expression profiles found in the analyses of breast cancer tissue must be evaluated with caution. Different profiles may simply be result of differences in the surgical trauma and timing of when samples are taken, and not necessarily associated with tumor biology. A clinical analysis by Parvaiz et al compared image guided core biopsy with free hand biopsy for symptomatic breast lesions. 400 core biopsies were studied, and the found that 76 % of biopsies performed free hand had no significant difference in diagnostic sensitivity compared with image guided biopsies. In a

selected group, free hand biopsies provide the added advantage of early diagnosis and subsequent treatment.(90).

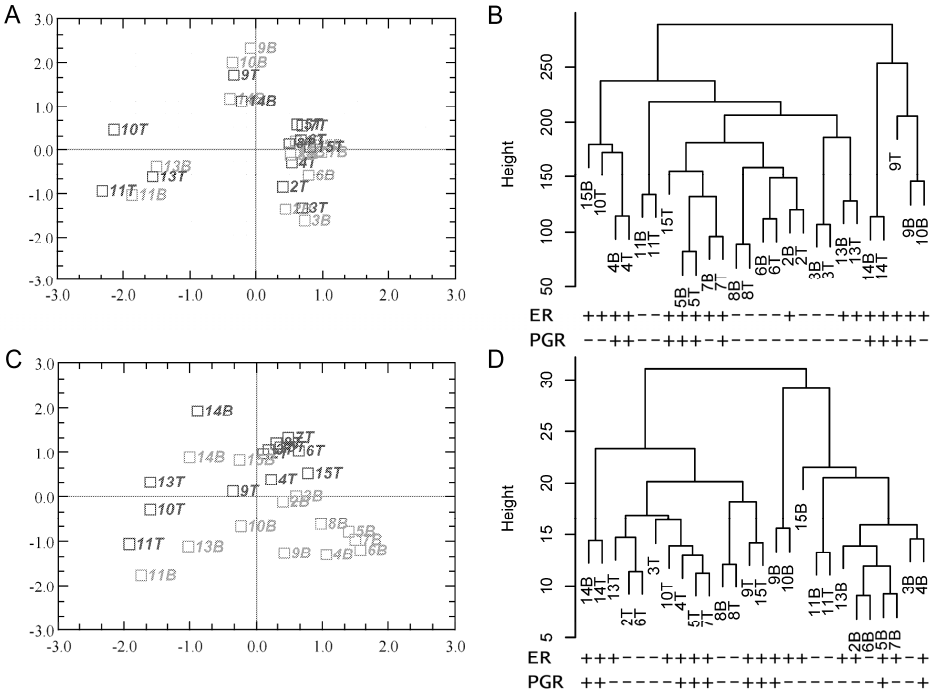


Figure 1 Unsupervised (A) and supervised (C) CA plots of mRNA expression in pre- and postoperative breast cancer tumours, and their corresponding clustering maps (B and D). The samples are marked by their respective numbers followed by either B, which define the preoperative samples, or T, which define the postoperative samples. The unsupervised chart was made using all 18,189 genes (24,105 probes) expressed on the microarrays whereas for the supervised only the 228 genes (235 probes) with a FDR < 2.5% from paired SAM where used.

7.2 Paper II

Due to early diagnostics through mammographic screening breast cancers today are of predominantly T1 (0.1–2.0 cm) or T2 (2–5 cm) categories. Nevertheless even in early stage breast cancer without lymph node involvement, 30% of the patients get recurrence of their disease. Chemotherapy and hormonal therapy reduce the risk of distant metastases by approximately 30%. However, the majority of the patients would have survived without adjuvant treatment and improved selection criteria could save patients from possible severe side effects and the community from unnecessary health expenses. The purpose of this study was to search for molecular signatures characterizing clinical T1 and T2 tumors and to better identify patients of high risk of relapse and/or metastasis. The gene list was validated on a collection of different cohorts from different groups. This allowed us to perform Kaplan-Meyer analysis and evalate how our newly developed signature was able to evaluate relapse-free survival as opposed to size. We observed highly significant separation of two risk groups for DMFS in the patient group with pT1 size tumors; suggesting that this signature can identify tumors that although categorized as T1 have a transcriptional signature that can point to poorer survival. In the pT2 tumor subgroup (n = 459), T-size signature achieved also significant separation for the risk prediction, suggesting that this signature can identify patients with good prognosis among those categorised as stage II. We found that our molecular signature was better than size in predicting prognosis of disease. This together with the specific genes involved can be used to select patients for targeted therapy and spare others for unnecessary potential toxic therapy.

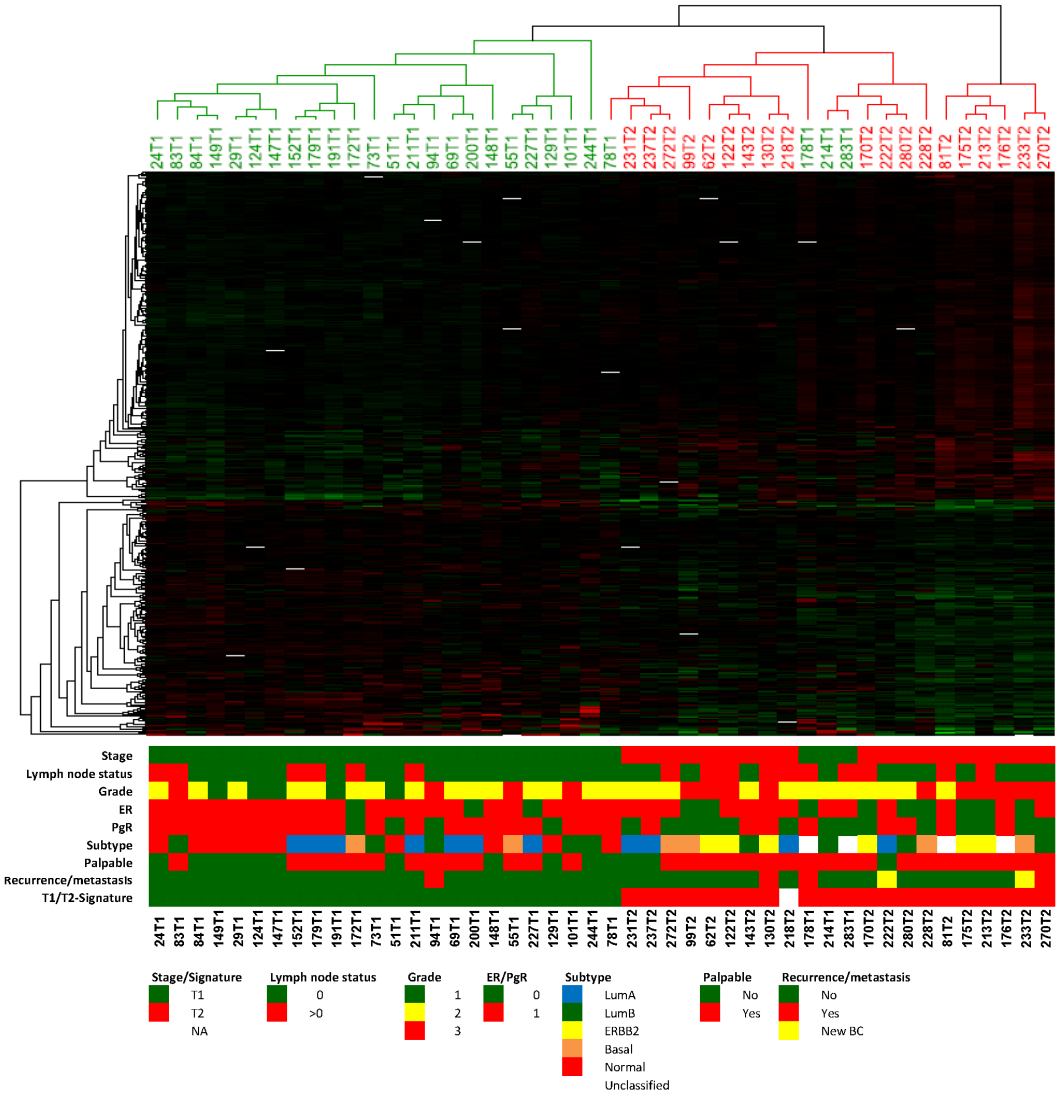


Figure 1 Unsupervised hierarchical clustering using the 441 significant probes after adjustment for clinical parameters. Genes are listed vertically and each patient is represented in the columns. Clinical T1 tumors are shown in green while clinical T2 tumors are shown in red.

7.3 Paper III

Polycomb Group (PcG) proteins are epigenetic silencers involved in maintaining cellular identity, and their deregulation can result in cancer(91). Mel-18 and Bmi-1 are both members of PcG. Bmi-1 was initially considered an oncogene, although recent studies have suggested thatBmi-1 overexpression is associated with good outcome in breast cancer. Mel-18 is considered a tumor suppressor gene. There are different mechanisms to this tumor suppressiveness.

Expression of Mel-18 and Bmi-1 have been studied in tumor tissue, but not in adjacent noncancerous breast epithelium. Our study compares the expression of the two genes in normal breast epithelium of cancer patients and compares this to the level of expression in breastepithelium of healthy women and suggests that the ratio of the expression of the two may be an important marker of disease progression.

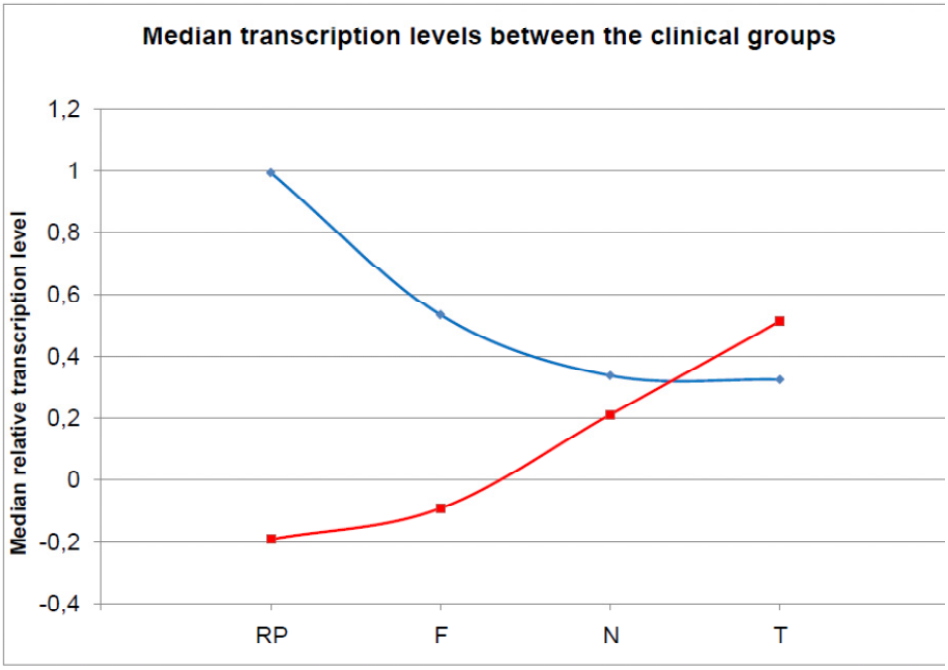


Figure 3: Median transcription levels between the groups. Comparison of the median transcription level of Bmi-1 and Mel-18 for each group shows an inverse correlation of the transcription level.

8 Discussion

As we have seen there are several gene signatures developed within breast cancer. This was revolutonary science more than a decade ago. It has truly visualized breast cancer as the heterogenous disease it is. But the signatures have not yet reached the clinics in a perspective that may improve breast cancer diagnostics and treatment. Our work is one attempt to add a piece in the great puzzle of breast cancer.

Fan et al (60) compared the prognostic ability of the intrinsic subtype and four prognostic signatures in 295 patients. They found that even though different gene sets were used for prognostication in patients with breast cancer, four out of five showed significant agreement

in outcome prediction for thje individual patients. This evaluation of gene signatures was performed in one dataset. Some years later there was a large metaanalysis of gene expression profiles in breast cancer which included 2833 breast tumors(92). They performed a prototype based expression module analysis where three prototype genes representing three key biological processes in breast cancer, namely proliferation, ER, and HER2 amplification signaling. The selected genes were respectively ESR1, ERBB2, and AURKA. This revealed the importance of proliferation genes as a common driving force behind the performance of all the prognostic signatures in the study. The clinical importance of these prognostic signatures will be validated in the MINDACT (Microarray in Node-Negative Disease May Avoid Chemotherapy)(59, 93) and TAILORX (Trial Assigning IndividuaLized Options for Treatment RX)(94). Since the levels of expression of proliferated genes are usually high in ER-negative cancers, first generation prognostic signatures almost invariably classify ER-negative cancers as of poor prognosis(92, 95-98). Most of these signatures correlate with response to conventional chemotherapy agents, which preferentially target proliferating cells(96-99). It was initially claimed that these first generation signatures could replace the classical clinicopathological features. Metaanalysis however have proved that tumor size and lymph node status provide prognostic information that is independent of that offered by prognostic signatures(55, 92, 95, 98, 100). The accuracy of the oucome prections of most prognostic signatures seems to be time-dependent, with more accurate prediction at 5 years than 10 years after diagnosis(58, 59). Evidence exists that these prognostic signatures do not provide more prognostic information beyond that offered by semiquantitative analysis of ER, PR, HER2, and Ki67(98, 101, 102).

Efforts were made to develop second generation signatures which better could classify the poor prognosis tumors. Immune response signature revealed ER-negative basal breast cancer

is a heterogeneous disease with at least four main subtypes and the heterogeneity in clinical outcome of ER-negative breast cancer is related to the variability in expression levels of complement and immune response pathway genes, independent of lymphocytic infiltration(103). Alexe et al(104) defined a subgroup of the HER2+ node negative breast cancer with better prognosis, and they found that these patients had high expression of lymphocytes associated genes.

In our work we have applied microarray to study the gene expression in different types of breast tissue. We have learned more about a selection of genes which may be used in the clinic both diagnostically and therapeutically. From being a disease with high mortality and little hope, breast cancer is a disease we now have treatment for, and the mortality has decreased. Prognosis today is based mainly on histopathological characteristics, but molecular profiling is slowly making an entrance. Can molecular studies add to the clinical parameters of classification to enable us to identify breast cancers in early stages and to identify patients who will benefit from adjuvant therapy? Can personalized cancer medicine be enabled through analysis of gene expression patterns(29)? These were the main questions and thoughts driving our work..

9. Further Perspective

How can our work help in diagnosis and treatment of breast cancer?

9.1 Micrometastases to the Sentinel Node

Sentinel lymph node dissection and examination in breast cancer surgery is considered a safe procedure with few complications(33). All patients with positive sentinel node are offered an

additional axillary dissection. Complete axillary dissection is extensive surgery which may result in severe arm morbidity(105). Micrometastases in the sentinel node is defined by metastasis of less than 2mm. National guidelines suggested axillary dissection in patients with micrometastasis to the sentinel node. However only 15% of patients with micrometastases to sentinel node have metastatic spread to non sentinel lymph nodes. The guidelines ”eased off” from autumn 2011 and we started to offer axillary node dissection only to patients who had performed mastectomy. The rationale behind this was that those patients who had breast conserving therapy would have irradiation towards the breast and one thought that any micrometastasis in the axilla would be erradicated in this radiation. In the summer of 2012 NBCG formally changed the guidelines of a selection of patients that did not need to have the sentinel node procedure with peroperative diagnostics performed due to the small possibility of metastases to the axilla.

Several attempts have been made to predict metastatic spread to non sentinel node in patients with micrometastasis to the sentinel node. A Danish group, Tvedskov et a(106)l, performed a study to evaluate TIMP-1, Ki67, and HER2 as markers for non-sentinel node metastases in breast cancer patients with micrometastases to the sentinel node. These are all known prognostic markers in breast cancer. However one was not able to obtain statistical significance in prognostication of these markers in considering which patients with micrometastasis to the sentinel node would have metastasis to non-sentinel lymph nodes.

In paper I(89) we studied the expression of Bmi1 both in genes and in proteins in different sources of breast tissue. We found that there was a significant difference in the expression of Bmi1 in histologically normal breast tissue of women operated for breast cancer as compared normal tissue in non cancer controls. Bmi-1 has been studied in plasma of breast cancer cells with

healthy women as controls and results show that levels of Bmi-1 expression may be a surrogate marker of poor prognosis(107).

In the future we want to measure Bmi1 in a case and a control group. The case group would be patients with micrometastases to the sentinel node in addition to metastases to non sentinel lymph node. The control group would be patients with micrometastases to the sentinenl node but without metastases to other axillary lymph nodes. We expect to get more patients in the control group so we can therefore calculate two controls for each case. The aim would be to see if this could be a possible prognostic marker that could predict metastatic spread to non-sentinel nodes in breast cancer patients with micrometastases to sentinel node.

9.2 Gene Expression Profiles Separating Tumor Tissue from Breast Cancer Patients and Histologically Normal Breast Epithelium both in Vicinity of Cancer and from Cancer Free Controls

Gene expression profiling of breast tumors has identified different subtypes with different clinical outcome and responses to therapy(50, 53, 108, 109).

These initial reports did not specifically study the stroma component of the tumors. Bergamaschi et al(110) investigated whether invasive breast tumors could be classified on the basis of the expression of extracellular components and whether such classification is representative of different clinical outcomes. Indeed they found that primary breast tumors could be classified based upon their extracellular matrix composition. This supports the hypothesis that clinical outcome is strongly related to stromal characteristics.

Laser capture microdissection (LCM) allows one to isolate nearly pure cell populations from a heterogenous environment. Combined with DNA microarray it is possible to perform a cellular based, rather than a tissue based, gene expression profile analysis. Epithelium derived

from regions of the breast adjacent to tumor, considered normal histopathologically, has been shown to have a distinct gene expression profile from tumor tissue(111). Patient-matched "normal" samples (adjacent to tumor) are highly similar to those from mammoplasty reduction specimen which tells us that normal tissue adjacent to tumor is more similar to normal tissue of non cancer patients than to the tumor tissue itself, which is as expected. In this respect the normal tumor adjacent tissue could be used as a baseline control for evaluating tumor progression which was used for further evaluation in mentioned paper(111).

They indeed found that alterations confering the potential for invasive growth are already present in the preinvasive stages(111).

How normal is the normal stroma? To focus on this question, Finak et al(112) did LCM to study normal epithelium and normal stroma derived from patients undergoing breast reduction mammoplasty or surgical treatment of breast cancer. They found no statistically difference in gene expression profiles on the two types of histopathologically normal tissue, which is in agreement with previous studie(111). Some of the same scientists however found a stroma-derived prognostic predictor (SDPP) that stratifies disease outcome independently of standard clinical prognostic factors and published expression-based predictors(113).

SDPP reveal the strong capacity of differential immune responses as well as angiogenic and hypoxic responses, highlighting the importance of stromal biology in tumor progression.

In our biobank we have gathered samples from tumor tissue, tumor adjacent tissue, and tissue from women having had reduction mammoplasties. In addition we have pairs of tumor tissue and normal adjacent tissue. Gene expression analysis was performed on these different tissue types with the intention to study if there are some characteristics in the stroma which are predictive of breast cancer which could be useful in a clinical setting. We have already found this in a previous study where two separate genes (Bmi-1 and Mel-18) and their gene products

were found to be upregulated and expressed already in normal tissue from cancer patients as opposed to normal tissue from breast cancer patients Riis et al(89).

9.3 Screening vs Interval Cancer

Interval breast cancers are those that are acknowledged between screening situations. Screening diagnosed breast cancer is thought generally to be less aggressive than these interval cancers. In some contries they are less aggressive in treatment of these screening diagnosed cancer. This has been studied with standard clinical markers which showed a difference in distribution of molecular subtypes according to mode detection(114). To support these results one may perform gene expression studies and generate a genetic profile of intervall cancer versus screening detected cancer

This may help us in evaluating screeningdiagnosed breast cancer and might allow us to treat these patient less aggressively but still give the same overall results. This is work that can be done retrospectively on expression data we already have worked out.

Reference list

1. Ferlay J, Shin HR, Bray F, Foreman D, Mathers C, Parkin DM. GLOBOCAN 2008 v1.2, Cancer Incidence and Mortality Worldwide. IARC CancerBase(Internet)2010.
2. SEER Cancer Statistics Review, 1975-2008 [database on the Internet]2011.
3. Tornberg S, Kemetli L, Lynge E, Helene OA, Hofvind S, Wang H, et al. Breast cancer incidence and mortality in the Nordic capitals, 1970-1998. Trends related to mammography screening programmes. Acta Oncologica. 2006;45(5):528-35.
4. Olsen AH, Njor SH, Lynge E. Estimating the benefits of mammography screening: the impact of study design. Epidemiology. 2007;18(4):487-92.
5. Haldorsen TGH, Solveig Hofvind, Ole-Erik Iversen, Rune Kvåle, Bente, Nygård KJaM. Cancer in Norway 2009. Special issue: Cancer screening in Norway. 2009;2009.
6. Chevarley F, White E. Recent trends in breast cancer mortality among white and black US women. American Journal of Public Health. 1997;87(5):775-81.
7. Aagnes BKAB, Freddie Bray, Aage F. Johansen, Tom Børge Johannesen,, Mette Kalager HL, Barbara Mortensen, Bjørn Møller, Jan F. Nygård, Børge Sildnes, Bjørge Sæther,, Svein Erling Tysvær FL. Cancer in Norway 2005. 2005.
8. Russo J, Russo IH. Development of the human breast. Maturitas. 2004;49(1):2-15.
9. Vorherr H. Breast Cancer: Epidemiology, Endocrinology, Biochemistry and Pathobiology. Baltimore: Urban & Schwarzenberg; 1980. p. 374-408.
10. Deome KB, Faulkin LJ, Jr., Bern HA, Blair PB. Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice. Cancer Research. 1959;19(5):515-20.
11. Vorherr H. The breast: morphology, physiology, and lactation. New York: Academic Press; 1974.
12. Diseases of the breast. Harris JR, editor. Philadelphia: Lippicott Williams & Wilkins; 2004.
13. Rudland PS. Stem cells and the development of mammary cancers in experimental rats and in humans. Cancer and Metastasis Reviews. 1987;6(1):55-83.
14. Rudland PS. Epithelial stem cells and their possible role in the development of the normal and diseased human breast. Histology and Histopathology. 1993;8(2):385-404.
15. Bissell MJ, Radisky D. Putting tumours in context. Nature reviews Cancer. 2001 Oct;1(1):46-54.
16. McCawley LJ, Matrisian LM. Tumor progression: defining the soil round the tumor seed. Current Biology. 2001;11(1):R25-R7.
17. Wiseman BS, Werb Z. Stromal effects on mammary gland development and breast cancer. Science. 2002;296(5570):1046-9.
18. Wernert N. The multiple roles of tumour stroma. Virchows Arch. 1997 Jun;430(6):433-43.
19. Townsend CM, R. D. Beauchamp, B. M. Evers, and K. L. Mattox Sabiston textbook of surgery: the biological basis of modern surgical practice. Townsend C, Sabiston DC, editors. Philadelphia: Elsevier Saunders; 2004.
20. Jacquemier J, Padovani L, Rabayrol L, Lakhani SR, Penault-Llorca F, Denoux Y, et al. Typical medullary breast carcinomas have a basal/myoepithelial phenotype. Journal of Pathology. 2005 Nov;207(3):260-8.
21. Singletary SE, Connolly JL. Breast cancer staging: working with the sixth edition of the AJCC Cancer Staging Manual. CA Cancer JClin. 2006;56(1):37-47.
22. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer JClin. 2005;55(2):74-108.
23. Paajanen H, Kyhala L, Varjo R, Rantala S. Effect of screening mammography on the surgery of breast cancer in Finland: a population-based analysis during the years 1985-2004. American Surgeon. 2006;72(2):167-71.

24. Polychemotherapy for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. Lancet. 1998;352(9132):930-42.
25. Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. BrJCancer. 1957;11(3):359-77.
26. Ciatto S, Cecchini S, Iossa A, Grazzini G. "T" category and operable breast cancer prognosis. Tumori. 1989;75(1):18-22.
27. Fitzgibbons PL, Page DL, Weaver D, Thor AD, Allred DC, Clark GM, et al. Prognostic factors in breast cancer. College of American Pathologists Consensus Statement 1999. ArchPatholLab Med. 2000;124(7):966-78.
28. Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group. Lancet. 1998;351(9114):1451-67.
29. van't Veer LJ, Bernards R. Enabling personalized cancer medicine through analysis of gene-expression patterns. Nature. 2008;452(7187):564-70.
30. Cody HS, 3rd. Clinical aspects of sentinel node biopsy. Breast Cancer Res. 2001;3(2):104-8.
31. Recht A, Pierce SM, Abner A, Vicini F, Osteen RT, Love SM, et al. Regional nodal failure after conservative surgery and radiotherapy for early-stage breast carcinoma. Journal of Clinical Oncology. 1991;9(6):988-96.
32. Vujovic O, Yu E, Cherian A, Dar AR, Stitt L, Perera F. The number of axillary nodes removed as a predictor of regional recurrence in node negative breast cancer. Radiother Oncol. 2009 Apr;91(1):38-41.
33. Veronesi U, Paganelli G, Viale G, Luini A, Zurrida S, Galimberti V, et al. A randomized comparison of sentinel-node biopsy with routine axillary dissection in breast cancer. New England Journal of Medicine. 2003;349(6):546-53.
34. Fisher B, Anderson S, Bryant J, Margolese RG, Deutsch M, Fisher ER, et al. Twenty-year follow-up of a randomized trial comparing total mastectomy, lumpectomy, and lumpectomy plus irradiation for the treatment of invasive breast cancer. New England Journal of Medicine. 2002;347(16):1233-41.
35. van Dongen JA, Voogd AC, Fentiman IS, Legrand C, Sylvester RJ, Tong D, et al. Long-term results of a randomized trial comparing breast-conserving therapy with mastectomy: European Organization for Research and Treatment of Cancer 10801 trial. Journal of the National Cancer Institute. 2000;92(14):1143-50.
36. Veronesi U, Cascinelli N, Mariani L, Greco M, Saccozzi R, Luini A, et al. Twenty-year follow-up of a randomized study comparing breast-conserving surgery with radical mastectomy for early breast cancer. New England Journal of Medicine. 2002;347(16):1227-32.
37. Bitterman A, Kessner R, Goldman I, Shiloni E, Steiner M. Intraoperative radiotherapy for breast cancer. The Israel Medical Association journal : IMAJ. 2012 Apr;14(4):256-9.
38. Trifiro G, Viale G, Gentilini O, Travaini LL, Paganelli G. Sentinel node detection in pre-operative axillary staging. Eur J Nucl Med Mol Imag. 2004;31:S46-S55.
39. Wei JS, Khan J. Purification of Total RNA from Mammalian Cells and Tissues. In: Bowtell D, Sambrook J, editors. DNA Microarrays: A Molecular Cloning Manual. New York: Cold Spring Harbor Laboratory Press; 2002. p. 110-9.
40. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 2004;5(10):R80.
41. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A. 1924;98(9):5116-21.
42. Mevik B-H, Wehrens R. The pls Package: Principal Component and Partial Least Squares Regression in R. Journal of Statistical Software. 2007 2007; 18(2):1-24.
43. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2010.
44. Huang dW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. NatProtoc. 2009;4(1):44-57.

45. Huang dW, Sherman BT, Tan Q, Collins JR, Alvord WG, Roayaei J, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol. 2007;8(9):R183.
46. Schemper M. The relative importance of prognostic factors in studies of survival. Statistics in medicine. 1993;12(24):2377-82.
47. Harrell F, Lee K, Mark D. Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in medicine. 1996;15:361-87.
48. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001 Dec;25(4):402-8.
49. Simon R. Challenges of microarray data and the evaluation of gene expression profile signatures. Cancer Investigation26(4)()(pp 327-332), 2008Date of Publication: May 2008. 2008(4):327-32.
50. Perou CM, Sorlie T, Eisen MB, Van De RM, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000;406(6797):747-52.
51. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. ProcNatlAcadSciUSA. 1998;95(25):14863-8.
52. van tV, Dai H, van dV, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature. 2002;415(6871):530-6.
53. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. ProcNatlAcadSciUSA. 2003;100(18):10393-8.
54. Ma XJ, Wang Z, Ryan PD, Isakoff SJ, Barmettler A, Fuller A, et al. A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. Cancer Cell. 2004;5(6):607-16.
55. Weigelt B, Baehner FL, Reis-Filho JS. The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. Journal of Pathology. 2010;220(2):263-80.
56. Ein-Dor L, Zuk O, Domany E. Thousands of samples are needed to generate a robust gene list for predicting outcome in cancer. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(15):5923-8.
57. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet. 2005;365(9460):671-9.
58. Desmedt C, Piette F, Loi S, Wang Y, Lallemand F, Haibe-Kains B, et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. Clinical Cancer Research. 2007;13(11):3207-14.
59. Buyse M, Loi S, Van't VL, Viale G, Delorenzi M, Glas AM, et al. Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. Journal of the National Cancer Institute. 2006;98(17):1183-92.
60. Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DS, Nobel AB, et al. Concordance among gene-expression-based predictors for breast cancer. New England Journal of Medicine. 2006;355(6):560-9.
61. Zhao X, Rodland EA, Sorlie T, Naume B, Langerod A, Frigessi A, et al. Combining gene signatures improves prediction of breast cancer survival. PloS one. 2011;6(3):e17845.
62. Venet D, Dumont JE, Detours V. Most random gene expression signatures are significantly associated with breast cancer outcome. PLoS Comput Biol. 2011 Oct;7(10):e1002240.
63. Spangrude GJ, Heimfeld S, Weissman IL. Purification and characterization of mouse hematopoietic stem cells. Science. 1988;241(4861):58-62.
64. Baum CM, Weissman IL, Tsukamoto AS, Buckle AM, Peault B. Isolation of a candidate human hematopoietic stem-cell population. ProcNatlAcadSciUSA. 1992;89(7):2804-8.
65. Smalley M, Ashworth A. Stem cells and breast cancer: A field in transit. NatRevCancer. 2003;3(11):832-44.

66. Dontu G, Al-Hajj M, Abdallah WM, Clarke MF, Wicha MS. Stem cells in normal breast development and breast cancer. Cell Proliferation. 2003;36 Suppl 1:59-72.
67. Williams JM, Daniel CW. Mammary ductal elongation: differentiation of myoepithelium and basal lamina during branching morphogenesis. Developmental Biology. 1983;97(2):274-90.
68. Smith GH. Experimental mammary epithelial morphogenesis in an in vivo model: evidence for distinct cellular progenitors of the ductal and lobular phenotype. Breast Cancer Research and Treatment. 1996;39(1):21-31.
69. Kordon EC, Smith GH. An entire functional mammary gland may comprise the progeny from a single cell. Development. 1998;125(10):1921-30.
70. Reya T, Morrison SJ, Clarke MF, Weissman IL. Stem cells, cancer, and cancer stem cells. Nature. 2001;414(6859):105-11.
71. Bruce WR, Van Der GAAG. A QUANTITATIVE ASSAY FOR THE NUMBER OF MURINE LYMPHOMA CELLS CAPABLE OF PROLIFERATION IN VIVO. Nature. 1963;199:79-80.
72. Park CH, Bergsagel DE, McCulloch EA. Mouse myeloma tumor stem cells: a primary cell culture assay. Journal of the National Cancer Institute. 1971;46(2):411-22.
73. Singh SK, Clarke ID, Terasaki M, Bonn VE, Hawkins C, Squire J, et al. Identification of a cancer stem cell in human brain tumors. Cancer Research. 2003;63(18):5821-8.
74. Al-Hajj M, Wicha MS, ito-Hernandez A, Morrison SJ, Clarke MF. Prospective identification of tumorigenic breast cancer cells. ProcNatlAcadSciUSA. 2003;100(7):3983-8.
75. Pecorino L. Molecular biology of cancer: mechanisms, targets, and therapeutics. Oxford: Oxford University Press; 2008.
76. Reya T, Duncan AW, Ailles L, Domen J, Scherer DC, Willert K, et al. A role for Wnt signalling in self-renewal of haematopoietic stem cells. Nature. 2003;423(6938):409-14.
77. Liu S, Dontu G, Wicha MS. Mammary stem cells, self-renewal pathways, and carcinogenesis. Breast Cancer Res. 2005;7(3):86-95.
78. Liu S, Dontu G, Mantle ID, Patel S, Ahn NS, Jackson KW, et al. Hedgehog signaling and Bmi-1 regulate self-renewal of normal and malignant human mammary stem cells. Cancer Research. 2006;66(12):6063-71.
79. Wodarz A, Nusse R. Mechanisms of Wnt signaling in development. Annual Review of Cell and Developmental Biology. 1998;14:59-88.
80. Polakis P. Wnt signaling and cancer. Genes and Development. 2000;14(15):1837-51.
81. Liu BY, McDermott SP, Khwaja SS, Alexander CM. The transforming activity of Wnt effectors correlates with their ability to induce the accumulation of mammary progenitor cells. ProcNatlAcadSciUSA. 2004;101(12):4158-63.
82. Li Y, Welm B, Podsypanina K, Huang S, Chamorro M, Zhang X, et al. Evidence that transgenes encoding components of the Wnt signaling pathway preferentially induce mammary cancers from progenitor cells. ProcNatlAcadSciUSA. 2003;100(26):15853-8.
83. Pasca di MM, Hebrok M. Hedgehog signalling in cancer formation and maintenance. NatRevCancer. 2003;3(12):903-11.
84. Mumm JS, Kopan R. Notch signaling: from the outside in. Developmental Biology. 2000;228(2):151-65.
85. Lessard J, Sauvageau G. Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. Nature. 2003;423(6937):255-60.
86. Park IK, Qian D, Kiel M, Becker MW, Pihalja M, Weissman IL, et al. Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. Nature. 2003;423(6937):302-5.
87. Molofsky AV, Pardal R, Iwashita T, Park IK, Clarke MF, Morrison SJ. Bmi-1 dependence distinguishes neural stem cell self-renewal from progenitor proliferation. Nature. 2003;425(6961):962-7.
88. Crawford YG, Gauthier ML, Joubel A, Mantei K, Kozakiewicz K, Afshari CA, et al. Histologically normal human mammary epithelia with silenced p16(INK4a) overexpress COX-2, promoting a premalignant program. Cancer Cell. 2004;5(3):263-73.

89. Riis ML, Luders T, Nesbakken AJ, Vollan HS, Kristensen V, Bukholm IR. Expression of BMI-1 and Mel-18 in breast tissue--a diagnostic marker in patients with breast cancer. BMCCancer. 2010;10:686.:686.
90. A. Parvaiz AI, C. Richardson and B. Isgar. Free hand breast core biopsies in a selected group are as good as image guided biopsies Ann Oncol (2012) 23 (suppl 2): ii27-ii31. 2012.
91. Sparmann A, van LM. Polycomb silencers control cell fate, development and cancer. NatRevCancer. 2006;6(11):846-56.
92. Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, et al. Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures. Breast Cancer Res. 2008;10(4):R65.
93. Bogaerts J, Cardoso F, Buyse M, Braga S, Loi S, Harrison JA, et al. Gene signature evaluation as a prognostic tool: challenges in the design of the MINDACT trial. Nature clinical practice Oncology. 2006 Oct;3(10):540-51.
94. Sparano JA. TAILORx: trial assigning individualized options for treatment (Rx). Clinical breast cancer. 2006 Oct;7(4):347-50.
95. Desmedt C, Haibe-Kains B, Wirapati P, Buyse M, Larsimont D, Bontempi G, et al. Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. Clinical Cancer Research. 2008;14(16):5158-65.
96. Reis-Filho JS, Weigelt B, Fumagalli D, Sotiriou C. Molecular profiling: moving away from tumor philately. Science translational medicine. 2010 Sep 1;2(47):47ps3.
97. Iwamoto T, Bianchini G, Booser D, Qi Y, Coutant C, Shiang CY, et al. Gene pathways associated with prognosis and chemotherapy sensitivity in molecular subtypes of breast cancer. Journal of the National Cancer Institute. 2011 Feb 2;103(3):264-72.
98. Reis-Filho JS, Pusztai L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. Lancet. 2011 Nov 19;378(9805):1812-23.
99. Iwamoto T, Pusztai L. Predicting prognosis of breast cancer with gene signatures: are we lost in a sea of data? Genome medicine. 2010;2(11):81.
100. Sotiriou C, Pusztai L. Gene-expression signatures in breast cancer. New England Journal of Medicine. 2009;%19;360(8):790-800.
101. Weigelt B, Reis-Filho JS. Molecular profiling currently offers no more than tumour morphology and basic immunohistochemistry. Breast Cancer Res. 2010;12 Suppl 4:S5.
102. Cuzick J, Dowsett M, Pineda S, Wale C, Salter J, Quinn E, et al. Prognostic value of a combined estrogen receptor, progesterone receptor, Ki-67, and human epidermal growth factor receptor 2 immunohistochemical score and comparison with the Genomic Health recurrence score in early breast cancer. Journal of Clinical Oncology. 2011 Nov 10;29(32):4273-8.
103. Teschendorff AE, Miremadi A, Pinder SE, Ellis IO, Caldas C. An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer. Genome Biol. 2007;8(8):R157.
104. Alexe G, Dalgin GS, Scanfeld D, Tamayo P, Mesirov JP, DeLisi C, et al. High expression of lymphocyte-associated genes in node-negative HER2+ breast cancers correlates with lower recurrence rates. Cancer Research. 2007 Nov 15;67(22):10669-76.
105. Gartner R, Jensen MB, Nielsen J, Ewertz M, Kroman N, Kehlet H. Prevalence of and factors associated with persistent pain following breast cancer surgery. JAMA. 2009 Nov 11;302(18):1985-92.
106. Tvedskov TF, Bartels A, Jensen MB, Paaschburg B, Kroman N, Balslev E, et al. Evaluating TIMP-1, Ki67, and HER2 as markers for non-sentinel node metastases in breast cancer patients with micrometastases to the sentinel node. APMIS. 2011 Dec;119(12):844-52.
107. Silva J, Garcia V, Garcia JM, Pena C, Dominguez G, Diaz R, et al. Circulating Bmi-1 mRNA as a possible prognostic factor for advanced breast cancer patients. Breast Cancer Res. 2007;9(4):R55.
108. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. ProcNatlAcadSciUSA. 2001;98(19):10869-74.

109. van dV, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, et al. A gene-expression signature as a predictor of survival in breast cancer. New England Journal of Medicine. 2002;347(25):1999-2009.
110. Bergamaschi A, Tagliabue E, Sorlie T, Naume B, Triulzi T, Orlandi R, et al. Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome. Journal of Pathology. 2008;214(3):357-67.
111. Ma XJ, Salunga R, Tuggle JT, Gaudet J, Enright E, McQuary P, et al. Gene expression profiles of human breast cancer progression. ProcNatlAcadSciUSA. 2003;100(10):5974-9.
112. Finak G, Sadekova S, Pepin F, Hallett M, Meterissian S, Halwani F, et al. Gene expression signatures of morphologically normal breast tissue identify basal-like tumors. Breast Cancer Res. 2006;8(5):R58.
113. Finak G, Bertos N, Pepin F, Sadekova S, Souleimanova M, Zhao H, et al. Stromal gene expression predicts clinical outcome in breast cancer. Nature Medicine. 2008;14(5):518-27.
114. Caldarella A, Puliti D, Crocetti E, Bianchi S, Vezzosi V, Apicella P, et al. Biological characteristics of interval cancers: a role for biomarkers in the breast cancer screening. J Cancer Res Clin Oncol. 2012 Sep 9.

Abbreviations

SEER	- Survaillance Epidemiology and End Results
LCIS	- Lobular carcinoma in situ
DCIS	- Ductal carcinoma in situ
UICC	- International Union Against Cancer
AJCC	- American Joint Committee on Cancer
T	- Tumor size
N	- Lymph node status
M	- Metasatasis
HER2	- Human Epidermal Growth Factor Receptor 2
DNA	- Deoxyribonucleic acid
RNA	- Ribo Nucleic acid
Tcf/LEF	-T-cell factor/lymphoid enhancing factor
APC	- adenomatous polyposis coli -
GSK	- glycogen synthase kinase
CKI	- glycogen synthase kinase
LRP5/6	- low-density lipoprotein receptor-related proteins 5 and 6
LEF/TCF	- lymphoid enhancer factor/T cell factor
PKC	- protein kinase C
CanKII	- calmodulin-dependent protein kinase II
Shh	- Sonic Hedgehog
Dhh	- desert Hedgehog
Ihh	- indian Hedgehog
Hip1	- hedgehog-interacting protein 1
Ptch	- Patched
Smo	- Smoothened

Fu	- Protein fused
SuFu	- Suppressor of fused
HDAC	- histone deacetylase
HAT	- histone acetyltransferase
Bmi1	- B lymphoma Mo MLV insertion region 1 homolog
MINDACT	- Microarray in Node-Negative Disease May Avoid Chemotherapy
TAILOR-X	-Trial Assigning IndividuaLized Options for Treatment RX
ER	- estrogen receptor
PR	- progesterone receptor
NBCG	- Norwegian Breast Cancer Group
SDPP	- stroma-derived prognostic predictor
PLS	- Partial Least Square

I

International Scholarly Research Network
ISRN Oncology
Volume 2012, Article ID 450267, 12 pages
doi:10.5402/2012/450267

Research Article

Molecular Profiles of Pre- and Postoperative Breast Cancer Tumours Reveal Differentially Expressed Genes

Margit L. H. Riis,[1,2] Torben Lüders,[2,3] Elke K. Markert,[4] Vilde D. Haakensen,[2,5] Anne-Jorun Nesbakken,[6] Vessela N. Kristensen,[2,3,5] and Ida R. K. Bukholm[1,2,7]

[1] *Department of Surgery, Akershus University Hospital, 1478 Lørenskog, Norway*
[2] *Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, 0318 Oslo, Norway*
[3] *Department of Clinical Molecular Biology and Laboratory Sciences (EpiGen), Akershus University Hospital, 1478 Lørenskog, Norway*
[4] *The Simons Center for Systems Biology, Institute for Advanced Study, Princeton, NJ 08540, USA*
[5] *Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, 0310 Oslo, Norway*
[6] *Department of Pathology, Akershus University Hospital, 1478 Lørenskog, Norway*
[7] *Institute of Health Promotion, Akershus University Hospital, 1478 Lørenskog, Norway*

Correspondence should be addressed to Vessela N. Kristensen, v.n.kristensen@medisin.uio.no

Received 31 August 2012; Accepted 18 October 2012

Academic Editors: E. A. Rakha and S. Ran

Gene expression studies on breast cancer have generally been performed on tissue obtained at the time of surgery. In this study, we have compared the gene expression profiles in preoperative tissue (core needle biopsies) while tumor is still in its normal milieu to postoperative tissue from the same tumor obtained during surgery. Thirteen patients were included of which eleven had undergone sentinel node diagnosis procedure before operation. Microarray gene expression analysis was performed using total RNA from all the samples. Paired significance analysis of microarrays revealed 228 differently expressed genes, including several early response stress-related genes such as members of the *fos* and *jun* families as well as genes of which the expression has previously been associated with cancer. The expression profiles found in the analyses of breast cancer tissue must be evaluated with caution. Different profiles may simply be the result of differences in the surgical trauma and timing of when samples are taken and not necessarily associated with tumor biology.

1. Introduction

Breast cancer is detected either by clinical signs such as palpable tumour or in mammographic screening. In both cases biopsies are taken from the tumour to determine whether the tumour is benign or malign. If malignancy is detected, the patient will be scheduled for surgery within a few weeks. Before the surgery, sentinel node (SN) diagnostics is generally performed to examine the spread of cancer cells to axillary lymph nodes. The SN can be identified using a blue dye, a radioactive colloid, or a combination of the two [1, 2].

Microarray technology enables scientists to study thousands of genes simultaneously. The resulting molecular profile can be used to study complex multifactorial diseases such as breast cancer [3, 4]. Gene signatures have been shown to correlate with clinically relevant clinicopathological parameters and prognosis [5–7]. These molecular signatures may be used to predict the individuals for whom therapy is beneficial and spare unnecessary treatment for over 80% of the others [6, 8–10].

The time of procurement, which refers to the point of when the biopsies are taken [11] as well as the postoperative handling [12], has been found to be a confounding factors in microarray data analysis in breast cancer. Most of the previously published studies consist of tumour tissue taken in connection to surgery. Biopsies taken from the tumour, while the tumour is within the breast prior to any manipulation, must be as near to the true expression state as possible. In this study, we analyzed whether there are differences in genes

expressed in preoperative biopsies obtained in connection with mammography and postoperative biopsies taken from the tumour immediately after its removal from the patient.

2. Materials and Methods

2.1. Patients. This study includes 13 patients from which both a pre- and postoperative samples were available. Histopathological characteristics are listed in Table 1. All of the patients had been operated with ablatio mammae (surgical removal of the entire breast). Sentinel node diagnostics using 99 m Tc-labelled colloids were performed in eleven patients as a part of the surgical procedure [13], while the remaining two underwent direct complete axillary dissection. No recurrence of disease has been observed so far for these patients, but the follow-up time is short. All women participating in this study have signed an informed consent and the study design is approved by the Regional Committee for Medical and Health Research Ethics (REK).

2.2. Tissue Collection. Preoperative needle biopsies were obtained by an experienced radiologist using a 16 Gauge core needle device through a small skin incision in a sterile field. Three samples were processed for routine histological diagnosis while one sample for molecular analysis was put directly into RNAlater (Sigma Aldrich, St Louis, MO, USA). The postoperative samples were taken by the breast cancer surgeon upon removal of the breast and were preserved in RNAlater. The RNAlater-stabilised tissue samples were stored at −80°C. The time delay between the sampling of the pre- and postoperative specimens were 2–8 weeks.

2.3. RNA Isolation. RNA was prepared using the method of Wei and Khan [14] but modified to also include miRNA. Briefly, frozen tissue samples were homogenized in TriReagent (Ambion, Austin, TX, USA) using a 5 mm steal bead in a Mixer Mill MM301 (Retsch, Haan, Germany) at 30 Hz for 2 min. After phase separation with 0.2 vol chloroform, the aqueous phase (containing RNA) was mixed with 1.5 vol 100% ethanol and transferred to RNeasy Mini columns (Qiagen, Hilden, Germany). Further processing (including on-column DNase digestion) was per the manufacturer's protocol and the purified RNA was eluted with RNase-free water. RNA concentration was measured using NanoDrop ND-1000 UV-VIS Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and the RNA quality analyzed on a 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA). The purified RNA was stored at −80°C.

2.4. Microarray Analysis. mRNA amplification, labelling, and hybridization were done following the manufacturer's instructions (Agilent One-Color Microarray-Based Gene Expression Analysis; Version 5.7). Briefly, 500 ng RNA was amplified and labelled with Cy3 using the Quick Amp labelling kit and the labelled cRNA purified using the Qiagen RNeasy Mini Kit. Amplification and labelling efficiency were controlled on the NanoDrop before 1.65 μg cRNA was fragmented and applied to Agilent Whole Human Genome 4 × 44 k microarrays (G4112F). After hybridisation for 17 h at 65°C the microarray slides were washed and scanned with the Agilent Microarray Scanner. Microarray data were extracted using Agilent Feature Extraction (v. 10.7.1.1) and further quantile normalized and analyzed using J-Express 2009 [15]. For expression values the gProcessedSignal from Feature Extraction were used and controls and bad spots were filtered with maximum 20% allowed missing values. The expression values were log2-transformed and missing values imputed using the LSImpute Adaptive method. Differential expression was analyzed using SAM as implemented in J-Express with 1000 permutations and only genes with false discovery rate (FDR) < 2.5% were considered significant. The microarray data are available at the ArrayExpress Archive (http://www.ebi.ac.uk/microarray-as/ae/) accession number E-MTAB-470.

Gene functional classification of the significant genes was performed using DAVID [16, 17] and pathway analysis was done through the use of Ingenuity Pathways Analysis (IPA; Ingenuity Systems, Redwood City, CA, USA).

2.5. Quantitative RT-PCR. To confirm the results of the microarray experiment, qRT-PCR using TaqMan low density arrays (Applied Biosystems, Carlsbad, CA, USA) were performed using primer-probe pairs for 13 of the significant genes (Table 2). The genes were selected to contain both up- and downregulated genes. 500 ng RNA was reverse transcribed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) per the manufacturer's instructions. Due to lack of material, qRT-PCR was only performed for ten of the sample pairs. The samples were further processed using TaqMan Gene Expression Master Mix (Applied Biosystems) and run on the 7900HT Real-Time PCR System (Applied Biosystems) as per the manufacturer's instructions. Relative changes in gene expression were analyzed using the ΔΔCt-method [18] with the preoperative sample as control sample for each pair. As endogenous controls the average of *GAPDH*, *18S*, and *ACTB* were used.

3. Results

3.1. mRNA Expression. RNA was isolated from matching samples taken both before and after breast cancer surgery of 13 patients. After filtering, expression data were available from 24,105 different probes representing 18,189 different genes. Comparing the gene expressions of the 13 pairs showed that there was relatively little difference between the pre- and postoperative samples (Figures 1(a) and 1(b) and Supplementary Figure 1 available online at doi:10.5402/2012/450267). Paired significance analysis of microarrays (SAM) [19], however, showed differently expression for 235 probes with false discovery rate (FDR) <2.5%, corresponding to 228 different genes (Supplementary Table 1) that separates the pre- and postoperative samples (Figures 1(c) and 1(d) and Supplementary Figure 2). The majority (201) of these genes were upregulated and only 27 were downregulated in the postoperative samples. The differentially expressed genes contained genes involved in early response such as *FOSB*, response to oxidative stress such

Table 1: Histopathological characteristics of the patients included in the study.

Case	Age	Tumour type*	TNM	Tumour size (cm)	Preoperative				Postoperative		
					Tumor content (%)	ER† (%)	PGR†	HER2†	Tumor content (%)	ER† (%)	PGR†
2	54	IDC	T2gr3N2M0	4.3	30	+ (>50)	−	+	30	−	−
3	42	IDC	T2gr3N0M0	4.8	40	−	−	−	60	−	−
4	67	IDC	T2gr2N1M0	3.5	50	+ (>50)	+	−	55	+ (>50)	+
5	82	IDC	T1cgr3N0M0	1.8	30	+ (>50)	+	−	40	+ (>50)	+
6	52	IDC	T2gr3N3M0	2.0	5	−	−	+	5	−	−
7	68	IDC	T1cgr3N1M0	1.3	35	+ (>50)	−	+	40	+ (>50)	+
8	76	IDC	T1cgr2N0M0	1.4	40	−	−	−	35	−	−
9	70	IDC	T1bgr1N0M0	0.9	25	+ (>50)	+	−	25	+ (>50)	+
10	77	IDC	T2gr3N1M0	2.6	50	+ (>50)	−	−	15	+ (>50)	−
11	61	IDC	T2gr3N1M0	2.5	50	−	−	+	40	−	−
13	79	IDC	T2gr3N0M0	2.3	45	+ (>10)	−	+	55	+ (>1)	−
14	70	IDC	T2gr2N0M0	2.3	35	+ (>50)	+	−	35	+ (>50)	+
15	68	ILC	T2gr2N0M0	2.5	50	+ (>50)	−		30	+ (>50)	+

*IDC: infiltrating ductal carcinoma; ILC: infiltrating lobular carcinoma.
†ER: oestrogen receptor status; PR: progesterone receptor status; HER2: HER2 receptor status.

Table 2: TaqMan assays used for validation qRT-PCR and correlation between the microarray and the qRT-PCR results. The *P* values given are for positive correlation. *r*: Pearson's product-moment correlation.

Gene	Array probe ID	TaqMan assay ID	Task	*r*	*P*
18S	NA	Hs99999901_s1	Endogenous control	NA	NA
ACTB	A_23_P31323 A_24_P226554 A_24_P226554 A_32_P137939	Hs99999903_m1	Endogenous control	0.904	$1.7E-4$
ANGEL2	A_24_P28622	Hs00404357_m1	Target	0.238	0.254
CYP2D6	A_23_P143734 A_23_P155123	Hs02576167_m1	Target	0.262	0.232
CYR61	A_23_P46426 A_24_P370946	Hs00155479_m1	Target	0.894	$2.4E-4$
DUSP1	A_23_P110712	Hs00610256_g1	Target	0.983	$1.7E-7$
DUSP9	A_24_P417189	Hs00154830_m1	Target	−0.630	0.965
EVI2B	A_23_P66694	Hs00272421_s1	Target	0.945	$1.8E-5$
FOSB	A_23_P429998	Hs00171851_m1	Target	0.749	0.006
GAPDH	A_23_P13899	Hs99999905_m1	Endogenous control	0.632	0.025
MALAT1	A_24_P497244	Hs00273907_s1	Target	0.332	0.174
MAPK3	A_23_P37910	Hs00385075_m1	Target	−0.018	0.520
NFRκB	A_23_P24485	Hs00196269_m1	Target	−0.169	0.680
PTPRE	A_24_P213494 A_24_P213503	Hs00369944_m1	Target	0.413	0.118
RASD1	A_24_P348006 A_23_P118392	Hs02568415_s1	Target	0.797	0.003
TMEM19	A_24_P358976	Hs00217586_m1	Target	0.248	0.245

Table 3: Selected genes that are differently expressed between pre- and postoperative samples.

Gene name	Agilent ID	Description	SAM			Fold Change	
			Called	FDR	q-val	Mean	Range
CX3CL1	A_24_P390495	Chemokine (C-X3-C motif) ligand 1 (*CX3CL1*), mRNA (NM_002996)	21	0	0	2.99	0.90–4.09
CYP2D6	A_23_P143734	Cytochrome P450, family 2, subfamily D, polypeptide 6 (*CYP2D6*), transcript variant 1, mRNA (NM_000106)	18	0	0	1.65	0.58–3.55
	A_23_P155123		103	2.08	1.55	1.51	0.46–3.38
CYR61	A_23_P46426	Cysteine-rich, angiogenic inducer, 61 (*CYR61*), mRNA (NM_001554)	163	1.75	1.65	4.07	0.61–12.88
	A_24_P370946		196	1.82	1.79	5.51	0.36–22.80
DUSP1	A_23_P110712	Dual specificity phosphatase 1 (*DUSP1*), mRNA (NM_004417)	217	2.3	2.20	3.18	0.49–12.69
DUSP9	A_24_P417189	Dual specificity phosphatase 9 (*DUSP9*), mRNA (NM_001395)	7	0	0	2.12	1.15–2.88
FOSB	A_23_P429998	FBJ murine osteosarcoma viral oncogene homolog B (*FOSB*), transcript variant 1, mRNA (NM_006732)	203	2.11	2.01	2.79	0.96–24.26
MALAT1	A_24_P497244	Metastasis associated lung adenocarcinoma transcript 1 (nonprotein coding) (*MALAT1*), noncoding RNA (NR_002819)	3	0	0	3.54	1.31–9.97
MAPK3	A_23_P37910	Mitogen-activated protein kinase 3 (*MAPK3*), transcript variant 1, mRNA (NM_002746)	42	1.70	1.30	2.01	0.40–2.55
NFRκB	A_23_P24485	nuclear factor related to kappaB binding protein (*NFRκB*) , transcript variant 2, mRNA (NM_006165)	183	1.95	1.79	1.61	0.63–3.25
RAB17	A_23_P5778	*RAB17*, member RAS oncogene family (*RAB17*), mRNA (NM_022449)	10	0	0	1.84	0.23–4.36
RASAL1	A_23_P139600	RAS protein activator like 1 (GAP1 like) (*RASAL1*), mRNA (NM_004658)	140	2.04	1.65	1.46	0.38–2.41
RASD1	A_23_P118392	RAS, dexamethasone-induced 1 (*RASD1*), mRNA (NM_016084)	105	2.04	1.55	3.14	0.88–21.55
	A_24_P348006		27	0	0	2.69	0.85–12.90
RHOB	A_23_P51136	ras homolog gene family, member B (*RHOB*), mRNA (NM_004040)	16	0	0	1.99	0.51–2.96
RHOU	A_23_P114814	ras homolog gene family, member U (*RHOU*), mRNA (NM_021205)	122	1.75	1.55	2.68	0.71–3.65
RHOV	A_23_P117912	Rho-related GTP-binding protein RhoV (Wnt-1 responsive Cdc42 homolog 2)(WRCH-2)(CDC42-like GTPase 2)(GTP-binding protein-like 2) (Rho GTPase-like protein ARHV) (ENST00000220507)	184	1.94	1.79	1.60	0.25–5.45

as *DUSP1,9* as well as genes earlier identified as differentially expressed in cancer (*MAPK*, *MALAT1*, *RASD1*, etc) (Table 3).

Gene functional classification in DAVID of the upregulated genes showed enrichment for four groups (kinase/phosphatase, Ras, negative regulation of transcription, and transmembrane) while the downregulated genes mainly correspond to transmembrane proteins (Table 4). Gene function was also analyzed by Ingenuity Pathways Analysis (IPA) and includes "cellular movements," "connective tissue development and movement" and "cellular growth and proliferation" (Figure 2). IPA also identified molecular networks connecting several of the genes: *FOSB*, *ERK*, *MAPK3*, *CYR61*, and the *RAS*-genes (Figure 3(a)); *DUSP1*, *ERK1/2*, *P38MAPK*, *DUSP9*, and *RASD1* (Figure 3(b)); *CYR61* and *NFRκB* (Figure 3(c)) amongst other (Supplementary Figure 3).

Table 4: Gene functional classification (DAVID) of the differently expressed genes.

Gene name	Agilent ID	Description
	Gene group 1 Kinase/phosphatase	Enrichment score: 2.11
DCAKD	A_24_P58331	Dephospho-CoA kinase domain containing
SIK2	A_23_P138957	Salt-inducible kinase 2
ITPKC	A_23_P208369	Inositol 1,4,5-trisphosphate 3-kinase C
DAK	A_23_P36129	Dihydroxyacetone kinase 2 homolog (*S. cerevisiae*)
RIPK4	A_23_P211267	Receptor-interacting serine-threonine kinase 4
CHKA	A_23_P136135	Choline kinase alpha
DDR1	A_23_P93311, A_24_P367289	Discoidin domain receptor tyrosine kinase 1
STK35	A_24_P940537	Serine/threonine kinase 35
ACTR3B	A_23_P123193	ARP3 actin-related protein 3 homolog B (yeast)
INO80	A_24_P39454	INO80 homolog (*S. cerevisiae*)
EPHA1	A_23_P157333	EPH receptor A1
BCR	A_24_P15270	Breakpoint cluster region
CAMK1D	A_23_P124252	Calcium/calmodulin-dependent protein kinase ID
HISPPD2A	A_23_P205818	Histidine acid phosphatase domain containing 2A
	Gene group 2 Ras	Enrichment score: 1.35
RHOB	A_23_P51136	Ras homolog gene family, member B
RHOV	A_23_P117912	Ras homolog gene family, member V
RASD1	A_24_P348006, A_23_P118392	RAS, dexamethasone-induced 1
RAB17	A_23_P5778	RAB17, member RAS oncogene family
RHOU	A_23_P114814	Ras homolog gene family, member U
	Gene group 3 Negative regulation of transcription	Enrichment score: 1.26
ARID5B	A_23_P97871	AT rich interactive domain 5B (MRF1-like)
COBRA1	A_23_P148150	Cofactor of BRCA1
TH1L	A_24_P222126	TH1-like (Drosophila)
FOXD3	A_23_P46560	Forkhead box D3
EID2	A_23_P365844	EP300 interacting inhibitor of differentiation 2
	Gene group 4 Transmembrane	Enrichment score: 0.42
PQLC1	A_24_P181677	PQ loop repeat containing 1
RNF215	A_32_P420563	Ring finger protein 215
KIAA1305	A_23_P129005	KIAA1305
TMEM49	A_32_P9753	Transmembrane protein 49
F11R	A_24_P319369	F11 receptor
RBM8A	A_23_P305335	Gonadotropin-releasing hormone (type 2) receptor 2
KIAA0922	A_23_P257250	KIAA0922
TSPAN12	A_23_P145984	Tetraspanin 12
DGCR2	A_24_P125881	DiGeorge syndrome critical region gene 2
PCDH1	A_23_P213359	Protocadherin 1
LMBRD2	A_32_P8952	LMBR1 domain containing 2
GPR65	A_23_P14564	G protein-coupled receptor 65
EVI2B	A_23_P66694	Ecotropic viral integration site 2B
RTF1	A_24_P93741	RFT1 homolog (*S. cerevisiae*)
TMEM19	A_24_P358976	Transmembrane protein 19

Table 4: Continued.

Gene name	Agilent ID	Description
GPR155	A_23_P335958	G protein-coupled receptor 155
OSMR	A_24_P145134	Oncostatin M receptor
TMEM97	A_32_P201521	Transmembrane protein 97
PTPRE	A_24_P213503, A_24_P213494	Protein tyrosine phosphatase, receptor type, E

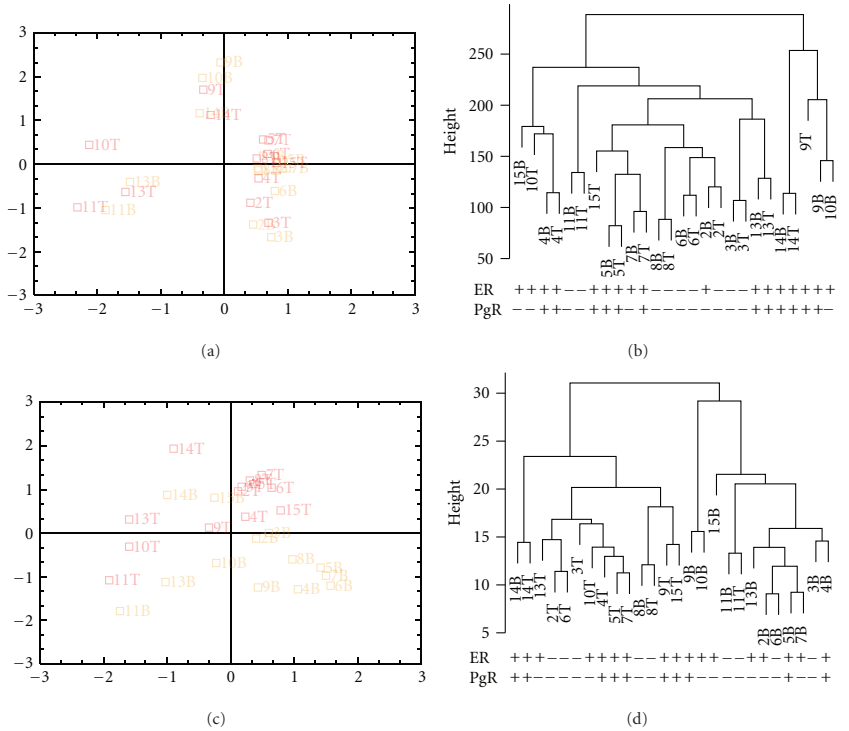


Figure 1: Unsupervised (a) and supervised (c) CA plots of mRNA expression in pre- and postoperative breast cancer tumours, and their corresponding clustering maps (b, d). The samples are marked by their respective numbers followed by either B, which defines the preoperative samples, or T, which defines the postoperative samples. The unsupervised chart was made using all 18,189 genes (24,105 probes) expressed on the microarrayss whereas for the supervised only the 228 genes (235 probes) with FDR < 2.5% from paired SAM were used.

3.2. Quantitative RT- PCR Validation. To confirm the results of the microarray experiment, qRT-PCR was performed using primer-probe pairs for the top significant genes. The genes were selected to contain both up- and downregulated genes. The microarray and the qRT-PCR results were in agreement with the following genes (Figure 4 and Table 2): *ACTB, CYR61, DUSP1, EVI2b, FOSB, GAPDH*, and *RASD1*.

3.3. Histological Analysis versus Gene Expression Analysis. Immunohistochemistry was performed on the pre- and the postoperative samples. Overall the tumour content in the two samples were comparable and there was no systematic bias (Table 1), indicating that the gene expression as measured by microarray is comparable in the pairs. In addition, ER and PGR status for the pre- and postoperative specimen were similar (Table 1).

4. Discussion

Microarray studies have influenced breast cancer research over the last decade revealing breast cancer as a heterogeneous disease opening for individual treatment in a clinical

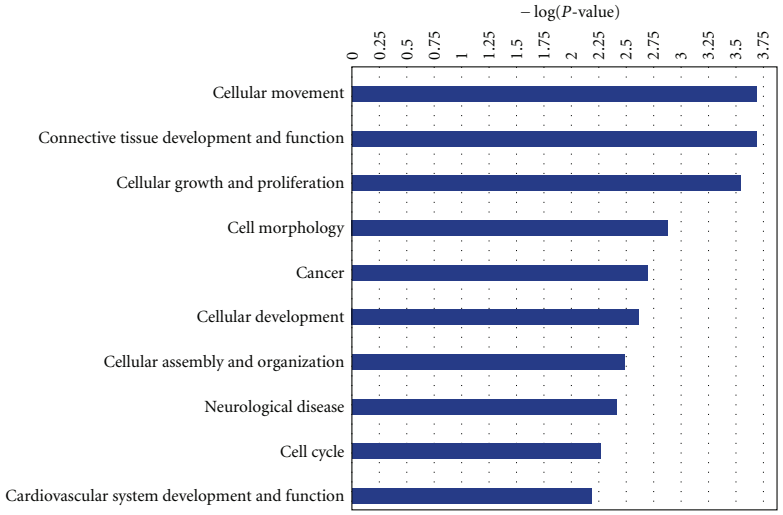


Figure 2: The most significantly enriched biological categories as identified with Ingenuity Pathway Analysis. For each category −log(*P* value) is reported.

perspective. Therefore, the results from microarray studies need to be validated. Multiple studies have generated different gene list and studied the reproducibility and correlation with prognosis [20–22]. Despite the difference in development of these signatures and the limited overlap in gene identity, they show similar prognostic performance, adding to the growing evidence that these prognostic signatures are of clinical importance [20]. There are two prospective ongoing studies, the MINDACT trial [23] in Europe and TAILORx [24] in USA which will evaluate the prognostic potential of this technology.

One important question may be if the differences in gene expressions are related to tumour biology or reflect the surgical trauma of the patient or the manipulation of the tumour tissue during the operative procedure or the time of specimen handoff. If altered gene expression is caused by such exogenous factors, the results may differ considerably between studies depending on the operative procedure and the time spent at the operation before taking the tissue samples. It is therefore important to evaluate if gene expression patterns differ between biopsies taken before and after surgical procedure. This has been done in our study with 13 patients and the gene list of 228 genes was dominated by stress-related genes like *CYR61*, *MALAT1*, *RASD1*, *CX3CL1*, *FOSB*, and *CYP2D6*. Some of these genes have been studied by others in relation to oxidative stress [25–27] and also psychological stress [28]. These genes have different functions all included in very important pathways with strong hubs such as *MAPK3*, *NFRκB*, *FOS,* and *ERK*.

Upregulation of Fos has been associated with breast cancer in a number of studies [29–31]. The fos-gene family consists of 4 members: *FOS*, *FOSB*, *FOSL1*, and *FOSL2*. These genes encode leucine zipper proteins that can dimerise with proteins of the jun-family, and the Fos-proteins have been implicated as regulators of cell proliferation, differentiation, and transformation. Another gene, *CYR61* (cysteine-rich, angiogenic inducer, 61), most strongly associated to differential expression in pre- and postoperative samples, belongs to the CCN-family [32] and mediates cell proliferation, survival, and apoptosis. Acting as an extracellular matrix-associated signalling molecule, *CYR61* promotes the adhesion of endothelial cells through interaction with the integrin $\alpha v\beta 3$ and augments growth factor-induced DNA synthesis in the same cell type [33]. In this aspect, it is both chemotactic and angiogenic, two properties important for tumour growth and vascularisation. *CYR61* is claimed to play a critical role in oestrogen, as well as growth factor-dependent breast tumour growth [34]. In our list of genes, *CYR61* is repeatedly connected in most of the involved pathways. Further studies will be necessary to confirm and explain this association.

It is of particular importance to take into consideration knowledge about gene expression differences in pre- and postoperative tissue samples in the case of treatment response studies in the neoadjuvant setting, when the first sample is frequently taken by biopsy and the second during operation. In a study comparing gene expression profiles before and after doxorubicin and cyclophosphamide neoadjuvant chemotherapy [35] one of the genes upregulated after the first chemotherapy treatment was *DUSP1*. Expression of this gene may be associated with resistance to further administration of chemotherapy. In our study *DUSP1* was one of the significantly upregulated genes in postoperative tissue. *DUSP1* is a stress response gene of the mitogen-activated protein (*MAP*) kinase phosphatase family and is located in the cytoplasm, mitochondria, and the nucleus.

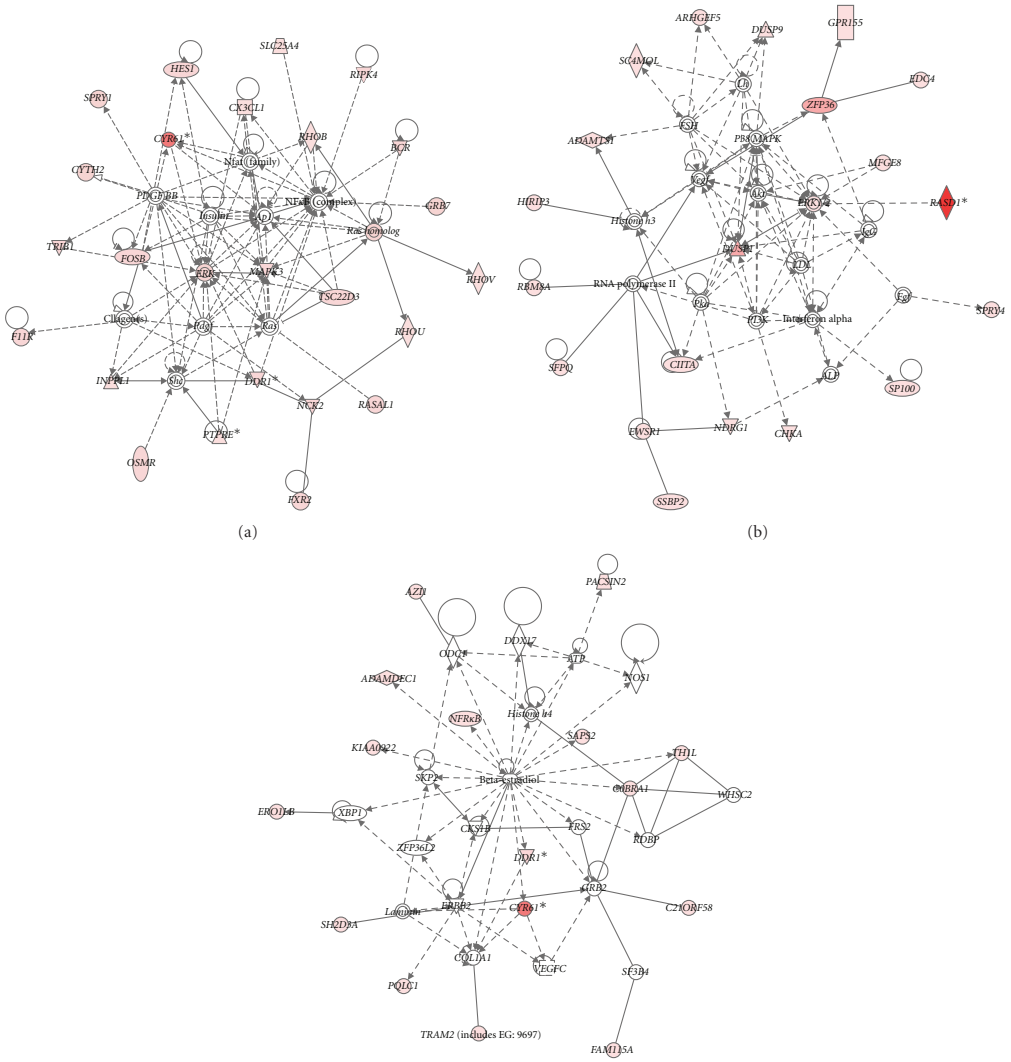


Figure 3: Most enriched molecular networks showing interactions between the significant genes (according to SAM)—(a) *FOSB*, *ERK*, *MAPK3*, *CYR61*, and the *RAS*-genes; (b) *DUSP1*, *ERK1/2*, *P38MAPK*, *DUSP9*, and *RASD1*; (c) *CYR61* and *NFRκB*. The gene identifiers and corresponding expression values were uploaded into in the Ingenuity Pathway Analysis. Networks were then algorithmically generated based on their connectivity in Ingenuity's Knowledge Base. Molecules are represented as nodes, and the biological relationship between two nodes is represented as an edge (line). The intensity of the node colour indicates the degree of (red) up- or (green) downregulation. Nodes are displayed using various shapes that represent the functional class of the gene product.

The gene has been shown to be overexpressed in human breast cancer [36] through different signalling pathways. One important pathway is in response to stress which is mediated in part through the p38 *MAPK* pathway. Later studies have implicated that *DUSP1* is controlled by p53 during cellular response to oxidative stress [37]. A similar discussion could be relevant on molecular profiling of inflammatory breast cancer [29], where *DUSP1* was also among the genes suggested to be useful diagnostic and prognostic markers. Our study suggests that such findings have to (1)

Table 5: Intrinsic subtypes of the tumours. Samples with all correlations < 0.1 were not assigned to any subtype.

Patient no.	Sample type	Score						Subtype
		LumA	LumB	ERBB2	Normal	Basal	Max	
2	Preoperative	−0.050	0.231	0.103	−0.191	−0.034	0.231	Luminal B
	Postoperative	−0.081	0.193	0.010	0.011	0.062	0.193	Luminal B
3	Preoperative	−0.465	0.213	0.215	−0.061	0.616	0.616	Basal-like
	Postoperative	−0.454	0.177	0.190	0.010	0.668	0.668	Basal-like
4	Preoperative	0.064	0.152	−0.058	−0.268	−0.085	0.152	Luminal B
	Postoperative	0.085	0.063	0.017	−0.271	−0.227	0.085	NA
5	Preoperative	−0.001	0.032	−0.028	0.014	−0.116	0.032	NA
	Postoperative	−0.050	0.088	−0.061	−0.056	−0.079	0.088	NA
6	Preoperative	−0.454	0.067	0.477	−0.017	0.374	0.477	ERBB2
	Postoperative	−0.576	0.096	0.531	−0.015	0.410	0.531	ERBB2
7	Preoperative	0.234	0.031	−0.165	−0.212	−0.224	0.234	Luminal A
	Postoperative	0.123	−0.115	−0.106	0.067	−0.185	0.123	Luminal A
8	Preoperative	−0.117	−0.011	0.222	−0.066	0.031	0.222	ERBB2
	Postoperative	−0.066	−0.060	0.144	−0.065	−0.039	0.144	ERBB2
9	Preoperative	−0.127	−0.189	0.149	0.188	0.055	0.188	Normal-like
	Postoperative	−0.040	−0.220	0.077	0.266	0.005	0.266	Normal-like
10	Preoperative	0.032	0.039	−0.025	−0.085	−0.138	0.039	NA
	Postoperative	−0.003	0.218	−0.082	−0.201	−0.035	0.218	Luminal B
11	Preoperative	−0.269	0.260	0.242	−0.163	0.217	0.260	Luminal B
	Postoperative	−0.147	0.292	0.132	−0.174	0.086	0.292	Luminal B
13	Preoperative	−0.266	**0.130**	0.166	−0.002	**0.211**	0.211	**Basal-like***
	Postoperative	−0.105	**0.246**	0.089	−0.197	**0.007**	0.246	**Luminal B***
14	Preoperative	**0.169**	−0.111	−0.142	**−0.017**	−0.170	0.169	**Luminal A***
	Postoperative	**0.068**	−0.175	−0.092	**0.196**	−0.026	0.196	**Normal-like***
15	Preoperative	0.356	−0.013	−0.328	−0.188	−0.338	0.356	Luminal A
	Postoperative	0.223	−0.065	−0.131	−0.094	−0.349	0.223	Luminal A

*Different subtypes in the pre- and postoperative samples.

demonstrate upregulation above the one observed here by us attributable to the pre- and postoperative factors and (2) that deregulation attributable to the pre- and postoperative factors is similar in the compared case/control or treatment arm groups.

In the present study, it is not possible to separate the effects of operative manipulation, anaesthesia, or the injection of radioactive substance to examine spread of cancer cells. For the latter, we should have had a biopsy after the application of radioactivity not before operation. However, both ethical and logistical considerations make collection of such a sample infeasible. Wong et al. [11] studied the effects of timing of fine-needle aspiration biopsies. Using hierarchical clustering analysis, they found 12 genes to be differentially expressed before and after surgery, which were in agreement with our study all fos-related. However, it was unclear whether any other treatment, like sentinel node, was given to patients between the two time points. It has been previously shown that both fine-needle aspiration biopsy and central core biopsy yield a similar quality and quantity of total RNA and that microarray profiles are mainly the same [38]. Microscopic cell counts have demonstrated that there are more stromal cells present in core biopsies compared to fine-needle biopsies [38], and the core biopsy is therefore needed for the complete histological examination.

Another confounding factor in the analysis of gene expression profiles of breast cancers is intratumour heterogeneity [39, 40]. Even though this study was not designed to analyze this, molecular subclassification [41] of the samples did in a few cases give different result for the pre- and postoperative samples (Table 5). Interference from surrounding normal tissue is not likely since the overall gene expression profiles of the pre- and postoperative samples were very similar and distinct from that of adjacent normal tissue (Supplementary Figure 1), thus, suggesting true cases of intratumour heterogeneity.

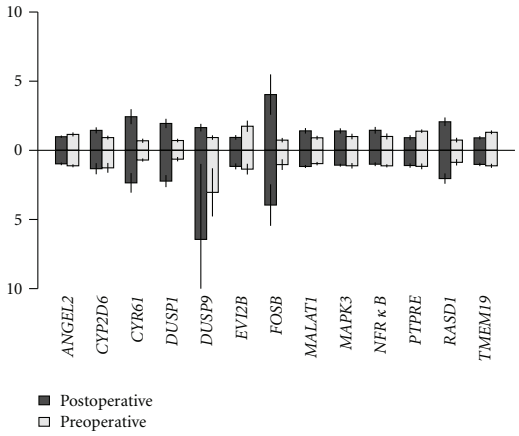


Figure 4: Relative expression of selected genes on microarrays (upwards) and qRT-PCR (downwards). Values shown are mean expression ±SE. For easier comparison, both the microarray and qRT-PCR values were gene-centered.

As we have seen there are genes that are differently expressed between the pre- and postoperative samples. We compared our gene lists to some of the publically available gene list to see if there were any overlapping genes. The Oncotype DX consists of 21 genes, 16 cancer related genes, and 5 reference genes [42]. One of these is also found in our gene list which differs between pre- and postoperative samples, namely, *GRB7*, which is upregulated in the postoperative samples. *GRB7* was associated with an increased risk of recurrence in TNBC (tripple negative breast cancer) treated with adjuvant doxorubicin-containing chemotherapy, suggesting that *GRB7* or *GRB7*-dependent pathways may serve as potential biomarkers for therapeutic targets [43, 44]. We have shown that this gene is upregulated in the postoperative sample. Even though the gene has been well characterized *in vitro* [45–47], it is of interest that we find it in the list of genes separating pre- and postoperative samples.

We also wanted to compare the 70 genes listed in the Mammaprint which were based on the intrinsic gene list [6]. As with the Oncotype DX, there was only one single gene (*NDRG1*) in common for the 70 gene list in Mammaprint with our gene list separating pre- and postoperative samples. *NDRG1* (N-myc downstream-regulated gene 1) is a member of the N-myc downregulated gene family which belongs to the alpha/beta hydrolase superfamily. The protein encoded by this gene is a cytoplasmic protein involved in stress responses, hormone responses, cell growth, and differentiation. The encoded protein is necessary for p53-mediated caspase activation and apoptosis. Expression of this gene may be a prognostic indicator for several types of cancer (provided by RefSeq, May 2012). The gene is significantly upregulated in the postoperative samples of the present study. It is known to be induced by stress, through hypoxia [48], like many of the other genes mentioned above.

Low expression of *NDRG1* is correlated with poor clinical outcome in breast cancer [49]. It has also been shown that expression of *NDRG1* is downregulated upon estradiol stimulation, and its expression is correlated with favorable prognosis in breast cancer patients [50]. On the other hand, induction of its differentiation is considered a promising alternative or complementary to standard anticancer chemotherapy. One may speculate why this gene is upregulated in the postoperative samples. Stress is probably the cause, but since the gene is a positive predictive factor, can we then say that we place the tumor into a different prognostic group simply because of the stress of the procedure. Fotovati et al. [51] concluded that *NDRG1* could be used as a biomarker for differentiation of breast cancer for both diagnostic and therapeutic purposes. Still it is very important to be aware of at what material the gene is measured upon.

Our study shows the expression profiles found in the analyses of breast cancer tissue must be evaluated with caution. Different profiles may simply be result of differences in the surgical trauma and timing of when samples are taken, and not necessarily associated with tumor biology.

Authors' Contribution

M. L. H Riis, T. Lüders, V. N. Kristensen, and I. R. K. Bukholm contributed equally to the work and should be considered as co-first authors and cosenior authors, respectively.

Acknowledgments

The authors are grateful to Andliena Tahiri, Karin Helmersen, and Helen Larsen from the department of clinical molecular biology for help with the miRNA microarray experiments and to Himanshu Joshi, department of clinical molecular biology, for help with the pathway analysis. They would also like to thank their colleagues at the Department of Radiology and the Department of Surgery, Akershus University Hospital, for help with the specimen collection. This work was supported by Grant of the South-Eastern Norway Regional Health Authority (Helse Sør-Øst) no. 2789119 and of the Akershus University Hospital no. 2679030 and no. 2699015 to V. N. Kristensen.

References

[1] C. M. Townsend, R. D. Beauchamp, B. M. Evers et al., *Sabiston Textbook of Surgery: The Biological Basis of Modern Surgical Practice*, Elsevier Saunders, Philadelphia, Pa, USA, 2004.

[2] U. Veronesi, G. Paganelli, G. Viale et al., "A randomized comparison of sentinel-node biopsy with routine axillary dissection in breast cancer," *The New England Journal of Medicine*, vol. 349, no. 6, pp. 546–553, 2003.

[3] C. M. Perou, T. Sørile, M. B. Eisen et al., "Molecular portraits of human breast tumours," *Nature*, vol. 406, no. 6797, pp. 747–752, 2000.

[4] T. Sørlie, R. Tibshirani, J. Parker et al., "Repeated observation of breast tumor subtypes in independent gene expression data sets," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 14, pp. 8418–8423, 2003.

[5] C. Sotiriou, S. Y. Neo, L. M. McShane et al., "Breast cancer classification and prognosis based on gene expression profiles from a population-based study," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 18, pp. 10393–10398, 2003.

[6] L. J. Van't Veer, H. Dai, M. J. van de Vijver et al., "Gene expression profiling predicts clinical outcome of breast cancer," *Nature*, vol. 415, no. 6871, pp. 530–536, 2002.

[7] H. Zhao, A. Langerød, Y. Ji et al., "Different gene expression patterns in invasive lobular and ductal carcinomas of the breast," *Molecular Biology of the Cell*, vol. 15, no. 6, pp. 2523–2536, 2004.

[8] M. J. van de Vijver, Y. D. He, L. J. Van 'T Veer et al., "A gene-expression signature as a predictor of survival in breast cancer," *The New England Journal of Medicine*, vol. 347, no. 25, pp. 1999–2009, 2002.

[9] Y. Wang, J. G. M. Klijn, Y. Zhang et al., "Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer," *The Lancet*, vol. 365, no. 9460, pp. 671–679, 2005.

[10] B. Weigelt, J. L. Peterse, and L. J. Van't Veer, "Breast cancer metastasis: markers and models," *Nature Reviews Cancer*, vol. 5, no. 8, pp. 591–602, 2005.

[11] V. Wong, D. Y. Wang, K. Warren et al., "The effects of timing of fine needle aspiration biopsies on gene expression profiles in breast cancers," *BMC Cancer*, vol. 8, article 277, 2008.

[12] L. de Cecco, V. Musella, S. Veneroni et al., "Impact of biospecimens handling on biomarker research in breast cancer," *BMC Cancer*, vol. 9, article 409, 2009.

[13] G. Trifirò, G. Viale, O. Gentilini, L. Lavinia Travaini, and G. Paganelli, "Sentinel node detection in pre-operative axillary staging," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 31, supplement 1, pp. S46–S55, 2004.

[14] J. S. Wei and J. Khan, "Purification of total RNA from mammalian cells and tissues," in *DNA Microarrays: A Molecular Cloning Manual*, D. Bowtell and J. Sambrook, Eds., pp. 110–119, Cold Spring Harbor Laboratory Press, New York, NY, USA, 2002.

[15] B. Dysvik and I. Jonassen, "J-Express: exploring gene expression data using Java," *Bioinformatics*, vol. 17, no. 4, pp. 369–370, 2001.

[16] D. W. Huang, B. T. Sherman, and R. A. Lempicki, "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists," *Nucleic Acids Research*, vol. 37, no. 1, pp. 1–13, 2009.

[17] D. W. Huang, B. T. Sherman, and R. A. Lempicki, "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources," *Nature Protocols*, vol. 4, no. 1, pp. 44–57, 2009.

[18] K. J. Livak and T. D. Schmittgen, "Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method," *Methods*, vol. 25, no. 4, pp. 402–408, 2001.

[19] V. G. Tusher, R. Tibshirani, and G. Chu, "Significance analysis of microarrays applied to the ionizing radiation response," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 9, pp. 5116–5121, 2001.

[20] B. Haibe-Kains, C. Desmedt, F. Piette et al., "Comparison of prognostic gene expression signatures for breast cancer," *BMC Genomics*, vol. 9, article 394, 2008.

[21] C. Sotiriou and L. Pusztai, "Gene-expression signatures in breast cancer," *The New England Journal of Medicine*, vol. 360, no. 8, pp. 752–800, 2009.

[22] B. Weigelt, F. L. Baehner, and J. S. Reis-Filho, "The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade," *Journal of Pathology*, vol. 220, no. 2, pp. 263–280, 2010.

[23] F. Cardoso, M. Piccart-Gebhart, L. Van't Veer, and E. Rutgers, "The MINDACT trial: the first prospective clinical validation of a genomic tool," *Molecular Oncology*, vol. 1, no. 3, pp. 246–251, 2007.

[24] S. Paik, "Development and clinical utility of a 21-gene recurrence score prognostic assay in patients with early breast cancer treated with tamoxifen," *Oncologist*, vol. 12, no. 6, pp. 631–635, 2007.

[25] S. F. Ma, D. N. Grigoryev, A. D. Taylor et al., "Bioinformatic identification of novel early stress response genes in rodent models of lung injury," *American Journal of Physiology*, vol. 289, no. 3, pp. L468–L477, 2005.

[26] M. Ruel, C. Bianchi, T. A. Khan et al., "Gene expression profile after cardiopulmonary bypass and cardioplegic arrest," *Journal of Thoracic and Cardiovascular Surgery*, vol. 126, no. 5, pp. 1521–1530, 2003.

[27] G. Tarozzo, M. Campanella, M. Ghiani, A. Bulfone, and M. Beltramo, "Expression of fractalkine and its receptor, CX3CR1, in response to ischaemia-reperfusion brain injury in the rat," *European Journal of Neuroscience*, vol. 15, no. 10, pp. 1663–1668, 2002.

[28] K. Morita, T. Saito, M. Ohta et al., "Expression analysis of psychological stress-associated genes in peripheral blood leukocytes," *Neuroscience Letters*, vol. 381, no. 1-2, pp. 57–62, 2005.

[29] I. Bièche, F. Lerebours, S. Tozlu, M. Espie, M. Marty, and R. Lidereau, "Molecular profiling of inflammatory breast cancer: identification of a poor-prognosis gene expression signature," *Clinical Cancer Research*, vol. 10, no. 20, pp. 6789–6795, 2004.

[30] S. Langer, C. F. Singer, G. Hudelist et al., "Jun and Fos family protein expression in human breast cancer: correlation of protein expression and clinicopathological parameters," *European Journal of Gynaecological Oncology*, vol. 27, no. 4, pp. 345–352, 2006.

[31] C. Lu, Q. Shen, E. DuPré, H. Kim, S. Hilsenbeck, and P. H. Brown, "cFos is critical for MCF-7 breast cancer cell growth," *Oncogene*, vol. 24, no. 43, pp. 6516–6524, 2005.

[32] C. C. Chen and L. F. Lau, "Functions and mechanisms of action of CCN matricellular proteins," *International Journal of Biochemistry and Cell Biology*, vol. 41, no. 4, pp. 771–783, 2009.

[33] A. M. Babic, M. L. Kireeva, T. V. Kolesnikova, and L. F. Lau, "CYR61, a product of a growth factor-inducible immediate early gene, promotes angiogenesis and tumor growth," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 11, pp. 6355–6360, 1998.

[34] D. Sampath, R. C. Winneker, and Z. Zhang, "Cyr61, a member of the CCN family, is required for MCF-7 cell proliferation: regulation by 17β-estradiol and overexpression in human breast cancer," *Endocrinology*, vol. 142, no. 6, pp. 2540–2548, 2001.

[35] M. A. A. K. Folgueira, H. Brentani, D. M. Carraro et al., "Gene expression profile of residual breast cancer after doxorubicin and cyclophosphamide neoadjuvant chemotherapy," *Oncology Reports*, vol. 22, no. 4, pp. 805–813, 2009.

[36] H. Y. Wang, Z. Cheng, and C. C. Malbon, "Overexpression of mitogen-activated protein kinase phosphatases MKP1, MKP2 in human breast cancer," *Cancer Letters*, vol. 191, no. 2, pp. 229–237, 2003.

[37] Y. X. Liu, J. Wang, J. Guo, J. Wu, H. B. Lieberman, and Y. Yin, "DUSP1 is controlled by p53 during the cellular response to oxidative stress," *Molecular Cancer Research*, vol. 6, no. 4, pp. 624–633, 2008.

[38] W. F. Symmans, M. Ayers, E. A. Clark et al., "Total RNA yield and microarray gene expression profiles from fine-needle aspiration biopsy and core-needle biopsy samples of breast carcinoma," *Cancer*, vol. 97, no. 12, pp. 2960–2971, 2003.

[39] M. Benetkiewicz, A. Piotrowski, T. D. de Ståhl et al., "Chromosome 22 array-CGH profiling of breast cancer delimited minimal common regions of genomic imbalances and revealed frequent intra-tumoral genetic heterogeneity," *International Journal of Oncology*, vol. 29, no. 4, pp. 935–945, 2006.

[40] F. C. Geyer, B. Weigelt, R. Natrajan et al., "Molecular analysis reveals a genetic basis for the phenotypic diversity of metaplastic breast carcinomas," *Journal of Pathology*, vol. 220, no. 5, pp. 562–573, 2010.

[41] T. Sørlie, C. M. Perou, R. Tibshirani et al., "Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 19, pp. 10869–10874, 2001.

[42] S. Paik, S. Shak, G. Tang et al., "A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer," *The New England Journal of Medicine*, vol. 351, no. 27, pp. 2817–2826, 2004.

[43] O. Giricz, V. Calvo, S. C. Pero et al., "GRB7 is required for triple-negative breast cancer cell invasion and survival," *Breast Cancer Research and Treatment*, vol. 133, no. 2, pp. 607–615, 2012.

[44] J. A. Sparano, L. J. Goldstein, B. H. Childs et al., "Relationship between quantitative GRB7 RNA expression and recurrence after adjuvant anthracycline chemotherapy in triple-negative breast cancer," *Clinical Cancer Research*, vol. 17, no. 22, pp. 7194–7203, 2011.

[45] B. Margolis, O. Silvennoinen, F. Comoglio et al., "High-efficiency expression/cloning of epidermal growth factor-receptor- binding proteins with Src homology 2 domains," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 19, pp. 8894–8898, 1992.

[46] D. Stein, J. Wu, S. A. W. Fuqua et al., "The SH2 domain protein GRB-7 is co-amplified, overexpressed and in a tight complex with HER2 in breast cancer," *EMBO Journal*, vol. 13, no. 6, pp. 1331–1340, 1994.

[47] K. Yokote, B. Margolis, C. H. Heldin, and L. Claesson-Welsh, "Grb7 is a downstream signaling component of platelet-derived growth factor and β-receptors," *The Journal of Biological Chemistry*, vol. 271, no. 48, pp. 30942–30949, 1996.

[48] H. Cangul, "Hypoxia upregulates the expression of the NDRG1 gene leading to its overexpression in various human cancers," *BMC Genetics*, vol. 5, article 27, 2004.

[49] S. Bandyopadhyay, S. K. Pai, S. Hirota et al., "Role of the putative tumor metastasis suppressor gene Drg-1 in breast cancer progression," *Oncogene*, vol. 23, no. 33, pp. 5675–5681, 2004.

[50] A. Fotovati, T. Fujii, M. Yamaguchi et al., "17β-estradiol induces down-regulation of Cap43/NDRG1/Drg-1, a putative differentiation-related and metastasis suppressor gene, in human breast cancer cells," *Clinical Cancer Research*, vol. 12, no. 10, pp. 3010–3018, 2006.

[51] A. Fotovati, S. Abu-Ali, M. Kage, K. Shirouzu, H. Yamana, and M. Kuwano, "N-myc downstream-regulated gene 1 (NDRG1) a differentiation marker of human breast cancer," *Pathology and Oncology Research*, vol. 17, no. 3, pp. 525–533, 2011.

II

Hindawi Publishing Corporation
ISRN Oncology
Volume 2013, Article ID 924971, 12 pages
http://dx.doi.org/10.1155/2013/924971

Research Article

Gene Expression Profile Analysis of T1 and T2 Breast Cancer Reveals Different Activation Pathways

Margit L. H. Riis,[1,2] Xi Zhao,[3] Fateme Kaveh,[4] Hilde S. Vollan,[2,5] Anne-Jorunn Nesbakken,[6] Hiroko K. Solvang,[4] Torben Lüders,[2,5] Ida R. K. Bukholm,[1,2,7] and Vessela N. Kristensen[2,4,5]

[1] *Department of Surgery, Akershus University Hospital, 1478 Lørenskog, Norway*
[2] *Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, 0318 Oslo, Norway*
[3] *Department of Radiology, School of Medicine, Stanford Center for Cancer Systems Biology, Stanford University, Stanford, CA 94305-5488, USA*
[4] *Department of Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, 0310 Oslo, Norway*
[5] *Department of Clinical Molecular Biology and Laboratory Sciences (EpiGen), Akershus University Hospital, 1478 Lørenskog, Norway*
[6] *Department of Pathology, Akershus University Hospital, 1478 Lørenskog, Norway*
[7] *Institute of Health Promotion, Akershus University Hospital, 1478 Lørenskog, Norway*

Correspondence should be addressed to Ida R. K. Bukholm; i.r.k.bukholm@medisin.uio.no

Received 29 November 2012; Accepted 8 January 2013

Academic Editors: A. Abdollahi, Y. Ionov, and V. Lorusso

Breast cancers today are of predominantly T1 ($0.1 \geq 2.0$ cm) or T2 ($> 2 \leq 5$ cm) categories due to early diagnosis. Molecular profiling using microarrays has led to the notion of breast cancer as a heterogeneous disease both clinically and molecularly. Given the prognostic power and clinical use of tumor size, the purpose of this study was to search for molecular signatures characterizing clinical T1 and T2. In total 46 samples were included in the discovery dataset. After adjusting for hormone receptor status, lymph node status, grade, and tumor subclass 441 genes were differently expressed between T1 and T2 tumors. Focal adhesion and extracellular matrix receptor interaction were upregulated in the smaller tumors while p38MAPK signaling and immune-related pathways were more dominant in the larger tumors. The T-size signature was then tested on a validation set of 947 breast tumor samples. Using the T-size expression signatures instead of tumor size leads to a significant difference in risk for distant metastases ($P < 0.001$). If further confirmed, this molecular signature can be used to select patients with tumor category T1 who may need more aggressive treatment and patients with tumor category T2 who may have less benefit from it.

1. Introduction

Breast cancer is by far the most frequent cancer among women, and ranks second overall [1]. Guidelines for breast cancer treatment are based upon classical clinicopathological parameters: age, tumor size, grade, lymph node status, and histological type; in addition to hormone receptor status [2]. Lymph node (N) status is the most powerful single indicator of breast cancer prognosis [3], while tumor size, categorized into four groups (T1–4) is the second strongest indicator and is independent of lymph node status [3]. Here we attempted to identify the molecular background behind this prognostic effect of tumor size.

Mammographic screening has led to breast cancer diagnosis at preclinical stage and, as a consequence, most diagnosed cases present as T1 or T2, with significantly better survival in T1 tumors [4]. Nevertheless, T1 tumors may also give recurrence or metastases. Chemotherapy and hormonal treatment reduce the risk of recurrence or distant metastases by approximately 30% and according to the current guidelines whether a tumor is T1 or T2 is a critical factor in treatment decision. However, 70–80% of patients would have survived without adjuvant treatment [5]. How to distinguish the patients that would benefit from adjuvant treatment would therefore be of great value to the patient preventing

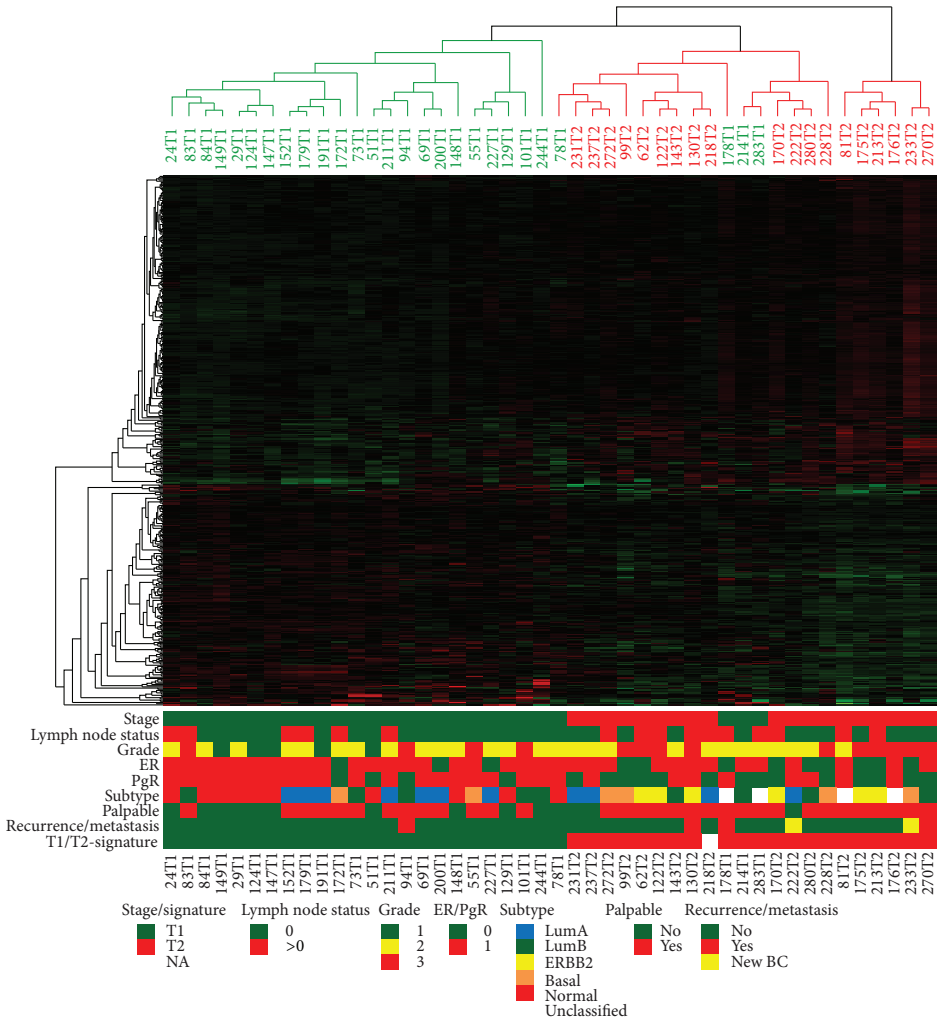


Figure 1: Unsupervised hierarchical clustering using the 441 significant probes after adjustment for clinical parameters. Genes are listed vertically and each patient is represented in the columns. Clinical T1 tumors are shown in green while clinical T2 tumors are shown in red.

possible severe side effects, and to the community saving from unnecessary health expenses.

Microarray technology has enabled to study thousands of genes simultaneously. Interpretation of the data requires advanced statistical analysis [6] and there has been a long way to clinical implication [7]. Hierarchical clustering has been the simplest algorithm applied to organize both genes and samples into groups based on similarity of gene expression [8]. Based on this, breast cancers have been separated into several molecular subclasses [9]. This implies breast cancer as a heterogeneous group of malignancy with distinct molecular signature. The molecular subgroups have been studied in respect of clinical implication and are significantly correlated to overall survival and recurrence of disease [10]. As opposed to this unsupervised approach, the principle of supervised analysis is one where predictive models are built based on existing knowledge of the clinical characteristics [11, 12]. This methodology has also been used to establish a good and a poor prognosis profile which is a more powerful predictor of outcome in young patients with breast cancer than the standard systems based on clinicohistological criteria [10, 13].

Since one of the most crucial factors of treatment decision is tumor size, we aimed to find a gene expression profile which will best characterize the two most common groups of

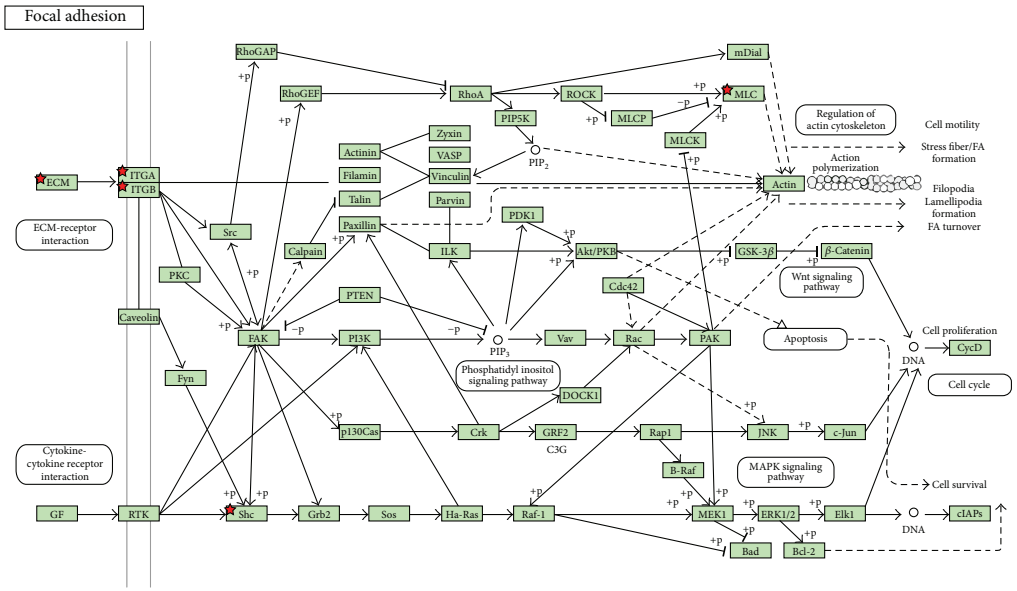


Figure 2: Focal adhesion is one of the pathways upregulated in clinical T1 and downregulated in clinical T2. Red star implies the gene is upregulated in clinical T1 and downregulated in clinical T2.

tumors: T1 and T2. We first identified the most differentially expressed genes between T1 and T2 tumors in 46 patients and characterized the biological pathways active in each category. We then validated this gene list on other publically available datasets.

2. Materials and Methods

2.1. Sample Collection. Core needle biopsies were collected at Akershus University Hospital, Norway, between 2003 and 2007. The tumors were detected clinically or through screening by mammography. The samples were taken under ultrasound guidance and immediately placed into RNA later (Sigma Aldrich, St. Louis, MO, USA). The stabilized samples were stored at −80°C. The study is approved by the Regional Committee for Medical and Health Research Ethics (REK) and all women included have signed a consent form.

This study includes in total 46 samples, 27 T1 and 19 T2 infiltrating ductal carcinomas. T1 lesion is defined as no more than 2 cm in size while T2 lesion is defined as above 2 cm up to 5 cm. The clinical parameters of the tumors are summarized in Table 1 and Figure 1. Within the T1 group two women had recurrence or metastasis, both of these are deceased. In addition two other women in this group are deceased but without cancer specific death. In the T2 group two patients had metastasis or recurrence, one of these has deceased. In addition there was one more case of mortality in this group; this patient developed malignant melanoma with liver metastasis which was the probable cause of death. Among the women in the T2 group two patients developed a new breast cancer.

2.2. RNA Isolation. Frozen biopsies were homogenized in 600 μL Trizol (Invitrogen, Carlsbad, CA, USA) using a 5 mm steal bead (Qiagen, Hilden, Germany) and a Mixer Mill MM301 (Retsch, Haan, Germany) at 20 Hz for 2 min before adding an additional 600 μL Trizol, followed by 240 μL chloroform (Sigma Aldrich). After centrifugation (15 min, 12000 ×g, 4°C) the upper aqueous phase was transferred to a new tube and RNA precipitated by adding an equal volume of isopropanol. After centrifugation the pellet was washed 2-3 times with 75% ethanol and dissolved in 40 μL RNase-free water (Ambion, Austin, TX, USA). Concentration was measured using NanoDrop (Thermo Fisher Scientific, Waltham, MA, USA) and RNA quality assayed on a 2100 Bioanalyzer (Agilent, Santa Clara, CA, USA). The purified RNA was stored at −80°C.

2.3. Microarray Analysis. 10 μg total RNA was amplified using Amino Allyl MessageAmp II aRNA Amplification Kit (Ambion) followed by posttranscriptional labeling with CyDye Cy3 or Cy5 (GE HealthCare, Chalfont St. Giles, UK). As a reference probe universal human reference RNA (UHR; Stratagene, La Jolla, CA, USA) was amplified and labeled as above. Amplification and labeling efficiency were controlled on the NanoDrop. Labeled cRNA corresponding to 20 picomoles cyanine dye each of experimental and reference samples were mixed and hybridized to Agilent

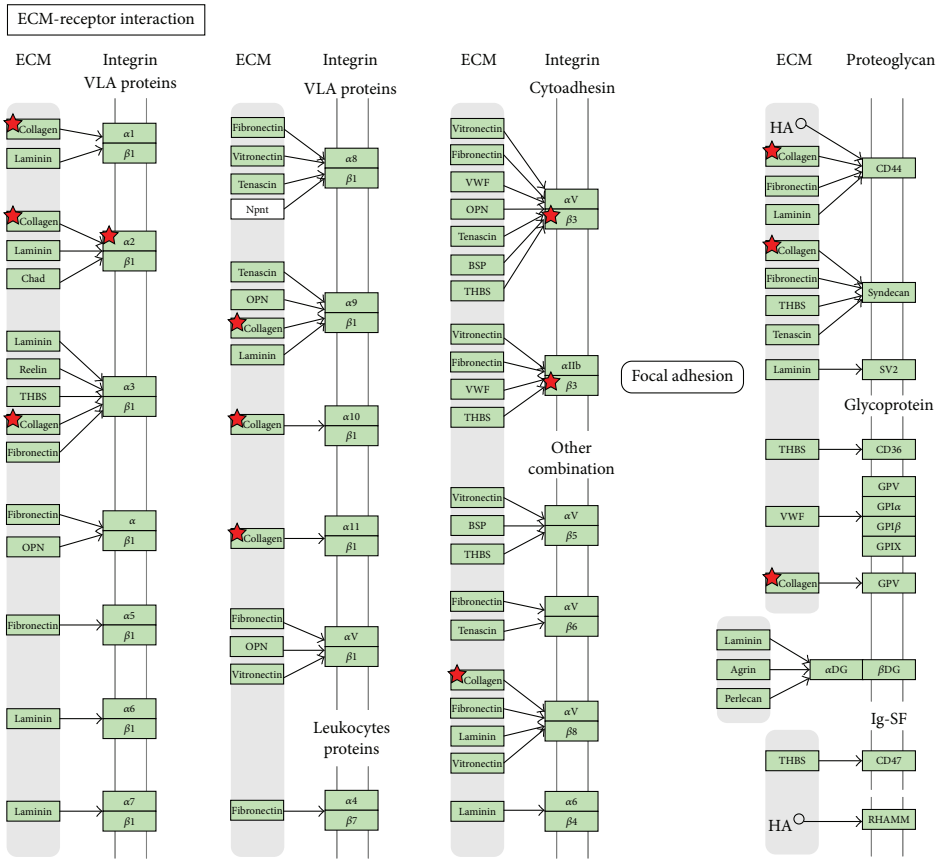


Figure 3: ECM receptor interactions are upregulated in clinical T1 and downregulated in clinical T2. Red star implies the gene is upregulated in clinical T1 and downregulated in clinical T2.

Whole Human Genome Oligo Microarrays (1 × 44 k format) per manufacturer's protocol (Ver. 4.1). After hybridization at 60°C for 17 hours the arrays were washed and scanned using an Agilent scanner.

Data collection and quality assessment were performed using Agilent Feature Extraction software v8.5 with default parameters. Preprocessing was performed using JExpress Pro v2.7 [14]. Poor spots flagged by Feature Extraction were filtered out and Loewess normalization applied. Missing values were calculated with the LSimpute function for genes with less than 5% missing values. To find significant changes of genes/probes between the two tumor stages, Statistical Analysis of Microarray (SAM) [15] was applied. To adjust for lymph node status, differential grade, estrogen and progesterone receptor status, and breast cancer subtype, a partial least squares regression analysis was performed with the pls package in R [16, 17]. To find biological functions pathway analysis was performed for the up- and downregulated genes in T1 and T2 using DAVID [18, 19]. The upregulated genes in T1 are simultaneously the downregulated genes in T2 and vice versa. To confirm and visualize the differential expression between T1 and T2 tumors, unsupervised hierarchical clustering using the genes/probes significantly deregulated by SAM was performed in JExpress Pro. The microarray data have been submitted to the ArrayExpress Archive (http://www.ebi.ac.uk/microarray-as/ae/), accession number: E-MTAB-1049.

2.4. Validation Set. To validate the T-size signature in an independent dataset, we collected expression profiles of 947 breast tumor samples [20] from six published microarray datasets [21–26] with updated followups. The datasets are accessible from NCBI's Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) with the following identifiers; GSE6532 for the Loi dataset [21], GSE3494 for the Miller dataset [22], GSE1456 for the Pawitan dataset [23], GSE7390 for the Desmedt dataset [24], and GSE2603 for

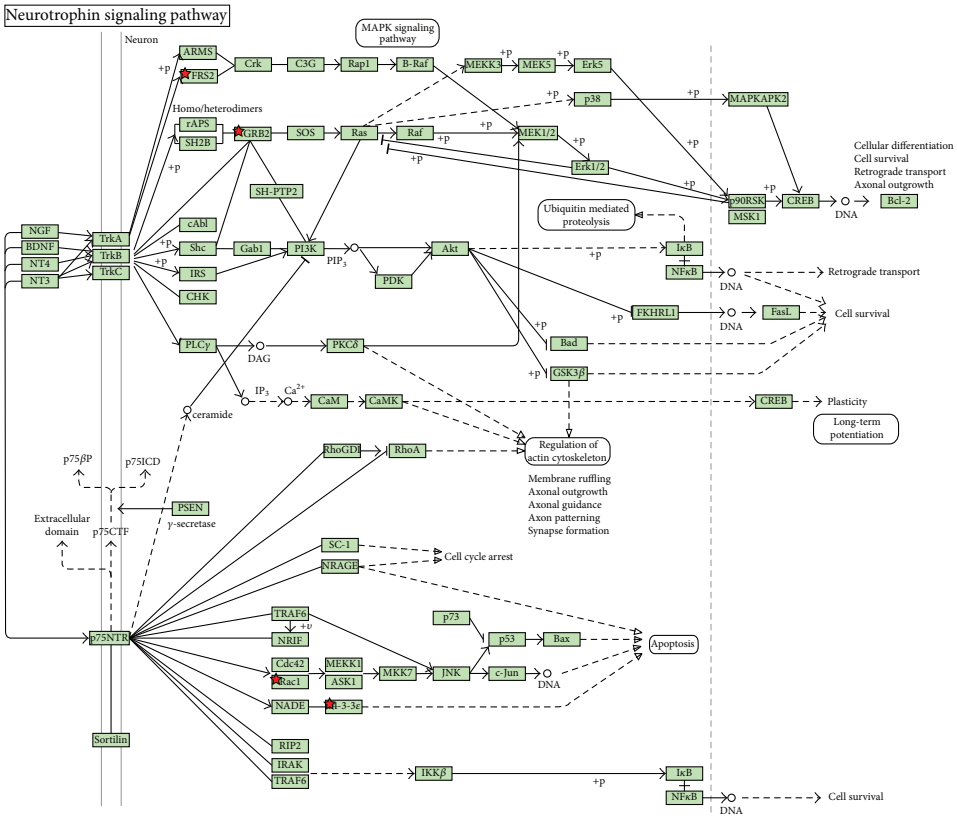


Figure 4: Neurotrophin signaling pathway is among the pathways downregulated in clinical T1 and upregulated in clinical T2. Red stars imply genes that are downregulated in clinical T1 and upregulated in T2.

the Minn dataset [25]. The Chin dataset [26] is available from ArrayExpress (http://www.ebi.ac.uk/) with identifier E-TABM-158.

These datasets were all measured on Human Genome HG-U133A Affymetrix arrays. Each dataset was RMA-normalized [27] and median centered per gene. All overlapping samples from the Desmedt and Loi datasets were excluded. The datasets were then merged based on the common probes. Gene centering has been shown to effectively remove many data set specific biases allowing effective integration of multiple data sets [28]. The merged dataset did not show batch effect after pulling (see Zhao et al. [20]).

2.5. Gene Signatures Evaluated on the Validation Set. For the T-size signature, tumors in the validation set were assigned to either T1-like group or T2-like group using the nearest of the T-size expression centroids (distances computed using correlation to the centroids). The risk group assignment corresponded to the label of the centroid with the highest correlation. We did not apply a correlation cutoff when assigning risk groups; every sample received a classification based on the T-size signature.

We further compared the prognostic power of the T-size signature with eight established prognostic gene signatures for breast cancer. These are Intrinsic [9, 29, 30], PAM50 [31], 70-gene or MammaPrint (Agendia, Amsterdam, The Netherlands) [13, 32], 76-gene [33], Genomic-Grade-Index (GGI) [21, 34], 21-gene-Recurrence-Score (RS) or Oncotype DX (Genomic Health Inc., Redwood City, CA) [35], Wound-Response (WR) signature [36, 37], and Hypoxia signature [38, 39]. All included gene signatures were implemented using the original algorithms. For Intrinsic and PAM50, in addition to subtype classification, a risk score per sample was computed by linear combination of the centroid correlations in ROR-S model (Risk-Of-Relapse scores by Subtype alone) [31]. A pseudo Oncotype DX Recurrence Score per patient was computed by the unscaled Recurrence Score [35]. For 76-genes, GGI and RS, rather than assigning risk groups based on published cutoffs, we used a population-based approach

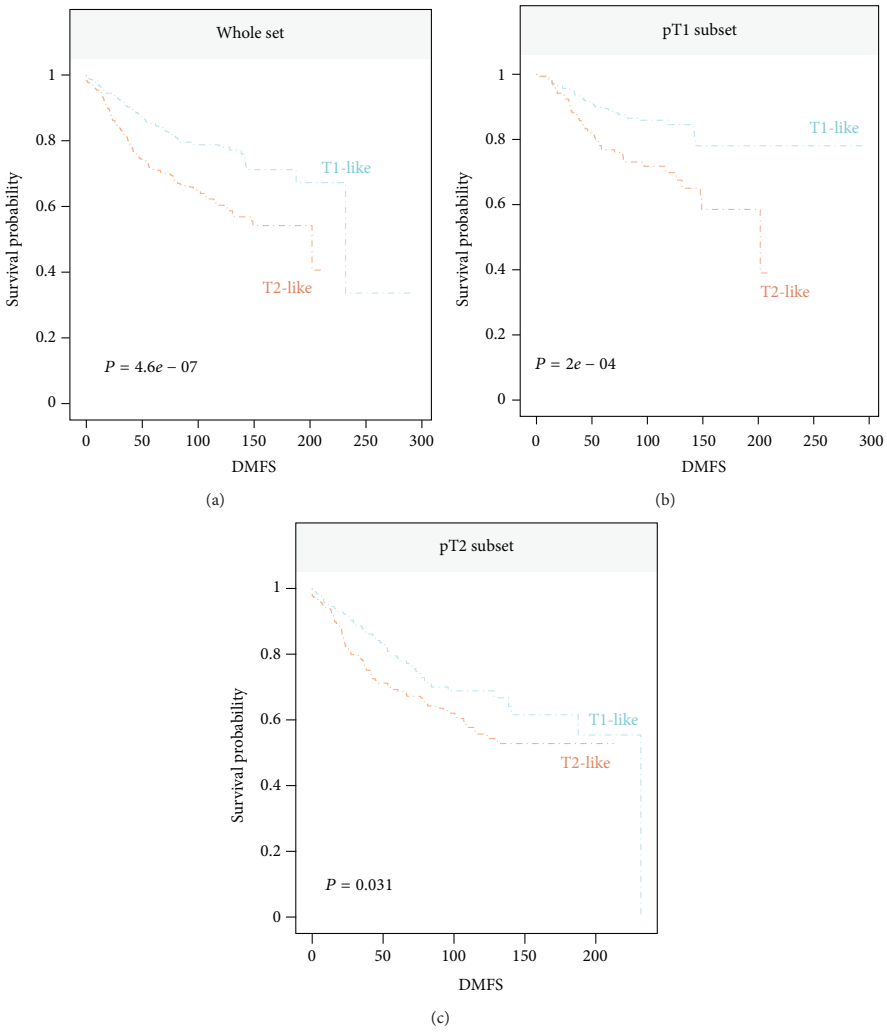


Figure 5: T-size signature for prediction for Distant Metastasis Free Survival (DMFS) on the Affy947 breast cancer dataset. Survival probabilities associated with the risk groups are shown by Kaplan Meier plot. A total of 912 patients had available DMFS status. Follow-up time is shown in month. (a) T-size signature in the complete set. (b) In pT1 tumor subset. (c) In pT2 tumor subset.

in which a fixed proportion of the population was assigned to each risk group. The proportions were derived from previous datasets associated with individual signatures [24, 34, 35]. We found this necessary as our analyses differed from the original methods in technical or methodological manners (see details in Zhao et al. [20]). To make a fair comparison across signatures, we assessed the signatures on the full dataset.

2.6. Survival Analysis. The signatures were evaluated for prediction of Distant Metastasis Free Survival (DMFS). A total of 912 patients on the validation set ($n = 947$) had available DMFS status with median followup for 81 months. The Kaplan-Meier survival curves were plotted for the corresponding risk groups. The differences in survival probabilities associated with the risk groups were tested by a logrank test.

A *likelihood ratio test* was used to assess the significance of the overall effect in a univariate comparison of predictors. *Deviance* was used to check the goodness of the model fit. The marginal contribution by a single predictor in the univariate setting was evaluated using the *proportion of*

Table 1: Summary of patient characteristics.

Sample	Stage	Size/cm	Grade	Node	ER	PgR	Recurrence/metastasis	Deceased
24	T1	0.8	2	1	Positive	Positive		
29	T1	0.8	2	0	Positive	Positive		
51	T1	1.0	1	0	Positive	Negative		
55	T1	0.8	3	0	Positive	Positive		
62	T2	2.5	3	2	Negative	Negative		
69	T1	1.4	2	0	Positive	Positive		
73	T1	1.8	2	0	Positive	Positive		
78	T1	1.2	2	0	Positive	Positive		
81	T2	2.3	2	1	Positive	Positive		
83	T1	1.9	3	1	Positive	Positive		
84	T1	1.0	2	0	Positive	Positive		
94	T1	1.9	3	0	Positive	Negative	Yes	Yes
99	T2	4.0	3	0	Negative	Negative		
101	T1	1.9	3	0	Positive	Positive		Yes
122	T2	3.4	3	1	Positive	Negative		Yes
124	T1	0.9	1	0	Positive	Positive		
129	T1	0.8	2	0	Positive	Negative		
130	T2	2.2	3	2	Positive	Positive	Yes	Yes
143	T2	2.8	2	0	Positive	Positive		
147	T1	1.0	1	0	Positive	Positive		
148	T1	0.9	2	0	Positive	Positive		
149	T1	1.0	1	0	Positive	Positive		
152	T1	1.5	2	1	Positive	Positive		
170	T2	4.0	2	1	Negative	Negative		
172	T1	1.9	2	1	Negative	Negative		Yes
175	T2	2.5	3	0	Negative	Negative		
176	T2	2.7	3	0	Positive	Positive		
178	T1	1.2	2	1	Negative	Positive	Yes	Yes
179	T1	1.8	2	1	Positive	Positive		
191	T1	1.4	1	0	Positive	Positive		
200	T1	1.5	2	0	Negative	Positive		
211	T1	1.7	2	2	Positive	Positive		
213	T2	3.0	3	1	Negative	Negative		
214	T1	1.2	2	0	Positive	Negative		
218	T2	2.1	1-2	1	Positive	Negative		
222	T2	2.1	2	0	Positive	Positive	New BC	
227	T1	1.8	2	0	Negative	Positive		
228	T2	3.0	3	0	Negative	Negative		
231	T2	2.1	2	0	Positive	Negative		
233	T2	3.0	3	0	Negative	Negative	New BC	
237	T2	2.1	2	0	Positive	Positive		
244	T1	0.6	2	0	Positive	Positive		
270	T2	3.0	3	0	Positive	Negative	Yes	
272	T2	2.3	2	1	Positive	Negative		
280	T2	3.2	2	0	Negative	Positive		
283	T1	1.3	2	1	Positive	Negative		

Table 2: Upregulated pathways in T1 breast cancer tumors compared to T2 tumors.

Term	Genes	Count	%	P value	Benjamini
Focal adhesion (KEGG_PATHWAY)	MYL7, ITGA2, ITGB3, COL4A6, SHC4	5	3.5	$3.6E-2$	$8.6E-1$
Arrhythmogenic right ventricular cardiomyopathy (ARVC) (KEGG_PATHWAY)	LEF1, ITGA2, ITGB3	3	2.1	$8.2E-2$	$9.0E-1$
Small cell lung cancer (KEGG_PATHWAY)	ITGA2, ITGB3, COL4A6	3	2.1	$9.7E-2$	$8.4E-1$
ECM-receptor interaction (KEGG_PATHWAY)	CDKN1B, ITGA2, COL4A6	3	2.1	$9.7E-2$	$8.4E-1$

Table 3: Downregulated pathways in T1 breast cancer tumors compared to T2 tumors.

Term	Genes	Count	%	P value	Benjamini
Neurotrophin signaling pathway (KEGG_PATHWAY)	YWHAZ, GRB2, RAC1, YWHAQ, FRS2	5	2.8	$4.2E-2$	$9.8E-1$
p38 MAPK Signaling Pathway (BIOCARTA)	GRB2, RAC1, MKNK1	3	1.7	$5.0E-2$	$9.8E-1$
Prion diseases (KEGG_PATHWAY)	C1QA, C1QB, C1QC	3	1.7	$5.3E-2$	$9.1E-1$
Jak-STAT signaling pathway (KEGG_PATHWAY)	OSM, IFNA2, GRB2, IL10RA, IL4R	5	2.8	$8.1E-2$	$9.2E-1$
Systemic lupus erythematosus (KEGG_PATHWAY)	C1QA, C1QB, HLA-DPB1, C1QC	4	2.2	$8.7E-2$	$8.7E-1$
Toll-like receptor signaling pathway (KEGG_PATHWAY)	IFNA2, MYD88, TICAM1, RAC1	4	2.2	$9.1E-2$	$8.2E-1$

variation explained in the outcome variable (PVE) [40], which is an indicator for the importance of covariates in the Cox model. The *Hazard Ratio* (HR) was used as an accuracy measure for the risk group prediction for different predictors. The *concordance index* (C-index) [41] was computed to assess the predictive discrimination ability of each of the predictors in the corresponding univariate Cox model. For a multivariate comparison of predictors, the relative importance of a covariate in a multivariate Cox model was measured by the partial PVE.

3. Results

After preprocessing 36,669 genes were included for further analyses. Comparing the gene expression profiles between T1 and T2 tumors and using partial least squares regression (PLS) analysis to adjust for lymph node status, differential grade, hormone receptor status, and breast cancer subtype, yielded 441 genes differentially expressed genes at FDR <1% (Supplementary Table S1 available online at http://dx.doi.org/10.1155/2013/924971). Unsupervised hierarchical clustering using these 441 probes resulted in T1 and T2 tumors to cluster for most part separately (Figure 1) except four T1 tumors that clustered with the T2 tumors. One of these patients developed metastasis to the lung and to the bone, and later died. Another one is still alive but had bilateral breast cancer in addition to primary lung cancer. The last two patients in this group are free of recurrence and metastasis.

3.1. Pathway Analysis. To further study the differences in genes between T1 and T2 tumors, we performed pathway analysis. Of the 441 significant probes, 184 probes were upregulated in T1 (downregulated in T2), and 257 probes were downregulated in T1 (upregulated T2). The genes upregulated in T1 were enriched for several pathways (Table 2), including Focal Adhesion (Figure 2) and ECM- (extracellular matrix) receptor interaction (Figure 3). Among the important upregulated genes are several collagens and integrins, and p27 (cyclin dependent kinase inhibitor 1B).

The downregulated genes in T1, upregulated in T2 were enriched for important pathways like Neurotrophin signaling pathway (Figure 4), p38MAPK signaling pathway, and several pathways involved in immune response (Table 3). Important genes in these pathways are MNK1 (MAP kinase interacting serine/threonine-protein kinase 1), GRB2 (Growth factor receptor bound protein 2), RAC1 (ras-related C3 botulinum toxin substrate 1), and several immune-related genes, such as IFN, IL6, MHCII, and complement component 1.

3.2. Validation of the T-Size Signature. In the validation of the T-size signature, a total of 480 samples were called as T1-like, and 467 were classified as T2-like. For all signatures except Hypoxia on the complete set for DMFS ($n = 912$), differences in DMFS between risk groups were highly significant (not shown; see Zhao et al. [20]). Specifically for the T-size signature, the separation between T1-like group and T2-like group was highly significant ($P < 0.001$; Figure 5(a)) with T2-like group associated with higher risk for distant metastasis. We also observed highly significant separation of two risk groups for DMFS in the patient group with pT1 size tumors ($n = 440$; $P < 0.001$; Figure 5(b)); while in the pT2 tumor subgroup ($n = 459$), T-size signature achieved less significant separation for the risk prediction ($P = 0.031$; Figure 5(c)).

We performed univariate analysis for the T-size signature and clinical parameters including tumor size (1–3), node status (positive versus negative), ER status (positive versus negative), and histological grade (1–3), respectively. The performance comparisons by using the likelihood ratio test, the deviance, the *proportion of variation explained* (PVE), the *concordance index* (C-index), and the *Hazard Ratio* (HR) are summarized in Table 4. A multivariate Cox model was used to simultaneously assess the T-size signature and the included clinical parameters in the study. Due to the known association

Table 4: Univariate comparison of predictors.

Covariate	HR [95% CI]	P	PVE	Deviance	C
T-size signature					
(Overall effect)		$4.30E-07$	$2.76E-02$	25.55	0.58
T2-like (versus T1-like)	1.92 [1.48–2.48]	$7.22E-07$			
Tumor size					
(Overall effect)		$5.63E-08$	$3.62E-02$	33.38	0.60
2 (versus 1)	1.95 [1.50–2.55]	$7.83E-07$			
3 (versus 1)	3.39 [1.96–5.88]	$1.37E-05$			
Node					
(Overall effect)		$2.40E-06$	$2.46E-02$	22.24	0.58
+ (versus −)	1.89 [1.46–2.45]	$1.35E-06$			
ER					
(Overall effect)		$2.07E-02$	$5.85E-03$	5.35	0.54
+ (versus −)	0.72 [0.55–0.94]	$1.78E-02$			
Histological grade					
(Overall effect)		$2.45E-04$	$2.11E-02$	16.63	0.60
2 (versus 1)	1.78 [1.15–2.77]	$1.04E-02$			
3 (versus 1)	2.37 [1.52–3.69]	$1.78E-02$			

Table 5: Multivariate comparison of predictors.

Covariate	HR [95% CI]	P	Partial PVE
T-size signature			
T2-like (versus T1-like)	1.70 [1.25–2.32]	$8.03E-04$	$1.42E-02$
Tumor size			
2 (versus 1)	1.74 [1.29–2.35]	$2.72E-04$	$1.62E-02$
3 (versus 1)	2.07 [0.97–4.40]	$5.98E-02$	$1.62E-02$
Node			
+ (versus −)	1.68 [1.24–2.28]	$8.11E-04$	$1.25E-02$
Histological grade			
2 (versus 1)	1.45 [0.92–2.29]	$1.14E-01$	$2.05E-02$
3 (versus 1)	1.49 [0.91–2.47]	$1.16E-01$	$2.05E-02$

between ER status and survival, we included ER status as stratification variable (Table 5).

4. Discussion

Approximately 15% of all women diagnosed with breast cancer die from their disease within 5 years of diagnosis [42] despite having been treated according to national clinical guidelines [2]. Both genomic and clinical variables should be induced in a common algorithm to yield the most accurate prediction model. Microarray has made it possible to study thousands of genes simultaneously. This generates information about gene expression profiles that can be computed in different ways. One of these is the clustering of patients according to the gene expression in their tumors. The majority of the gene lists are generated to distinguish patients from being subject to unnecessary adjuvant treatment or with the intention of individualizing therapy and treatment.

Several genetic signatures have been presented [13, 23, 32–34]. This work has led to the development of special kits such as MammaPrint (Agendia, Amsterdam, The Netherlands) [13, 32] and Oncotype DX (Genomic Health Inc., Redwood City, CA, USA) [35]. By combining information from multiple gene signatures, one would potentially increase the prediction power and bring out an overall picture of this disease. Zhao et al. aimed to develop an analytical framework that allows us to utilize the combined strength from individual gene signatures [43]. Such a framework and the resulting model will be broadly applicable for survival prediction across heterogeneous tumor groups capturing a broad spectrum of biological aspects. The tumor size associated signature presented here has the purpose to identify the molecular characteristics associated to size and does not claim to provide prognostic index superior to the existing ones. The signature specific difference in DMFS within the T1 subgroup and the T2 subgroup, shown here, are used only to suggest that it can be used as supplementary information to tumor size.

Most first generation signatures are good for predicting prognosis in early stage breast cancer. There is only a minor overlap in genes in the different signatures [44], but they produce similar risk group assignment in the same dataset.

Proliferation and the level of proliferation-related genes are the strongest prognostic factors in ER positive cancer. Proliferation-related genes are often highly expressed in ER negative cancers, so in the first generation signatures almost all ER negative cancers seem to have poor prognosis. It was initially meant that these prognostic signatures could replace the classical histopathological findings, but meta-analysis has revealed that tumor size and lymph node status give prognostic information independent of the molecular signatures [45].

The present study attempts to identify, independent of grade, receptor status and lymph node status, the molecular signature, and the underlying biological pathways associated to tumor size, which is an objective property without possibility of interobserver disagreement. The most significant pathways upregulated in T1 compared to T2 are focal adhesion, ECM-receptor interaction, and two organ specific pathways (Table 2). Important genes occur at several steps in these pathways. One of these genes being P27(Kip) (cyclin dependent kinase inhibitor 1B). The cell-cycle regulating protein $p27^{Kip1}$ (p27) has dual roles by acting as both a cdk inhibitor and as an assembly factor for different cdk complexes. Loss of p27 has been linked to malignant features in different tumors [46]. High levels of p27 are expressed in normal human mammary epithelium, but loss of p27 is frequent in breast cancer and has been demonstrated to have prognostic implications [47]. Patients with tumors expressing low levels of p27 were associated with poor prognosis, and it is especially pronounced in hormone-receptor positive tumors [48]. HER2 positive primary breast cancers often reveal low levels of p27 [49]. As mentioned, in our material p27 is upregulated in T1 tumors compared to T2 tumors and this is in coherence with earlier studies. Thus this could be a possible marker, among others, that could be used to select the T1 tumors that have a greater possibility of recurrence. The lower p27, the worse prognosis, consequently requiring stronger treatment.

Pathways downregulated in T1 and upregulated in T2 are shown in Table 3. These are all pathways associated with the immune response, and a majority of the actual downregulated genes are immune response related genes, like IFN, IL6, MHC II, and Complement component 1. This is consistent with a more aggressive lesion that requires more effort from the immune system. Among the genes downregulated in T1 tumors compared to T2 tumors is GRB2. Grb2 is an adaptor protein that is essential for a variety of cellular functions and acts as a critical downstream intermediary in several oncogenic signaling pathways [50]. In human breast cancer cells Grb2 is overexpressed. In an unpublished work we have demonstrated that there is a significant difference in the expression of this gene in normal tissue and breast cancer tissue, and also in normal tissue adjacent to tumor. The role of Grb2 as a signal transducer for several oncogenic growth factor receptors and the broad involvement of Grb2 in multiple steps of the metastasis cascade make it a good target for antitumor therapeutic strategies [50]. Like for p27, maybe this gene could be measured in the patients with smaller tumors to select those with worse prognosis.

RAC1 (ras-related C3 botulinum toxin substrate 1) may represent an attractive target. Rac GTPases, small G-proteins widely implicated in tumorigenesis and metastasis, transduce signals from tyrosine-kinase, G-protein-coupled receptors (GPCRs), and integrins, and control a number of essential cellular functions including motility, adhesion, and proliferation. In breast cancer cells Rac1 is a downstream effector of ErbB receptors and mediates migratory responses by ErbB1/EGFR ligands such as EGF or TGFα and ErbB3 ligands such as heregulins [51]. This gene is a potential target for use in therapy of breast cancer.

5. Conclusions

In summary we show here that there is a molecular profile that is associated to tumor size. Thus a gene-expression signature-based approach combined with the classical TNM classification as well as analysis of key genes may pave the way to improved individualized therapy.

Conflict of Interests

The authors declare that they have no conflict of interests.

Authors' Contribution

X. Zhao, F. Kaveh, I. R. K. Bukholm, and V. N. Kristensen contributed equally to the work and should be considered as second authors and cosenior authors, respectively.

Acknowledgments

The authors are grateful to our colleagues at the Department of Radiology and the Department of Surgery, Akershus University Hospital for their help with the specimen collection. They also would like to thank the Department of Pathology for immunohistochemistry help. This work was supported by Grant of the South-Eastern Norway Regional Health Authority (Helse Sør-Øst) no. 2789119 and of the Akershus University Hospital no. 2679030 and no. 2699015 to V. N. Kristensen.

References

[1] J. Ferlay, H. Shin, F. Bray et al., "GLOBOCAN 2008 v1.2, cancer incidence and mortality worldwide," in *IARC CancerBase*, 2012, http://globocan.iarc.fr.

[2] A. Goldhirsch, J. H. Glick, R. D. Gelber et al., "Meeting highlights: international expert consensus on the primary therapy of early breast cancer 2005," *Annals of Oncology*, vol. 16, no. 10, pp. 1569–1583, 2005.

[3] S. Ciatto, S. Cecchini, A. Iossa, and G. Grazzini, ""T' category and operable breast cancer prognosis," *Tumori*, vol. 75, no. 1, pp. 18–22, 1989.

[4] A. H. Olsen, S. H. Njor, and E. Lynge, "Estimating the benefits of mammography screening: the impact of study design," *Epidemiology*, vol. 18, no. 4, pp. 487–492, 2007.

[5] 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, no. 9472, pp. 1687–1717, 2005.

[6] D. B. Allison, X. Cui, G. P. Page, and M. Sabripour, "Microarray data analysis: from disarray to consolidation and consensus," *Nature Reviews Genetics*, vol. 7, no. 1, pp. 55–65, 2006.

[7] L. D. Miller and E. T. Liu, "Expression genomics in breast cancer research: microarrays at the crossroads of biology and medicine," *Breast Cancer Research*, vol. 9, no. 2, article 206, 2007.

[8] M. B. Eisen, P. T. Spellman, P. O. Brown, and D. Botstein, "Cluster analysis and display of genome-wide expression patterns," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 25, pp. 14863–14868, 1998.

[9] C. M. Perou, T. Sørlie, M. B. Eisen et al., "Molecular portraits of human breast tumours," *Nature*, vol. 406, no. 6797, pp. 747–752, 2000.

[10] L. J. Van't Veer, H. Dai, M. J. van de Vijver et al., "Gene expression profiling predicts clinical outcome of breast cancer," *Nature*, vol. 415, no. 6871, pp. 530–536, 2002.

[11] R. Tibshirani, T. Hastie, B. Narasimhan, and G. Chu, "Diagnosis of multiple cancer types by shrunken centroids of gene expression," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 10, pp. 6567–6572, 2002.

[12] T. Sørlie, "Molecular portraits of breast cancer: tumour subtypes as distinct disease entities," *European Journal of Cancer*, vol. 40, no. 18, pp. 2667–2675, 2004.

[13] M. J. van de Vijver, Y. D. He, L. J. van't Veer et al., "A gene-expression signature as a predictor of survival in breast cancer," *The New England Journal of Medicine*, vol. 347, no. 25, pp. 1999–2009, 2002.

[14] B. Dysvik and I. Jonassen, "J-Express: exploring gene expression data using Java," *Bioinformatics*, vol. 17, no. 4, pp. 369–370, 2001.

[15] V. G. Tusher, R. Tibshirani, and G. Chu, "Significance analysis of microarrays applied to the ionizing radiation response," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 9, pp. 5116–5121, 2001.

[16] B. H. Mevik and R. Wehrens, "The pls package: principal component and partial least squares regression in R," *Journal of Statistical Software*, vol. 18, no. 2, pp. 1–23, 2007.

[17] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2010.

[18] D. W. Huang, B. T. Sherman, Q. Tan et al., "The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists," *Genome Biology*, vol. 8, no. 9, article R183, 2007.

[19] D. W. Huang, B. T. Sherman, and R. A. Lempicki, "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources," *Nature Protocols*, vol. 4, no. 1, pp. 44–57, 2009.

[20] X. Zhao, E. A. Rødeland, T. Sørlie, and H. G. Russnes, "Systematic assessment of prognostic gene signatures for breast cancer shows distinct influence of time and ER status," *under review*.

[21] S. Loi, B. Haibe-Kains, C. Desmedt et al., "Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade," *Journal of Clinical Oncology*, vol. 25, no. 10, pp. 1239–1246, 2007.

[22] L. D. Miller, J. Smeds, J. George et al., "An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 38, pp. 13550–13555, 2005.

[23] Y. Pawitan, J. Bjohle, L. Amler et al., "Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts," *Breast Cancer Research*, vol. 7, no. 6, pp. R953–R964, 2005.

[24] C. Desmedt, F. Piette, S. Loi et al., "Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series," *Clinical Cancer Research*, vol. 13, no. 11, pp. 3207–3214, 2007.

[25] A. J. Minn, G. P. Gupta, P. M. Siegel et al., "Genes that mediate breast cancer metastasis to lung," *Nature*, vol. 436, no. 7050, pp. 518–524, 2005.

[26] K. Chin, S. DeVries, J. Fridlyand et al., "Genomic and transcriptional aberrations linked to breast cancer pathophysiologies," *Cancer Cell*, vol. 10, no. 6, pp. 529–541, 2006.

[27] R. A. Irizarry, B. M. Bolstad, F. Collin, L. M. Cope, B. Hobbs, and T. P. Speed, "Summaries of Affymetrix GeneChip probe level data," *Nucleic Acids Research*, vol. 31, no. 4, article e15, 2003.

[28] A. H. Sims, G. J. Smethurst, Y. Hey et al., "The removal of multiplicative, systematic bias allows integration of breast cancer gene expression datasets—improving meta-analysis and prediction of prognosis," *BMC Medical Genomics*, vol. 1, no. 1, article 42, 2008.

[29] C. M. Perou, S. S. Jeffrey, M. van de Rijn et al., "Distinctive gene expression patterns in human mammary epithelial cells and breast cancers," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 96, no. 16, pp. 9212–9217, 1999.

[30] T. Sørlie, C. M. Perou, R. Tibshirani et al., "Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 19, pp. 10869–10874, 2001.

[31] P. S. Bernard, J. S. Parker, M. Mullins et al., "Supervised risk predictor of breast cancer based on intrinsic subtypes," *Journal of Clinical Oncology*, vol. 27, no. 8, pp. 1160–1167, 2009.

[32] M. Buyse, S. Loi, L. van't Veer et al., "Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer," *Journal of the National Cancer Institute*, vol. 98, no. 17, pp. 1183–1192, 2006.

[33] Y. Wang, J. G. M. Klijn, Y. Zhang et al., "Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer," *The Lancet*, vol. 365, no. 9460, pp. 671–679, 2005.

[34] C. Sotiriou, P. Wirapati, S. Loi et al., "Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis," *Journal of the National Cancer Institute*, vol. 98, no. 4, pp. 262–272, 2006.

[35] S. Paik, S. Shak, G. Tang et al., "A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer," *The New England Journal of Medicine*, vol. 351, no. 27, pp. 2817–2826, 2004.

[36] H. Y. Chang, D. S. A. Nuyten, J. B. Sneddon et al., "Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 10, pp. 3738–3743, 2005.

[37] H. Y. Chang, J. B. Sneddon, A. A. Alizadeh et al., "Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds," *PLoS Biology*, vol. 2, no. 2, article E7, 2004.

[38] J. T. Chi, Z. Wang, D. S. A. Nuyten et al., "Gene expression programs in response to hypoxia: cell type specificity and

prognostic significance in human cancers," *PLoS Medicine*, vol. 3, no. 3, article e47, 2006.

[39] D. S. A. Nuyten, T. Hastie, J. T. A. Chi, H. Y. Chang, and M. J. van de Vijver, "Combining biological gene expression signatures in predicting outcome in breast cancer: an alternative to supervised classification," *European Journal of Cancer*, vol. 44, no. 15, pp. 2319–2329, 2008.

[40] M. Schemper, "The relative importance of prognostic factors in studies of survival," *Statistics in Medicine*, vol. 12, no. 24, pp. 2377–2382, 1993.

[41] F. Harrell, K. Lee, and D. Mark, "Tutorial in biostatistics multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors," *Statistics in Medicine*, vol. 15, pp. 361–387, 1996.

[42] H. Brenner, "Long-term survival rates of cancer patients achieved by the end of the 20th century: a period analysis," *The Lancet*, vol. 360, no. 9340, pp. 1131–1135, 2002.

[43] X. Zhao, T. Sørlie, B. Naume et al., "Combining gene signatures improves prediction of breast cancer survival," *PLoS ONE*, vol. 6, no. 3, Article ID e17845, 2011.

[44] C. Fan, D. S. Oh, L. Wessels et al., "Concordance among gene-expression-based predictors for breast cancer," *The New England Journal of Medicine*, vol. 355, no. 6, pp. 560–569, 2006.

[45] P. Wirapati, C. Sotiriou, S. Kunkel et al., "Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures," *Breast Cancer Research*, vol. 10, no. 4, article R65, 2008.

[46] M. Stendahl, S. Nilsson, C. Wigerup et al., "P27Kip1 is a predictive factor for tamoxifen treatment response but not a prognostic marker in premenopausal breast cancer patients," *International Journal of Cancer*, vol. 127, no. 12, pp. 2851–2858, 2010.

[47] I. M. Chu, L. Hengst, and J. M. Slingerland, "The Cdk inhibitor p27 in human cancer: prognostic potential and relevance to anticancer therapy," *Nature Reviews Cancer*, vol. 8, no. 4, pp. 253–267, 2008.

[48] P. L. Porter, W. E. Barlow, I. T. Yeh et al., "p27Kip1 and cyclin E expression and breast cancer survival after treatment with adjuvant chemotherapy," *Journal of the National Cancer Institute*, vol. 98, no. 23, pp. 1723–1731, 2006.

[49] L. Newman, W. Xia, H. Y. Yang et al., "Correlation of p27 protein expression with HER-2/neu expression in breast cancer," *Molecular Carcinogenesis*, vol. 30, no. 3, pp. 169–175, 2001.

[50] A. Giubellino, T. R. Burke, and D. P. Bottaro, "Grb2 signaling in cell motility and cancer," *Expert Opinion on Therapeutic Targets*, vol. 12, no. 8, pp. 1021–1033, 2008.

[51] E. Wertheimer, A. Gutierrez-Uzquiza, C. Rosemblit, C. Lopez-Haber, M. S. Sosa, and M. G. Kazanietz, "Rac signaling in breast cancer: a tale of GEFs and GAPs," *Cellular Signalling*, vol. 24, no. 2, pp. 353–362, 2012.

III

Riis et al. BMC Cancer 2010, 10:686
http://www.biomedcentral.com/1471-2407/10/686

BMC Cancer

RESEARCH ARTICLE Open Access

Expression of BMI-1 and Mel-18 in breast tissue - a diagnostic marker in patients with breast cancer

Margit LH Riis[1,2], Torben Lüders[2], Anne-Jorunn Nesbakken[3], Hilde S Vollan[2], Vessela Kristensen[2*†], Ida RK Bukholm[1,4†]

Abstract

Background: Polycomb Group (PcG) proteins are epigenetic silencers involved in maintaining cellular identity, and their deregulation can result in cancer. Expression of Mel-18 and Bmi-1 has been studied in tumor tissue, but not in adjacent non-cancerous breast epithelium. Our study compares the expression of the two genes in normal breast epithelium of cancer patients and relates it to the level of expression in the corresponding tumors as well as in breast epithelium of healthy women.

Methods: A total of 79 tumors, of which 71 malignant tumors of the breast, 6 fibroadenomas, and 2 DCIS were studied and compared to the reduction mammoplastic specimens of 11 healthy women. In addition there was available adjacent cancer free tissue for 23 of the malignant tumors. The tissue samples were stored in RNAlater, RNA was isolated to create expression microarray profile. These two genes were then studied more closely first on mRNA transcription level by microarrays (Agilent 44 K) and quantitative RT-PCR (TaqMan) and then on protein expression level using immunohistochemistry.

Results: Bmi-1 mRNA is significantly up-regulated in adjacent normal breast tissue in breast cancer patients compared to normal breast tissue from noncancerous patients. Conversely, mRNA transcription level of Mel-18 is lower in normal breast from patients operated for breast cancer compared to breast tissue from mammoplasty. When protein expression of these two genes was evaluated, we observed that most of the epithelial cells were positive for Bmi-1 in both groups of tissue samples, although the expression intensity was stronger in normal tissue from cancer patients compared to mammoplasty tissue samples. Protein expression of Mel-18 showed inversely stronger intensity in tissue samples from mammoplasty compared to normal breast tissue from patients operated for breast cancer.

Conclusion: Bmi-1 mRNA level is consistently increased and Mel-18 mRNA level is consistently decreased in adjacent normal breast tissue of cancer patients as compared to normal breast tissue in women having had reduction mammoplasties. Bmi-1/Mel-18 ratio can be potentially used as a tool for stratifying women at risk of developing malignancy.

Background

Breast cancer is the leading cause of cancer mortality in women [1]. The prognosis of breast cancer is dependent on stage at the diagnosis, tumors diagnosed at early stage having better prognosis. It is therefore important to detect breast tumors at as early stage as possible [2,3]. Benign diseases, like fibroadenomas, in the mammary gland are associated with increased risk of breast cancer in the same women, although in clinical practice it is difficult to recognize women with fibroadenomas who are at risk of developing breast cancer. There is a need for diagnostic tools which may help stratifying the risk for women with benign changes in their breasts. Bmi-1 and Mel-18 may be genes used for this purpose as our study will show.

There is increasing evidence that breast cancers arise from deregulation of normal pathways in stem or early progenitor cells due to mutations or epigenetic silencing [4,5]. Polycomb Group (PcG) proteins are epigenetic silencer genes involved in maintaining cellular identity, and their deregulation can result in cancer [5].

* Correspondence: vessela.kristensen@medisin.uio.no
† Contributed equally
[2]Department of Clinical Molecular Biology (EpiGen), Institute of Clinical Medicine, Akershus University Hospital, University of Oslo, Lørenskog, Norway
Full list of author information is available at the end of the article

The deregulation of PcG genes may be one of the first events in neoplasia of breast epithelium. Bmi-1 and Mel-18 are two members of the PcG family. Bmi-1 has been shown to maintain the stem cell pool by preventing premature senescence [6,7]. Bmi-1 expression has been detected in normal mammary epithelium and myoepithelial cells and later studies have implied Bmi-1 in connection with self-renewal of stem cells in breast tissue [8]. Bmi-1 has been shown to be up-regulated in breast tumors [7] as well as in several other tumor types [7,9-12]. Over-expression induces lymphomas [11,13,14]. Bmi-1 has shown to regulate cellular senescence and proliferation in rodent and human fibroblasts [15]. After a finite number of cell division, most human cells undergo cellular senescence, whereby cells irreversibly cease to divide [16,17]. Bmi-1 can also bypass senescence and immortalize human mammary epithelial cells [18]. Bmi-1 has been studied in plasma of breast cancer cells with healthy women as controls and results show that levels of Bmi-1 expression may be a surrogate marker of poor prognosis [19]. This may be a very useful noninvasive prognostic marker.

Mel-18 regulates cell proliferation and senescence via transcriptional repression of Bmi-1 and c-myc oncoproteins [17], and is considered to play a dual role, being either oncogenic in some tumor types or acting as a tumor suppressor gene in others. In breast cancer, Mel-18 is supposed to play a tumor suppressor role [20]. Mel-18 was originally cloned from B16 mouse melanoma cells and was shown to be highly expressed in many tumor cells including human melanoma and Hodgkin's lymphomas [21,22].

It has been of our interest to investigate to what extent normal breast tissue from breast cancer patients is actually normal and to what it reflects cancer-specific deregulation. Since both Mel-18 and Bmi-1 may play a role in renewal of stem cells, deregulation of these proteins may be one of the initial steps in development of neoplasia and may be present even in non-cancerous tissue adjacent to the tumor tissue. In the present study, we wanted to analyze whether expression of these two genes, both at the mRNA and protein level, differ between normal tissue taken from breast cancer patients compared to breast tissue from women who had no actual or previous history of any kind of malignancy in the breast.

Methods

Tissue Collection

Tissue samples from malignant tumors and normal counterpart were obtained from patients operated for breast cancer at Akershus University Hospital in the period 2003-2009. In addition to the operated tumor, in some cases large core needle biopsies were taken preoperatively at the radiologist department either when women came for screening or for diagnostic mammography due to palpable mass. Both tumors and biopsies were evaluated by a pathologist to confirm the diagnosis and estimate the tumor cell content. All tumor samples used in this study contained at least 60% tumor cells. Tissue samples were immediately stabilized in RNAlater, and then stored at -80°C. The women have signed a written consent to participate in the study. The study was approved by the Regional Committee for Medical and Health Research Ethics (REK). A total numbers of 71 samples from tumors were analyzed. Mostly invasive ductal carcinomas (64 samples) were included in the study. The remaining were either lobular (5 samples) or mixed (2 samples). Mean age at the operation was 60 years, median 60 years (range 34-83). 30 were classified as lymph node positive and 41 as lymph node negative. 55 were estrogen positive and 43 samples were positive for progesterone. Her-2 gene was found to be amplified in 4 samples (only 21 of the samples were tested for Her-2 gene amplification). Samples from 2 DCIS lesions as well as samples from 6 fibroadenomas were also included in the study. The clinical data are summarized in tables 1 and 2.

Tissue samples from non cancer controls (reduction mammoplasty) have been gathered since the autumn of 2008 and stabilized in the same way as the other samples. Mean age of these women was 39 years, median 36 years (range 20-68). Only two of these women were postmenopausal. Our intension was to have a cohort of healthy women to compare our results. These women also signed a written consent and their names and identification number were registered in our databases.

RNA isolation

The surgical specimens and the large core needle biopsies were macroscopically dissected to obtain a sample suitable for further processing. Frozen tissue was homogenized in Trizol (Invitrogen) with a 5 mm steal bead (Qiagen) using a Mixer Mill MM301 (Retsch) at 30 Hz for 2 min. RNA was isolated following the manufacturer's protocol or the protocol of Wei and Kahn [23]. Purified RNA was dissolved in RNase-free water (Ambion). Concentration was measured using NanoDrop and RNA quality assayed on an Agilent 2100 BioAnalyzer. The purified RNA was stored at -80°C.

Microarray Analysis and Statistical Analysis

500-1000 ng isolated RNA was converted to cDNA with reverse transcriptase and an oligo(dT) primer bearing a T7 promotor followed by in vitro transcription with T7 RNA polymerase to create amplified antisense RNA. The amplified RNA was labeled with Cy3 or Cy5. As a reference probe universal human reference RNA (UHR;

Table 1 Clinical data for the tumor patients

Histology	
Ductal	64
Lobular	5
Mixed	2
Palpable	
Yes	54
No	17
Grade	
1	8
2	40
3	23
Lymph node status	
N0	44
N1	19
N2	6
N3	2
Estrogen receptor status	
ER+	56
ER-	15
Progesterone receptor status	
PR+	41
PR-	30
HER2 receptor status	
HER2+	4
HER2-	17
Unknown	50
Recurrence/Metastasis	1
Death (not cancer specific)	2

Stratagene) was amplified and labeled as above. Amplification and labeling efficiency was controlled using a NanoDrop.

Labeled cRNA was hybridized to Agilent Whole Human Genome Oligo Microarrays per the manufacturer's protocol. After hybridization for 17 hours the arrays were washed and scanned using an Agilent scanner and microarray data extracted with Agilent Feature Extraction software. Preprocessing of the microarray data was done in J-Express Pro http://www.molmine.org while between-array quantile normalization was done in BioConductor [24]. The microarray data are submitted to The ArrayExpress Archive http://www.ebi.ac.uk/microarray-as/ae/ accession number E-MTAB-271.

Table 2 Age distribution for the patients within the different groups

	Cancers	Non cancer controls	DCIS	Fibroadenomas
Mean Age	60	39	63	42
Median Age	60	36	63	45
Range	34-83	20-68	60-66	26-52

Further statistical analysis was done in R. Between group comparison were done using Student's two-sided, two-class t-test and ANOVA using the function *aov*.

Quantitative Real-time PCR (qPCR)

Total RNA was reverse transcribed using the High Capacity RNA to-cDNA Master Mix (Applied Biosystems) and cDNA was diluted with high molecular grade water and stored at -20°C.

For qPCR, 25 ng cDNA, primer/probe sets and TaqMan Gene Expression Master Mix (2×) (Applied Biosystems) were pipetted on a MicroAmp Optical 96-Well Reaction Plate (Applied Biosystems) using the epMotion 5075 pipetting robot (Eppendorf). All samples were pipetted in triplicates and a no template control was run on each plate. The plate was run on the ABI PRISM 7900 HT Fast Real-Time PCR system (Applied Biosystems) with the thermal profile: 50°C for 2 min, 95°C for 10 min and 50 cycles at 95°C for 15 seconds and 60°C for 1 min. Analysis were done using ABI Prism SDS2.3 software and the RQ Manager 1.2 (Applied Biosystems). As for the microarray studies the UHR RNA was used as calibrator.

TaqMan® Gene Expression Assays from Applied Biosystems (BMI1: Hs00180411_m1, PCGF2: Hs00810639_m1, MRPL19: Hs00608519_m1, PPIA: Hs99999904_m1) were used to perform qPCR. All gene assays target exon-exon junctions to be mRNA specific. The final concentration of the TaqMan gene expression assay used was 900 nM for each primer and 250 nM for each probe.

Histology

The slides were evaluated by an experienced pathologist (AJN) and graded according to the Nottingham grading system (Nottingham modification of the Bloom-Richardson system) [25].

Immunohistochemistry

Immunohistochemistry was performed on 5 μm sections from formalin-fixed, paraffin-embedded tissue applied to coated slides. Deparaffinization, rehydration and epitope retrieval were performed in a Dako PT link (Dako) at 97°C for 20 minutes. Dako Autostainer Plus together with Envision™ Flex, high pH system (K8000, Dako) were used in the immunostaining procedure following the operating manual. The secondary antibody was incubated for 20 minutes. Sections were stained with anti-Bmi-1 (Santa Cruz Biotechnology) dilution 1:150 and anti-Mel-18 (Santa Cruz Biotechnology) dilution 1:75. Primary antibody incubated for 30 minutes. For Mel-18 visualization a FLEX + Rabbit (Linker) protocol was used. The choice of antibodies was made primarily on

the basis of literature studies, choosing clones used in previous publications [20,26]. The slides were counterstained with Hagen's Hematoxylin for visualization of tissue structures.

Evaluation of immunohistochemistry

The amount of positive cells and immunoreactivity intensity was evaluated semi- quantitative. For Bmi-1 only two grades were applied; if less than 10% of the epithelial cells were immunoreactive to Bmi-1, the sample was recorded as negative for Bmi-1 immunoreactivity, while samples showing more than 10% of cells were recorded as positive.

Soring grades for Mel-18 were as follows: Grade 1; < 5% of the cells positive for Mel-18. Grade 2; 6% to 35% showing positive immunoreactivity. Grade 3; 36-70% of the cells showed positive immunoreactivity, and grade 4 when more than 70% of the cells were positive for Mel-18 immunoreactivity. The intensity of the immunoreactivity was also recorded, and the grading was as follows: Grade 1, weak intensitivity, grade 2; moderate intensitivity and grade 3 when a strong intensitivity of the immunoreaction was observed. The immunoreactivity for both Mel-18 and Bmi-1 was evaluated by two independent investigators. There was no discrepancy between the two investigators.

Results

Bmi-1 transcription level

When comparing the transcription level of Bmi-1 in the different clinical groups, i.e. breast cancer, tissue taken in the vicinity of the tumor, fibroadenomas and breast tissue from non cancer controls, statistically significant differences were observed between the groups (table 3). Transcription level in non cancer breast controls was lowest, while transcription level in the normal adjacent tissue was more similar to that of the tumor (Figure 1 and 2).

Of special interest was the difference in transcription levels of Bmi-1 mRNA between non cancer controls and normal tissue from cancer patients (p = 0.041).

Mel-18 transcription level

The relative transcription level of Mel-18 in breast tissue from non cancer controls, fibroadenomas, tissue taken in the vicinity of the tumor and the tumor itself, is demonstrated in figure 3 and 4 and summarized in table 3. The mRNA transcription level of Mel-18 was statistically significantly higher in normal breast tissue from non cancer controls compared to normal tissue from cancer patients (p < 0.001).

There was a inverse relationship between the transcription level of Mel-18 and Bmi-1

When we compared the transcription level of Bmi-1 and Mel-18 in all categories (NC, non cancer controls, F, fibroadenomas N, normal and T tumor) we observed an inverse relationship between the transcription level of the two genes (figure 5).

Table 3 Mean transcription levels of Bmi-1 and Mel-18 in the different groups as well as the p-values and confidence intervals for the group-wise comparisons

Bmi-1	Group	mean	SD	p-value (confidence interval)		
				F	N	T
	NC	0.098	0.184	<0.001 (-0.424 -0.149)	0.0418 (-0.346 -0.007)	<0.001 (-0.790 -0.419)
	F	0.385	0.080	0.124 (-0.032 0.253)	<0.001 (-0.480 -0.157)	
	N	0.274	0.295		<0.001 (-0.620 -0.237)	
	T	0.703	0.659			
Mel-18	**Group**	**mean**	**SD**	**p-value (confidence interval)**		
				F	N	T
	NC	0.929	0.139	<0.001 (0.230 0.503)	<0.001 (0.362 0.707)	<0.001 (0.405 0.706)
	F	0.563	0.115		0.065 (-0.0115 0.347)	0.021 (0.030 0.348)
	N	0.395	0.352		0.827 (-0.172 0.214)	
	T	0.373	0.556			

Quantitative Real-time PCR

To validate the results obtained by the microarray analysis showing a clear inverse relationship between the transcription levels of the two genes, transcription levels of Bmi-1 and Mel-18 were also analyzed using qPCR. Some of the samples could not be validated due to insufficient mRNA amounts but all the groups were represented (9 non cancer controls, 4 fibroadenomas, 2 DCIS, 22 normals from cancer patients, and 69 tumors). Based on the work of McNeill et al [27] PPIA and MRPL19 were included as endogenous controls. However, analysis of our results showed large variations in transcription levels between the groups, especially for PPIA (results not shown). As an alternative approach, we therefore directly compared the transcriptional level of Bmi-1 to that of Mel-18 (figure 6).

Results of immunohistochemistry

In our material all of the samples stained positive both for Bmi-1 and Mel-18 but with different intensity. When protein expression of Bmi-1 was evaluated using immunohistochemsitry, as illustrated in figure 7 we

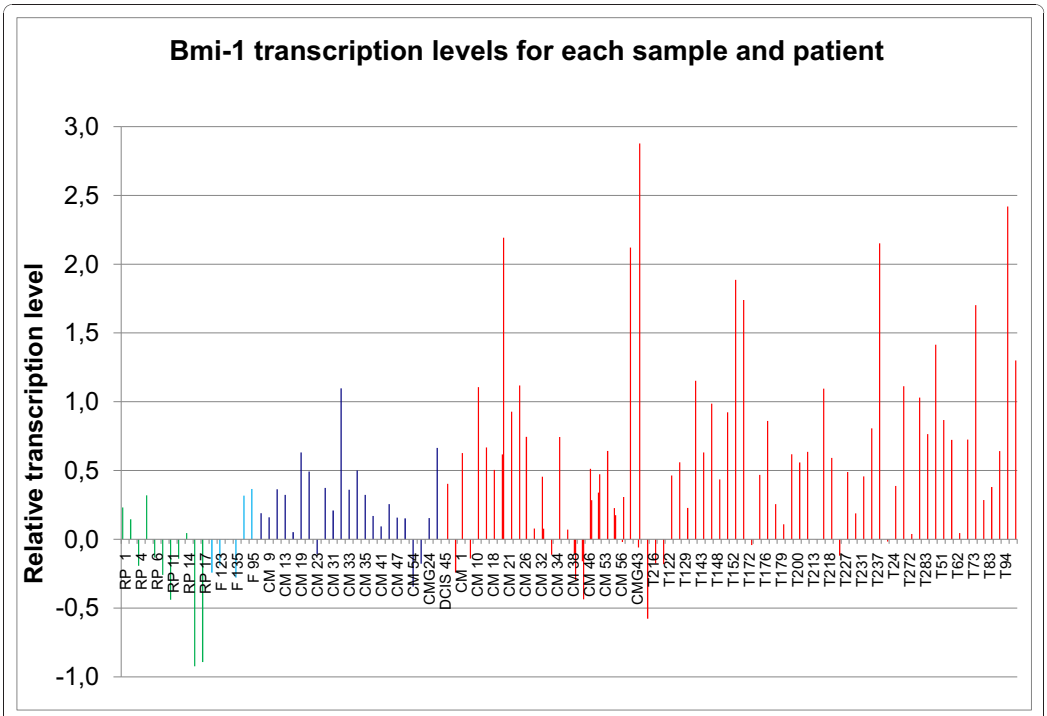


Figure 1 Bmi-1 transcription levels for each sample and patient. Green = non cancer controls (NC), light blue = fibroadenomas (F), dark blue = adjacent normal (N), and red = tumor (T). In cases where more than one tumor sample is shown for one patient, both preoperative and peroperative tumor tissue were available. On the y-axis there is a logarithmic transcription level of our sample compared to a Universal Human Reference ($\log_2$(expression of sample/expression of UHR)).

observed that almost for every case the expression intensity was stronger in normal tissue from breast cancer patients compared to normal breast tissue from non cancer patients, indicating differences in amount of protein in the cells between these two groups.

When evaluating the protein expression of Mel-18 using immunohistochemistry, breast tissue from non cancer individuals exhibited stronger expression intensity compared to normal tissue from cancer patients for almost every case, indicating probably higher amount of Mel-18 protein in breast tissue from non cancer controls.

However, both observations were subjectively made by a pathologist and no statistical differences were seen in numbers of positive cells between breast tissue from non cancer controls compared to normal tissue taken from a breast with a malignant tumor in the breast.

Discussion

In the present study we have analyzed levels of Mel-18 and Bmi-1 in normal breast tissue samples from patients operated for cancer, and compared to breast tissue samples from patients operated for non malignant condition and with no previously history of malignant disease.

We have shown that level of Bmi-1 is different in the normal breast tissue from cancer patients compared to normal breast tissue from non cancer controls at the mRNA level and there is also differences in staining intensity at the protein level, indicating that gene alterations associated with tumor development is already detected in the normal tissue, leading to higher risk for development of a malignant disease in the breast. At the same time we observed that at the protein level both genes were expressed in all studied tissue types, with no statistical differences in the numbers of positive cells between breast tissue from non cancer controls compared to normal tissue taken from a breast with a malignant tumor in that same breast. This may suggest that it is the change of expression above a certain threshold that may matter in exerting a cancer phenotype.

As Bmi-1 and Mel-18 are involved in cell aging [6,17], we looked for confounding effects of age of the patient

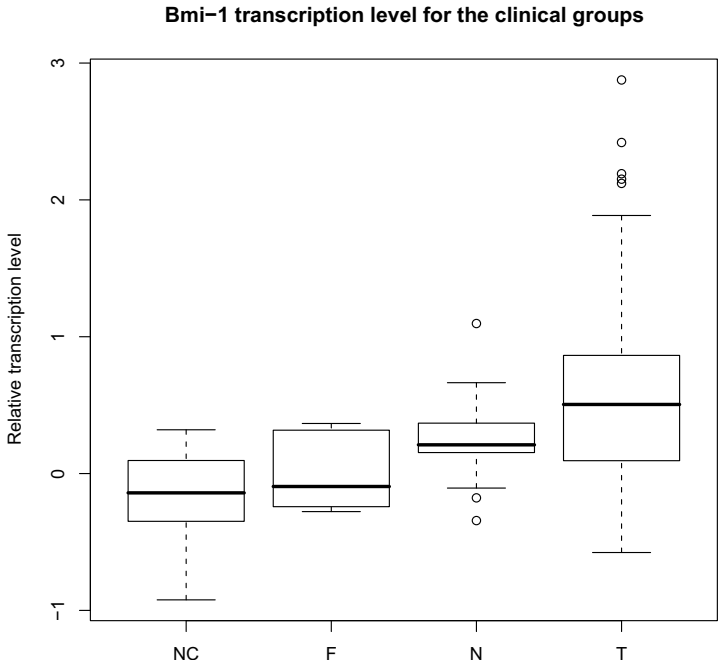


Figure 2 Bmi-1 transcription levels for the clinical groups. Boxplot showing the transcription level of Bmi-1 within the clinical groups. Group legends as in figure 1.

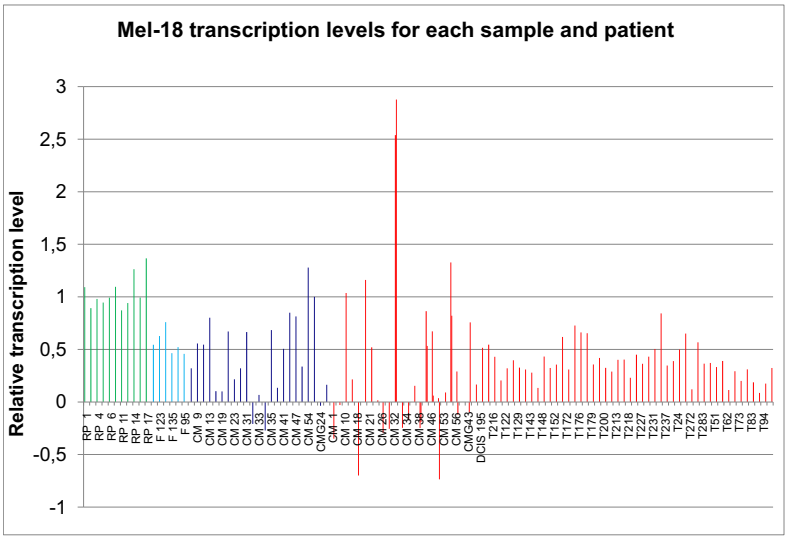


Figure 3 Mel-18 transcription levels for each sample and patient. Green = non cancer controls (NC), light blue = fibroadenomas (F), dark blue = adjacent normal (N) and red = tumor (T). In cases where more than one tumor sample is shown for one patient, both preoperative and peroperative tumor tissue were available. As in figure 1 there is logarithmic transcription level on the y-axis.

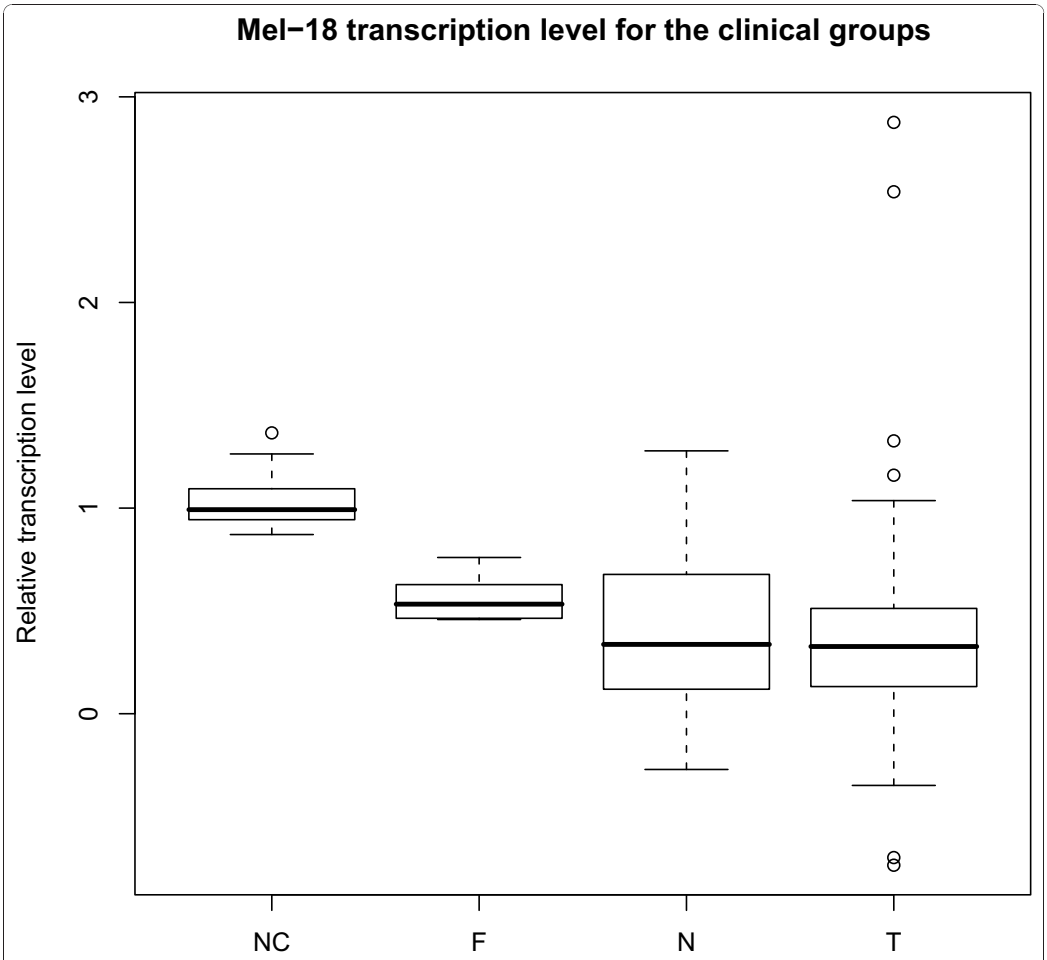


Figure 4 Mel-18 transcription levels for the clinical groups. Boxplot showing the transcription level of Mel-18 within the clinical groups. Group legends as in figure 3.

even though cell aging and aging of a person is not the same. Accounting for this we did regression analysis by ANOVA and found that age is not a significant factor in this respect ($p = 0.47$ for Bmi-1 and $p = 0.61$ for Mel-18). The different clinical groups on the other hand, were highly significant ($p = 0.0013$ for Bmi-1 and $p = 0.013$ for Mel-18).

The results from present study indicate that normal breast tissue in cancer patients carries different characteristics than that in women without previous history of malignant disease. One may speculate that it is possible to stratify women in risk groups for development of malignant tumor in the breast, according to transcription level ratios of genes like Bmi-1 and Mel -18.

After the introduction of mammography as a screening method in most countries, breast tumors are more often diagnosed at early stage. Most of the lesions detected by mammography are benign, where benign histology is confirmed by histological diagnosis of a biopsy. However, new cancer cases known as "interval cancers" still emerge and are diagnosed at later stage. The results from the present study indicate that it may be possible to distinguish between patients at risk by analyzing the normal breast tissue for genes like Bmi-1 and Mel-18.

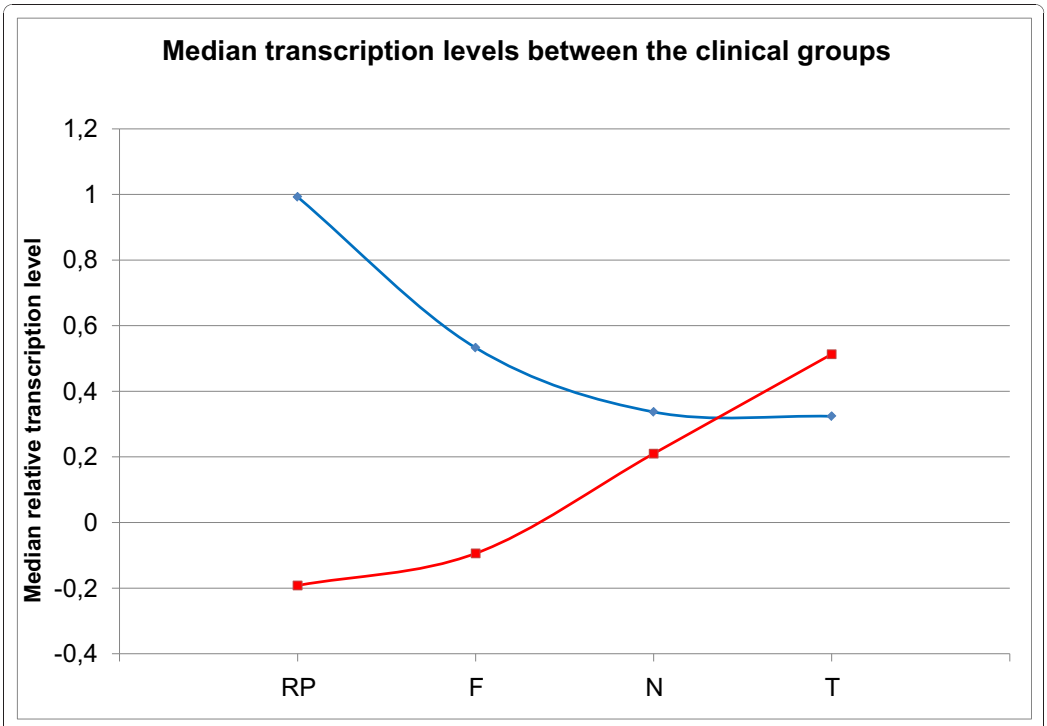


Figure 5 Median transcription levels between the clinical groups. Comparison of the median transcription level of Bmi-1 and Mel-18 for each clinical group shows an inverse correlation of the transcription level.

When comparing transcription level of mRNA for Bmi-1 between different sources of tissue, we observed a highly statistical difference in levels of transcription level. In normal tissue from breast cancer patients Bmi-1 mRNA was up-regulated, compared to tissue from breast without history of malignant disease. Again, there was no correlation between mRNA transcription level and the amount of cells positive for Bmi-1 in different groups, although intensitivity of immunoreactions were different between different groups and followed the same pattern as mRNA transcriptional level, indicating that there may be "more" protein in cells where mRNA transcriptional level was found to be high.

However, as in the case of Mel-18, mRNA transcription level is of transcripts isolated from all cells from which total RNA was isolated together and cannot directly correlate to the number of positive cells stained by immunohistochemistry. The number of cells in both specimens (for mRNA and protein analysis) is unknown and cannot be equal. Nevertheless the intensity of the immunohistochemistry, indicating amount of protein per cell, was inversely to what was observed for Mel-18 being higher in almost all samples from breast cancer patients compared to controls. Bmi-1 expression is necessary for normal cell cycle dynamics. It is possible that there may be "threshold" of Bmi-1 protein expression, and when this "threshold" is overridden, this protein may start to function as an oncogene. The highest transcription level of Bmi-1 was observed in invasive tumor tissue. Expression of Mel-18 was also analyzed. When transcription level of Mel-18 was analyzed in normal tissue from breast cancer patients, tumor tissue and tissue from patients without malignant disease, the lowest transcription level was observed in tumor tissue, and the highest in tissue from benign breast tissue. If this correlation reflects a direct functional interaction between Bmi-1 and Mel-18 is not possible to evaluate in this study. Nevertheless, we believe that expression analyses of both proteins may be an important tool for stratifying patients at risk for development of malignant disease in the breast. This is before the onset of an eventual malignant disease and as Silva et al suggest

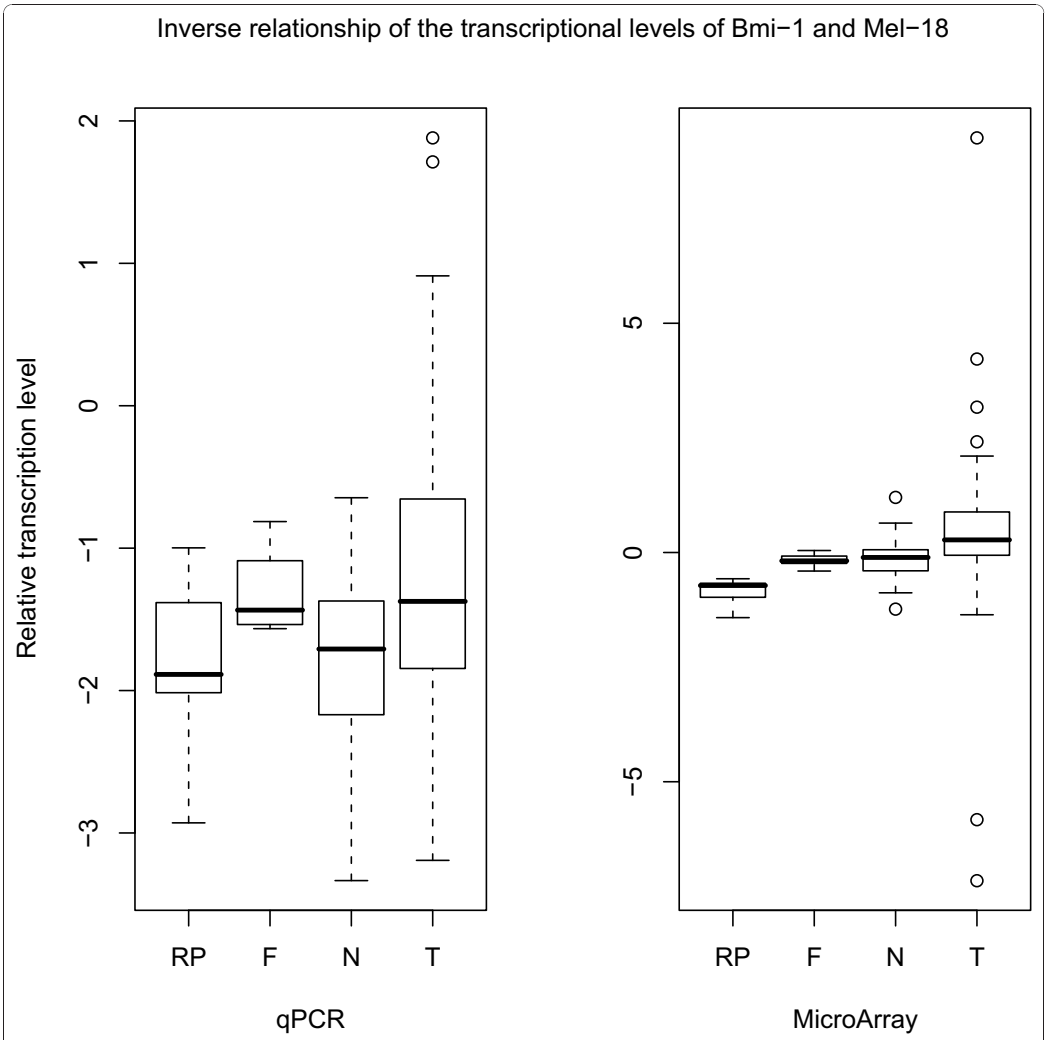


Figure 6 Quantitative real-time PCR shows the same inverse relationship between the transcriptional levels of Bmi-1 and Mel-18 as the microarray data. The qPCR results were calculated using the comparative C_T–method ($-\Delta\Delta C_T$) while the microarray results are given as the difference in expression of the two genes ($\log_2$(expression of Bmi-1/expression of Mel-18)).

[19] it can further be used as a diagnostic marker in patients who have already been diagnosed with breast cancer.

Data comprehensively comparing gene expression between histologically normal breast epithelium of breast cancer patients and cancer-free controls is limited. Similarly as our study Tripathi et al [28] did global gene expression of these two groups and conclude that cancer-related pathways are already perturbed in normal epithelium of breast cancer patients. This is cohesive to our study. Chen et al [29] also had the intention of studying malignancy-risk gene signature in histologically-normal breast tissue, but they compared the tumor tissue to the histologically normal tissue adjacent to the tumor, which is also part of what has been done in our study.

Saeki et al [30] reported a similar study on Bmi-1 where they found transcription level of the gene to be

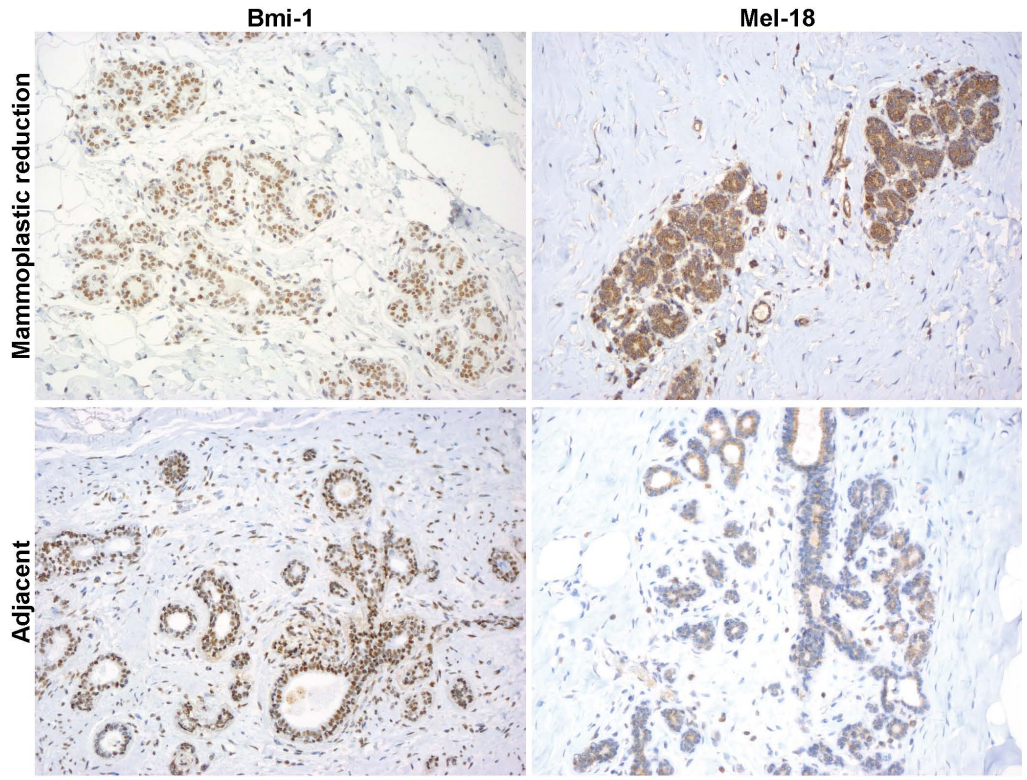


Figure 7 Immunostaining. Normal tissue from non cancer controls (upper panel) and normal tissue adjacent to a tumor (lower panel) of Bmi-1 (left) and Mel-18 (right). Demonstrating the increase in staining of Bmi-1 in tumor tissue, and the decrease in staining of Mel-18 in the same tissue. These are the outer limits and the staining of the groups in between are described in the text.

ten times higher in cancer tissue than non cancer controls. This is in coherence with our results except for the magnitude. A reason for this may be the observed large variation in tumor Bmi-1 levels in both materials and tumor heterogeneity may be responsible for the observed differences. By using laser micro dissection one could more accurately see the exact tissue from which mRNA was extracted.

Conclusion

In summary; we have in the present study demonstrated for first time that the levels of Mel-18 and Bmi-1 is different in normal tissue from breast cancer patients compared to breast tissue from non cancer controls. The transcription level of Bmi-1 and Mel-18 in all clinical categories (non cancer controls, fibroadenomas, normal, and tumor) was inversely correlated. Mel- 18 and Bmi-1 are two essential proteins in stem cell renewal pathways. Results from the present study indicate that expression profile analyses of that Mel-18 and Bmi-1 may be a tool for stratifying women at risk for development of malignant disease.

List of abbreviations

PcG: Polycomb Group of Proteins; DCIS: Ductal Carcinoma in Situ; UHR: Universal Human Reference; RP: Reduction Mammoplasty; F: Fibroadenoma; N: Normal; T: Tumor; and qPCR: Quantitative Real-time PCR;

Acknowledgements

We thank our surgical colleagues, Dr. S. Trønnes, Dr. B. Gravdehaug, and Dr. E. Kjæstad at the Surgical Department, Akershus University Hospital and colleagues at the Radiology Department same hospital for helping us collect the samples from cancer patients. Likewise we would like to thank colleagues at the Collosseum Clinic, Oslo, for providing samples from mammoplasties. We are grateful to T. Sonerud, Department for Clinical Molecular Biology (EpiGen), for help with performing the qPCR experiments.

MR is a fellow of the Faculty of Medicine, University of Oslo, this work was supported by grant number 2789119 from Helse Sør-Øst and internal grants from Akershus University Hospital, grant number 2689023.

Author details
[1]Department of Surgery, Akershus University Hospital, Lørenskog, Norway. [2]Department of Clinical Molecular Biology (EpiGen), Institute of Clinical Medicine, Akershus University Hospital, University of Oslo, Lørenskog, Norway. [3]Department of Pathology, Akershus University Hospital, Lørenskog, Norway. [4]Institute of Clinical Medicine, University of Oslo, Lørenskog, Norway.

Authors' contributions
MR was involved in the design of the study, collected the clinical data, contributed to the tissue collection and immunohistochemical analysis and is responsible for the preparation of the manuscript. TL performed the microarray experiments, contributed to the microarray, qPCR and statistical analysis, and is responsible for the preparation of the figures and the final formatting of the manuscript. AJN performed the immunohistochemical analysis. HV contributed to the microarray analysis. VK and IB designed the study, contributed to the preparation of the manuscript and supervised the work of MR, TL and HV. All authors have read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Received: 17 July 2010 Accepted: 16 December 2010
Published: 16 December 2010

References

1. Parkin DM, Bray F, Ferlay J, Pisani P: **Global cancer statistics, 2002.** *CA Cancer J Clin* 2005, **55**:74-108.
2. Paajanen H, Kyhala L, Varjo R, Rantala S: **Effect of screening mammography on the surgery of breast cancer in Finland: a population-based analysis during the years 1985-2004.** *Am Surg* 2006, **72**:167-171.
3. Weigel S, Batzler WU, Decker T, Hense HW, Heindel W: **First epidemiological analysis of breast cancer incidence and tumor characteristics after implementation of population-based digital mammography screening.** *Rofo* 2009, **181**:1144-1150.
4. Reya T, Morrison SJ, Clarke MF, Weissman IL: **Stem cells, cancer, and cancer stem cells.** *Nature* 2001, **414**:105-111.
5. Sparmann A, van LM: **Polycomb silencers control cell fate, development and cancer.** *Nat Rev Cancer* 2006, **6**:846-856.
6. Park , Morrison , Clarke : **Bmi1, stem cells, and senescence regulation.** *J Clin Invest* 2004, **113**:175-179.
7. Kim JH, Yoon SY, Jeong SH, Kim SY, Moon SK, Joo JH, Lee Y, Choe IS, Kim JW: **Overexpression of Bmi-1 oncoprotein correlates with axillary lymph node metastases in invasive ductal breast cancer.** *Breast* 2004, **13**:383-388.
8. Liu S, Dontu G, Mantle ID, Patel S, Ahn NS, Jackson KW, Suri P, Wicha MS: **Hedgehog signaling and Bmi-1 regulate self-renewal of normal and malignant human mammary stem cells.** *Cancer Res* 2006, **66**:6063-6071.
9. van Kemenade FJ, Raaphorst FM, Blokzijl T, Fieret E, Hamer KM, Satijn DP, Otte AP, Meijer CJ: **Coexpression of BMI-1 and EZH2 polycomb-group proteins is associated with cycling cells and degree of malignancy in B-cell non-Hodgkin lymphoma.** *Blood* 2001, **97**:3896-3901.
10. Lindstrom MS, Klangby U, Wiman KG: **p14ARF homozygous deletion or MDM2 overexpression in Burkitt lymphoma lines carrying wild type p53.** *Oncogene* 2001, **20**:2171-2177.
11. Bea S, Tort F, Pinyol M, Puig X, Hernandez L, Hernandez S, Fernandez PL, van LM, Colomer D, Campo E: **BMI-1 gene amplification and overexpression in hematological malignancies occur mainly in mantle cell lymphomas.** *Cancer Res* 2001, **61**:2409-2412.
12. Vonlanthen S, Heighway J, Altermatt HJ, Gugger M, Kappeler A, Borner MM, van LM, Betticher DC: **The bmi-1 oncoprotein is differentially expressed in non-small cell lung cancer and correlates with INK4A-ARF locus expression.** *Br J Cancer* 2001, **84**:1372-1376.
13. van Kemenade FJ, Raaphorst FM, Blokzijl T, Fieret E, Hamer KM, Satijn DP, Otte AP, Meijer CJ: **Coexpression of BMI-1 and EZH2 polycomb-group proteins is associated with cycling cells and degree of malignancy in B-cell non-Hodgkin lymphoma.** *Blood* 2001, **97**:3896-3901.
14. Alkema MJ, Jacobs H, van Lohuizen M, Berns A: **Perturbation of B and T cell development and predisposition to lymphomagenesis in eu-Bmi1 transgenic mice require the Bmi1 RING finger.** *Oncogene* 1997, **15**:899-910.
15. Itahana K, Zou Y, Itahana Y, Martinez JL, Beausejour C, Jacobs JJ, van LM, Band V, Campisi J, Dimri GP: **Control of the replicative life span of human fibroblasts by p16 and the polycomb protein Bmi-1.** *Mol Cell Biol* 2003, **23**:389-401.
16. Dimri GP: **What has senescence got to do with cancer?** *Cancer Cell* 2005, **7**:505-512.
17. Guo WJ, Datta S, Band V, Dimri GP: **Mel-18, a polycomb group protein, regulates cell proliferation and senescence via transcriptional repression of Bmi-1 and c-Myc oncoproteins.** *Mol Biol Cell* 2007, **18**:536-546.
18. Dimri GP, Martinez JL, Jacobs JJ, Keblusek P, Itahana K, van LM, Campisi J, Wazer DE, Band V: **The Bmi-1 oncogene induces telomerase activity and immortalizes human mammary epithelial cells.** *Cancer Res* 2002, **62**:4736-4745.
19. Silva J, Garcia V, Garcia JM, Pena C, Dominguez G, Diaz R, Lorenzo Y, Hurtado A, Sanchez A, Bonilla F: **Circulating Bmi-1 mRNA as a possible prognostic factor for advanced breast cancer patients.** *Breast Cancer Res* 2007, **9**:R55.
20. Guo WJ, Zeng MS, Yadav A, Song LB, Guo BH, Band V, Dimri GP: **Mel-18 acts as a tumor suppressor by repressing Bmi-1 expression and down-regulating Akt activity in breast cancer cells.** *Cancer Res* 2007, **67**:5083-5089.
21. Tagawa M, Sakamoto T, Shigemoto K, Matsubara H, Tamura Y, Ito T, Nakamura I, Okitsu A, Imai K, Taniguchi M: **Expression of novel DNA-binding protein with zinc finger structure in various tumor cells.** *J Biol Chem* 1990, **265**:20021-20026.
22. Dukers DF, van Galen JC, Giroth C, Jansen P, Sewalt RG, Otte AP, Kluin-Nelemans HC, Meijer CJ, Raaphorst FM: **Unique polycomb gene expression pattern in Hodgkin's lymphoma and Hodgkin's lymphoma-derived cell lines.** *Am J Pathol* 2004, **164**:873-881.
23. Wei J, Kahn J: **Purification of Total RNA from Mammalian Cells and Tissues.** In *DNA Microarrays: A Molecular Cloning Manual.* Edited by: Bowtell D, Sambrook J. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2002:110-119.
24. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, *et al*: **Bioconductor: open software development for computational biology and bioinformatics.** *Genome Biol* 2004, **5**:R80.
25. Bloom HJ, Richardson WW: **Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years.** *Br J Cancer* 1957, **11**:359-377.
26. Song LB, Zeng MS, Liao WT, Zhang L, Mo HY, Liu WL, Shao JY, Wu QL, Li MZ, Xia YF, *et al*: **Bmi-1 is a novel molecular marker of nasopharyngeal carcinoma progression and immortalizes primary human nasopharyngeal epithelial cells.** *Cancer Res* 2006, **66**:6225-6232.
27. McNeill RE, Miller N, Kerin MJ: **Evaluation and validation of candidate endogenous control genes for real-time quantitative PCR studies of breast cancer.** *BMC Mol Biol* 2007, **8**:107.
28. Tripathi A, King C, de la Morenas A, Perry VK, Burke B, Antoine GA, Hirsch EF, Kavanah M, Mendez J, Stone M, *et al*: **Gene expression abnormalities in histologically normal breast epithelium of breast cancer patients.** *Int J Cancer* 2008, **122**:1557-1566.
29. Chen DT, Nasir A, Culhane A, Venkataramu C, Fulp W, Rubio R, Wang T, Agrawal D, McCarthy SM, Gruidl M, *et al*: **Proliferative genes dominate malignancy-risk gene signature in histologically-normal breast tissue.** *Breast Cancer Res Treat* 2010, **119**:335-346.
30. Saeki M, Kobayashi D, Tsuji N, Kuribayashi K, Watanabe N: **Diagnostic importance of overexpression of Bmi-1 mRNA in early breast cancers.** *Int J Oncol* 2009, **35**:511-515.

Pre-publication history
The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2407/10/686/prepub

doi:10.1186/1471-2407-10-686
Cite this article as: Riis *et al.*: Expression of BMI-1 and Mel-18 in breast tissue - a diagnostic marker in patients with breast cancer. *BMC Cancer* 2010 **10**:686.