

Novel biomarkers predicting long-term survival in breast cancer

**Elin
Karlsson**

Göteborg 2009



**Department of Oncology
Institute of Clinical Sciences
The Sahlgrenska Academy at University of Gothenburg**



UNIVERSITY OF GOTHENBURG

Copyright © Elin Karlsson 2009

Printed by Intellecta Infolog AB
Gothenburg, Sweden 2009
ISBN 978-91-628-7787-3

ABSTRACT

Breast cancer is the most common malignancy among women, affecting over a million women worldwide every year. During the last decades, there has been a dramatic increase in the survival rates due to earlier detection and improved treatment. Breast cancer treatment today is getting more and more targeted, but still, many patients are being overtreated, and some undertreated. Therefore, the need for additional complementary prognostic markers is urgent. In this thesis, molecular differences in tumours from breast cancer survivors and deceased patients have been explored on the DNA, RNA and protein levels. The major findings include differences on the genomic level between lymph node-negative 10-year survivors and deceased patients; gains at 4q, 5q31-5qter, 6q12-6q16 and 12q14-12q22 and losses at 8p21.2-8p21.3, 8p23.1-8p23.2, 17p, 18p, Xp21.3, Xp22.31-Xp22.33 and Xq were significantly more frequent in tumours from deceased patients compared to tumours from 10-year survivors. Gains at 1q25.2-1q25.3 and 1q31.3-1q41 were more common in tumours from 10-year survivors. In addition, a gene signature consisting of 51 genes was generated. The expression profile of these 51 genes predicted clinical outcome in our material of node-negative patients as well as in an external tumour material with good accuracy. The protein expression of four genes (*ADIPOR1*, *ADORA1*, *BTG2* and *CD46*) that differed between the survival groups, both in DNA copy number alterations and in gene expression, was explored in a larger independent cohort of breast cancer patients. The protein expression of *BTG2* significantly more frequent in tumours from 5-year survivors compared to tumours from deceased patients. This finding indicates expression of *BTG2* as a possible prognostic biomarker. Furthermore, the prognostic biomarkers found in this work, may in the future facilitate the prognosis as well as predict course of treatment for breast cancer patients, following extensive validation.

LIST OF PAPERS

This academic thesis is based on the following papers:

- I **Karlsson E.**, Danielsson A., Delle U., Olsson B., Karlsson P., Helou K.
Chromosomal changes associated with clinical outcome in lymph node-negative breast cancer
Cancer Genetics and Cytogenetics 2007;172(2):139-46

- II **Karlsson E.**, Delle U., Danielsson A., Parris T., Olsson B., Karlsson P., Helou K.
High-resolution genomic profiling to predict 10-year survival in node-negative breast cancer
Manuscript, 2009

- III **Karlsson E.**, Delle U., Danielsson A., Olsson B., Abel F., Karlsson P., Helou K.
Gene expression variation to predict 10-year survival in lymph-node-negative breast cancer
BMC Cancer, 2008;8(1):254

- IV **Karlsson E.**, Kovács A., Delle U., Lövgren K., Danielsson A., Parris T., Brennan D., Jirström K., Karlsson P., Helou K.
Up-regulation of cell cycle arrest protein BTG2 correlates with increased survival in breast cancer
Manuscript, 2009

CONTENTS

ABSTRACT	5
LIST OF PAPERS.....	6
CONTENTS	7
ABBREVIATIONS	8
INTRODUCTION	11
Cancer.....	11
Cancer genetics	12
Breast cancer	12
Prognostic and predictive markers in breast cancer	13
Axillary lymph node Status	14
Novel prognostic molecular markers.....	15
Future prognostic molecular markers	16
AIMS	17
MATERIALS AND METHODS.....	18
Tumour material	18
Metaphase CGH.....	20
Microarrays.....	21
Array CGH.....	21
Expression microarray	22
Tissue array	24
Quantitative Real Time PCR.....	24
Statistics.....	25
RESULTS AND DISCUSSION	26
Genomic level.....	26
Gene expression level.....	29
Protein expression level.....	31
CONCLUDING REMARKS	35
FUTURE PERSPECTIVES	36
SAMMANFATTNING PÅ SVENSKA	37
ACKNOWLEDGEMENTS.....	38
REFERENCES	40

ABBREVIATIONS

ABBREVIATIONS

aCGH	array CGH
BAC	bacterial artificial chromosome
cDNA	complementary DNA
CGH	comparative genomic hybridisation
CNA	copy number alteration
Cy3	Cyanine 3
Cy5	Cyanine 5
DAPI	4,6-diamino-2-phenylindole
DNA	deoxyribonucleic acid
EA	expression microarray
FITC	fluorescein isothiocyanate
HER2	human epidermal growth factor receptor 2
IHC	Immunohistochemistry
mCGH	metaphase CGH
mRNA	messenger RNA
PAI-1	plasminogen activator inhibitor
PCR	polymerase chain reaction
QPCR	quantitative real time PCR
RIN	RNA integrity number
RNA	ribonucleic acid
TMA	tissue microarray
TRITC	tetramethylrhodamine isothiocyanate
uPA	urokinase plasminogen activator

INTRODUCTION

Cancer

Cancer affects approximately 10.9 million people worldwide every year (non-melanoma skin cancer not included) [1]. In Sweden, approximately 50 000 patients are diagnosed with cancer every year [2]. Breast cancer is the most frequent malignancy among women, in Sweden as in the world in general (Figure 1) [1, 2]. The most common cause of death due to cancer is however lung cancer, that accounts for more than 1 million deaths worldwide every year.

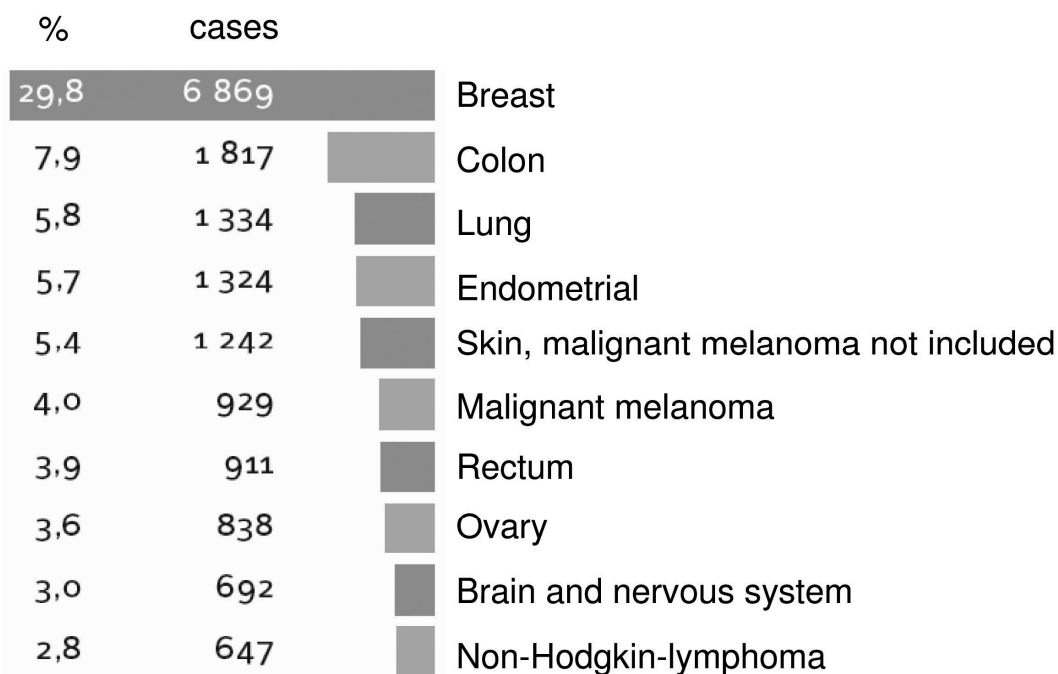


Figure 1. The ten most common malignancies among women in Sweden. Both percent and number of cases in Sweden per year are specified (the figure were originally published in Socialstyrelsens Cancer i siffror [2]).

Treatment for cancer patients imposes a considerable economic burden on the health care systems worldwide because of the high incidence rate of the disease. The selection of treatment is influenced by various prognostic factors, sometimes inadequate, resulting in over-treatment of many cancer cases. Therefore, many patients could benefit from accurate complements to the presently available prognostic markers, which may assist in the development of new therapeutic agents, vaccines and more individualised treatments.

Cancer genetics

Cancer is a heterogeneous genetic disease that arises from one single cell acquiring unlimited growth properties through genetic events. The specific genetic events are affected by the patient's genetic predisposition and environmental factors, such as diet, usage of tobacco, exposure to radiation, carcinogenic air pollution, food contaminants, viruses and microorganisms [3]. The genesis of cancer is a multistep process, where several genetic events are required for a normal cell to transform into a malignant one. It is suggested that most or maybe all tumours need to gain at least six essential alterations in cell physiology that collectively lead to malignant growth; self-sufficiency in growth signals, insensitivity to anti-growth signals, tissue invasion properties and metastasis formation, evading apoptosis, sustained angiogenesis, and limitless replicative potential [4]. These modifications of cell activity are due to changes in cancer-related genes; either oncogenes that gain function and thereby promote cell growth, tumour suppressor genes that decrease in expression or cease functioning, or DNA repair genes that lose their function resulting in genomic instability, which can facilitate other cancer promoting events to occur. These genetic proceedings vary enormously even within the same group of tumours, which makes cancer a complex disease to study.

Breast cancer

Breast cancer is by far the most common malignancy among women in the world; about 23% of all female cancer cases are breast cancer in the world, and in Sweden, 30% of all female cancer cases are breast cancer. In 2002, breast cancer accounted for 1.15 million new cases and 411 000 deaths. Furthermore, approximately 4.4 million living women around the world were diagnosed with breast cancer in the last five years. The breast cancer incidence rate is high in Europe and North America, accounting for more than half of all breast cancer cases around the world, while incidence rates in Africa and Asia are low. The highest rate is in North America (99.4 per 100 000), and Central Africa has the lowest incidence rate (16.5 per 100 000) [1]. In Northern Europe, the incidence rate is 82.5 per 100 000, and approximately 7 000 women are affected in Sweden every year [2]. During the last decades the survival rate of breast cancer patients has increased dramatically, due to earlier detection and new methods of treatment [5]. The 5-year survival rate in Sweden is approximately 86%, and 10-year survival is 75.5% [2]. Almost all breast cancer patients in Sweden are treated with radical surgery followed by different courses of treatment depending on characteristics determined by prognostic markers.

Prognostic and predictive markers in breast cancer

In breast cancer, different characteristics of the patient and the tumour are used to determine the risk of relapse and death, as well as proper treatment following surgery. However, many patients continue to receive treatment from which they do not benefit: many would have remained disease-free even without treatment or recurrent disease develop despite treatment. Some patients do not receive treatment they would have benefitted from, due to false favourable prognostic characteristics of their disease. This means that there is still a great need of additional prognostic markers (i.e. markers that predict prognosis) and predictive markers (i.e. markers predicting therapy response) in order to further tailor the treatment of each individual patient.

According to the St Gallen criterion, presence of steroid hormone receptors, lymph node status, size and differentiation grade of the tumour, as well as age at diagnosis are used to classify breast cancer patients into groups that determine which treatment the patients should receive after surgery [6]. Lately, the molecular marker HER2 and peritumoural vascular invasion have been taken into clinical use. Initially, all patients whose tumours present expression of any of the two steroid hormone receptors (oestrogen and progesterone), independent of any other marker, are considered endocrine responsive and are thereby in most cases given adjuvant endocrine treatment. In patients where the endocrine response of the disease is uncertain (low or insufficient detected expression of steroid hormone receptors), a combination of endocrine treatment and chemotherapy is used. Furthermore, patients whose tumours do not present steroid receptor expression are treated with chemotherapy. Patients with tumours > 10 mm that over express HER2 are treated with trastuzumab, an antibody directed against the HER2 receptor [7]. Additionally, patients are classified into risk groups depending on lymph node status, tumour size, grade, age at diagnosis, HER2 over expression and peritumoural vascular invasion. In general, node-negative patients with no risk attributes are classified as low risk patients. Node-negative patients presenting any of the risk factors are considered intermediate risk patients together with patients presenting 1-3 affected lymph nodes, endocrine responsive tumours and no HER2 over expression. The high-risk group includes patients presenting 1-3 affected lymph nodes and endocrine non-responsive tumours or HER2 over expression together with patients with more than four affected lymph nodes. The risk categories are summarised in Table 1.

INTRODUCTION

Table 1. Summary of the breast cancer risk categories according to the St Gallen criteria [6].

Risk category	
Low risk	node-negative and all following features: tumour size < 20 mm grade 1 absence of peritumoural vascular invasion oestrogen and/or progesterone receptor expressed HER2/ <i>neu</i> neither over expressed nor amplified age > 35 years
Intermediate risk	node-negative and not fulfilling any of the above mentioned features node-positive (1-3 involved nodes) and both of the following features oestrogen and/or progesterone receptor expressed HER2/ <i>neu</i> neither over expressed nor amplified
High risk	node-positive (1-3 involved nodes) and any of the following features oestrogen and progesterone receptor absent HER2/ <i>neu</i> over expressed or amplified node-positive (4 nodes or more involved)

AXILLARY LYMPH NODE STATUS

Approximately 70% of breast cancer patients diagnosed in Sweden today have axillary lymph nodes free from metastasis. The first three papers in this thesis are based on tumour samples from lymph node-negative patients.

Around 95% of the lymph drainage from the breast goes through the axillary lymph nodes (lymph nodes localised in the armpit) and therefore these nodes are usually the initial site of breast tumour metastasis. Lymph node status is the most important marker of tumour aggressiveness. Although metastasis-free lymph nodes are a sign of a less aggressive tumour, around 20% of lymph node-negative breast cancer patients die within fifteen years of diagnosis [8].

In the 21st century, sentinel node biopsy has replaced axillary dissection as the common way of examining lymph node involvement in breast cancer. In this way, axillary dissection can be avoided in patients that will not benefit from the procedure, which often results in temporary impaired function of the arm. A combination of coloured and radioactive fluid is injected into the breast at the start of operation and the first lymph node dyed is identified as the sentinel node. This node is removed and immediately analysed by a pathologist during the course of surgery of the actual breast tumour. The result of the sentinel node examination is delivered to the surgeon before the operation is terminated and axillary dissection is only performed if the node contains cancer cells.

NOVEL PROGNOSTIC MOLECULAR MARKERS

Presently, most evidence indicates that genetic alterations giving cells the capacity to metastasise and thereby eventually kills the patient are early events in tumour progression, and that the majority of cells, if any, in a primary tumour possess this capacity [9]. This implies that primary tumours are genetically predestined to evolve aggressive behaviour in the initial stages of tumour progression, making it possible to predict patient outcome by evaluating the molecular characteristics of the cells of the primary tumour.

The HER2/*neu* (Human Epidermal growth factor Receptor 2) marker is a recent molecular marker that is now in full use in the clinic. It is a transmembrane tyrosine kinase receptor amplified and overexpressed in approximately 10-20% of all breast cancers. If the protein transcribed from the proto-oncogene *ERBB2* is over expressed in the tumour, the patient is classified into a higher risk group and normally given tailored treatment [10], which is a monoclonal antibody directed against the extracellular part of HER2, blocking the receptor inhibiting tumour cell growth [7].

Expression profiling is widely used experimentally to classify breast tumours into molecular sub-categories, as well as in predicting clinical outcome [11-31]. This approach has been rather successful and two different gene expression profiles are presently being tested in clinical trials [32, 33]. One of these gene expression profiles, "MammaPrint", predicts disease-free 5-year survival in early breast cancers, whereas the other, "Oncotype DX RS", predicts recurrence-free survival in lymph node-negative, tamoxifen treated, breast cancers.

Proliferation is an important characteristic of tumour cells. Since cells in G1, S, G2 and mitosis express Ki67, but resting cells do not, Ki67 expression is an adequate marker of proliferation [34]. High expression of Ki67 has been indicated as a marker of both decreased overall survival and decreased disease-free survival, predominantly in node-negative patients [35]. There are however, contradictory reports [36].

Two other promising prognostic markers are urokinase plasminogen activator (uPA) and plasminogen activator inhibitor (PAI-1). These are markers of tumour proteolytic activity, which facilitates invasion through the extracellular matrix [37], making it possible for the tumour to metastasise. uPA and PAI-1 levels have been strongly associated with both disease-free and overall survival in node-negative as

INTRODUCTION

well as node positive patients [38-42]. A major drawback of these two markers is that they can only be measured in fresh frozen tissue.

FUTURE PROGNOSTIC MOLECULAR MARKERS

In the last decades, the only molecular marker that has actually been taken into clinical practise is HER2, despite massive research in the area. There are, however, numerous studies on less established prognostic markers, including genomic profiling, gene expression and various protein markers. Who knows what the future may bring in terms of both prognostic and predictive markers as well as novel treatments. Undoubtedly, there will be new ways to approach breast cancer patients, simply because of the effort put in to this matter.

AIMS

The overall prospective purpose of this study was to identify new molecular markers for long-term survival in breast cancer patients.

In the individual papers the aims were:

In **Paper I** the aim was to identify copy number changes of chromosomal regions in the tumour genome differing in frequency between patients that died from the breast cancer and patients surviving for at least ten years.

Paper II is a study similar to the investigation in Paper I, though using a new method holding a greatly increased resolution. The aim was to specify genetic alterations that affect 10-year survival in breast cancer patients.

Paper III is a screening study of gene expression and the aim was to develop a list of genes whose expression could predict clinical outcome in breast cancer.

In **Paper IV**, the findings from Papers II and III were combined in order to find a number of genes to study on the protein level in a larger independent set of tumours. The aim was to evaluate if these genes differed in protein expression between deceased breast cancer patients and long-term survivors.

MATERIALS AND METHODS

The work on this thesis began in 2003 utilising the, at the time, up-to-date genome wide screening method metaphase comparative genomic hybridisation (mCGH). Rather rapidly, the microarray methods became renowned, easy to use, and affordable for common researchers. The resolution accessible by microarrays is considerably higher than that of metaphase CGH and therefore, we pursued working with arrays, both on the DNA, RNA and protein level.

Tumour material

In order to determine molecular changes contributing to tumour development in human breast cancers, fresh frozen tumours have been collected for analysis between 1985 and 1998 in the Västra Götaland region of Sweden. These tumours have been investigated pathologically, analysed for oestrogen- and progesterone receptors, as well as S-phase determination. Based on these results, the stages and degrees of differentiation of the tumours were determined. These tumour samples have been stored for continued research at the Sahlgrenska University Hospital Oncology Lab Tumour Bank. In total, 67 of these tumours collected between 1990 and 1997 were used in **Papers I, II and III**. Of these, 39 tumours were analysed by all three methods, namely metaphase CGH (mCGH), microarray CGH (aCGH) and gene expression microarray (EA). Four samples were studied by both EA and mCGH, two tumours were analysed by EA and aCGH, one tumour by EA analysis only, and 21 tumour samples were analysed exclusively by mCGH (Figure 2).

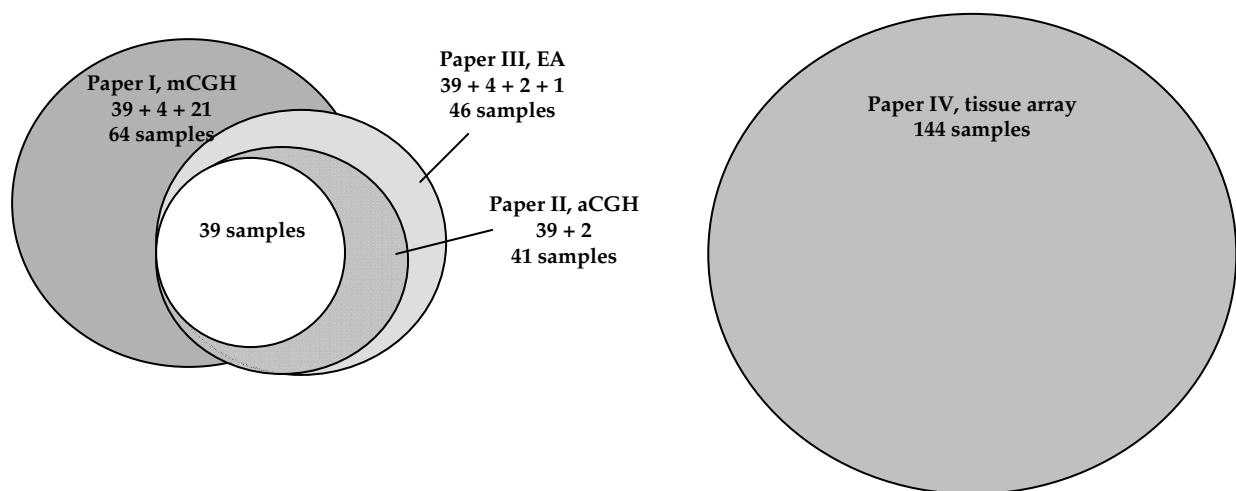


Figure 2. Distribution of tumour samples used in the four papers included in this thesis. In the first three papers, 39 samples were used in all studies, 4 in mCGH and EA, 2 in EA and aCGH and 21 in mCGH exclusively, and 1 in EA exclusively. In Paper IV, a new independent material was investigated.

The aim of these studies was to analyse molecular differences between 10-year survivors and deceased lymph node-negative breast cancer patients. In **Paper I**, initially tumour samples from node-negative patients in general were collected from the tumour bank, which naturally resulted in tumours primarily from survivors. In the next step, we wanted preferably tumours from deceased node-negative patients to balance the groups, and we used a list consisting of tumours from stage I node-negative patients (tumours smaller than 20 mm), to collect most of the samples from deceased patients. This resulted in an uneven distribution of tumour size and clinical stage in the two survival groups. This fact was taken into consideration, and we reasoned that this should not make the results less reliable. In fact, stage I tumours are normally less aggressive and here we used a subgroup of these stage I tumours that actually killed the patient, making the group of tumours from deceased patients to some extent more extreme, possibly giving us the opportunity to more distinct detect molecular differences between 10-year survivors and deceased patients. In the final mCGH analysis, seven samples were excluded: two of the patients lacked ten years of follow up and another five samples were from patients who died in intercurrent disease. The mCGH survival analysis consisted thereby of 57 samples, where 35 were tumours from 10-year survivors and 22 from deceased patients (Table 2). In **Paper III**, the tumours analysed in **Paper I** were used as a starting point. In expression analysis, the quality of the RNA is critical; therefore, some tumour samples were excluded due to poor RNA quality or insufficient material for RNA extraction. To further balance the survival groups in **Paper III**, three new specimens were used. Finally, this study consisted of 23 tumours from 10-year survivors and 23 tumours from patients that died within ten years (Table 2). In **Paper II**, all but five of the tumour samples used in **Paper III** were included, due to access of material, which resulted in 41 samples, 22 from 10-year survivors and 19 from deceased patients (Table 2). Clinical information for the samples used in these investigations is presented in each individual paper.

Table 2. Number of samples in the four studies. In Paper I-III the tumours were from 10-year survivors or from patients that died within ten years from diagnosis. In Paper IV, we compared tumours from 5-year survivors to tumours from patients that died within five years from diagnosis.

	total number of samples	tumours from survivors	tumours from deceased patients
Paper I	57	35	22
Paper II	41	22	19
Paper III	46	23	23
Paper IV	143	111	32

MATERIALS AND METHODS

In **Paper IV**, protein expression of markers of interest, discovered in **Papers II** and **III**, were explored in a new breast tumour set collected in Malmö in southern Sweden. This material consisted of 144 primary breast tumours attached to tissue microarray slides. Thirty-two of the primary tumours were from deceased patients, 111 from 5-year survivors and one lacked 5 years of follow up. Additional information of the patients is presented in **Paper IV**.

Metaphase CGH

In this molecular, cytogenetic method, tumour DNA is compared to normal DNA by means of competitive hybridisation to chromosome preparations after labelling the DNA with different fluorochromes, tumour DNA with green fluorescence and reference with red fluorescence. Regions in which the DNA sequence copy number is higher in the tumour DNA relative to normal DNA, genomic gain or amplification, will be identified as predominantly green fluorescing regions, whereas regions of predominantly red fluorescence represent loss or deletion of genetic material (Figure 3a). In the mCGH experiments, metaphase spreads from the tumour is not required, only genomic DNA is needed. This makes CGH ideal for the analysis of chromosomal changes in solid tumours where classical cytogenetics analysis may be restricted by technical limitations with metaphase preparations, such as low mitotic index or insufficiency in spreading of metaphases. A disadvantage of this method is that balanced rearrangements such as translocations and inversions are not detectable. Neither can mutations and copy number changes smaller than 10 Mb be detected [43]. Despite the limited resolution of this method, it has a substantial advantage in that it provides an overview of the genetic alterations in the tumour genome in one single experiment.

In our experiment, DNA was extracted from frozen tumours and reference DNA was extracted from lymphocytes drawn from a healthy female. CGH was performed essentially as described by Kallioniemi et al [43, 44] with minor modifications [45]. Briefly, tumour and reference DNA were differently labelled by nick translation. Equal amounts of labelled reference and tumour DNA were co-precipitated, denatured and hybridised to human metaphase slides made from lymphocytes from healthy females. The DNA probes were detected with the fluorochromes FITC (green) for the tumour DNA and TRITC (red) for the reference DNA. The metaphases were counterstained with DAPI for identification of the chromosomes. For each tumour, 10-19 (mean 14) metaphases were analysed using a Leica CW4000 software package where the FITC and TRITC images are merged

together to generate an average fluorescence ratio profile (Figure 3a). Chromosomal regions containing repetitive DNA sequences (1p32-1pter, 16p, 19 and 22 as well as chromosome telomeres and centromeres) have shown to be difficult to analyse using this method [43] and were therefore excluded from the analysis.

Microarrays

The novel research method of microarray analysis is a powerful tool for detecting genomic and expression levels of individual genes. The type of microarray analysis utilised in this thesis are CGH microarrays, gene expression microarrays, and tissue microarrays. aCGH uses genomic bacterial artificial chromosome (BAC) clones instead of chromosomes as hybridisation targets on the slides. In the gene expression and tissue microarray analyses, cDNA oligos and tissue sections were used. The DNA-clones or tissue sections are tightly bound to glass slides to create “micro-grids”, i.e. slides with hundreds to thousands of DNA-clones or tumours strictly ordered in lines. Because of the enormous amount of data generated from microarray experiments, it is important to use appropriate statistics when performing the data analysis.

ARRAY CGH

Array CGH is a development of metaphase CGH, where the targets are small fragments of DNA spotted out on a glass slide, instead of using chromosomes as targets. Test and reference DNA are differently labelled, and for each spot it is possible to determine the quote of fluorescent light emitted from hybridised test DNA versus reference DNA. Together, these spots create a genomic profile of gains, amplifications and losses of the total tumour genome. A schematic overview of the aCGH procedure is compiled in Figure 3b. The aCGH slides used in our experiment were constructed at the SCIBLU Genomics Center, Department of Oncology, Lund University, Sweden [46], and consisted of approximately 38 000 different DNA probes (BACs) that cover the entire genome in a tiling manner. As in mCGH, balanced rearrangements such as translocations and inversions are not detectable, and in these specific arrays, neither are alterations below 100 kb in size.

The same DNA as in the mCGH study was used, except for a few cases where new DNA was extracted. The DNA was purified using phenol/chloroform. Female reference DNA was purchased from Promega (Madison, WI, USA). aCGH was performed essentially as previously described [47]. In brief, equal amounts of Cy3 labelled test DNA (red) and Cy5 labelled reference DNA (green) were co-

MATERIALS AND METHODS

precipitated, denatured and hybridised to aCGH slides. The slides were scanned after the washing procedure and the Cy3 and Cy5 images were merged and analysed in the GenePix Pro software 6.0.1.12 to exclude inadequate spots.

EXPRESSION MICROARRAY

Expression microarray (EA) is a screening method where the expression levels of genes are studied using tumour mRNA. Labelled cDNA synthesised from mRNA is hybridised to a glass slide containing a quantity of spots (in our case 35 000 spots) consisting of oligo synthesised DNA-fragments. The expression level of each gene is measured by analysing the signal intensities of each spot. A schematic overview of the EA procedure is presented in Figure 3c.

The expression microarrays used in our study were produced at the Swegene DNA Microarray Resource Center, Department of Oncology, Lund University, Sweden [48]. Total mRNA was extracted from the tumour samples using TRIzol Reagent. The quality of the RNA was evaluated using the Agilent 2100 Bioanalyzer and specimens where the 28S/18S ratio was lower than 1.0 or the RNA integrity number (RIN)-value [49] was lower than 6.7 were excluded from the study. For each sample, cDNA probes labelled with Cy3 (red) were synthesised from the total tumour RNA by reverse transcription. Reference cDNA labelled with Cy5 (green) was synthesised from commercial reference RNA. Labelled tumour cDNA and reference cDNA were co-precipitated and hybridised to the microarray slide. The microarray slides were scanned, the Cy3 and Cy5 images were then merged and analysed in the GenePix Pro software 6.0.1.12 in order to exclude inadequate spots.

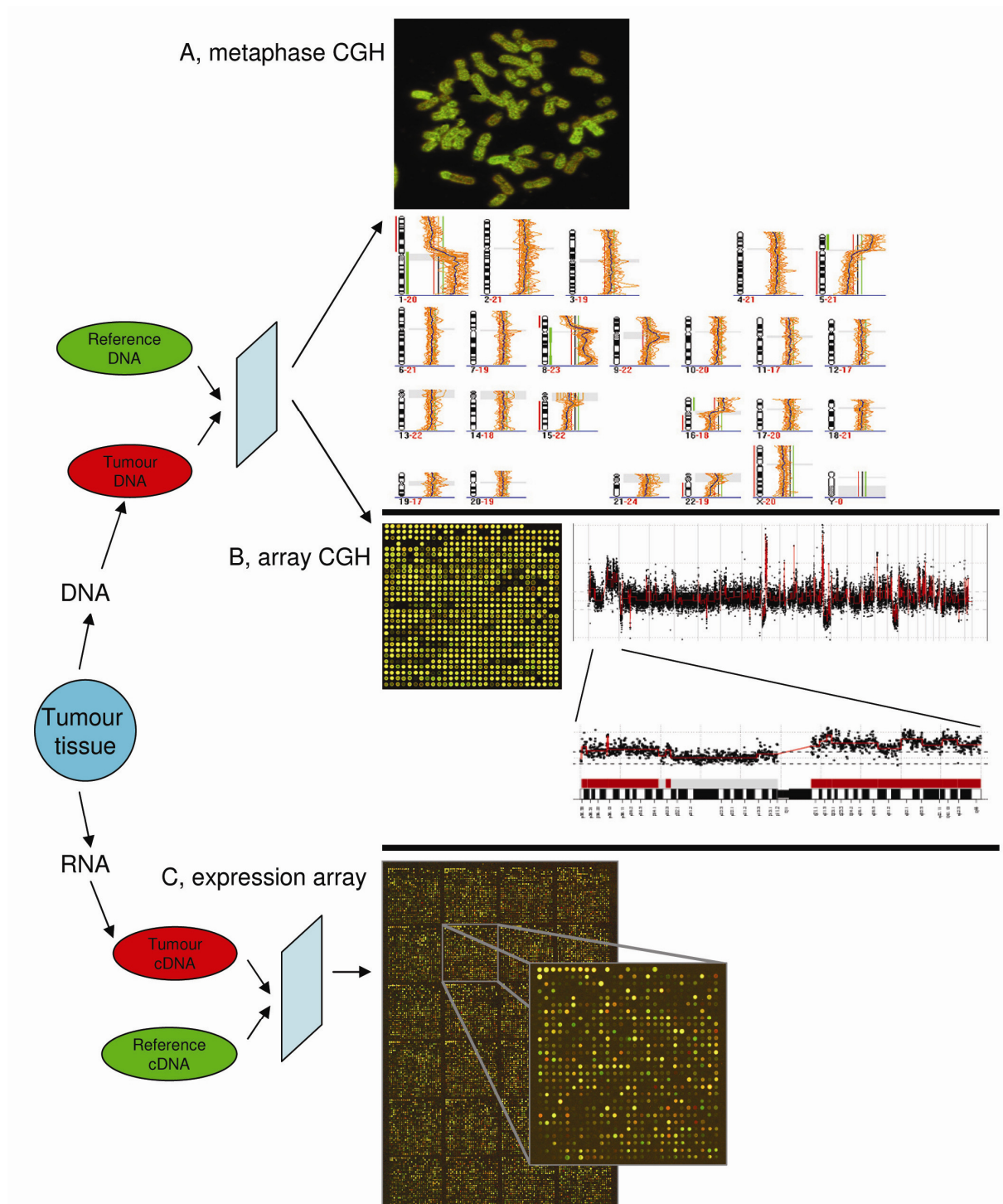


Figure 3. Schematic overview of metaphase CGH (A), array CGH (B) and gene expression microarray (C). DNA and RNA were extracted from tumour samples. The tumour DNA were analysed by metaphase CGH and array CGH. Differently labelled test and reference DNA were co-hybridised to glass slides that were coated with metaphase chromosome spreads (metaphase CGH) or spots containing BAC-DNA (array CGH). The slides were photographed or scanned and picture analysis was performed to generate genomic profiles. In the case of gene expression microarray, (C) RNA was converted into cDNA and simultaneously labelled by reverse transcription. Differently labelled tumour and reference cDNA were co-hybridised to glass slides containing cDNA oligonucleotides. It is also visible comparing the two array pictures that in (B), array CGH, almost every spot has a strong signal, as a reflection of the normal DNA condition of two copies of each fraction of the genome, whereas in expression array (B) the strength of the signals varies largely reflecting the diversity of gene expression.

MATERIALS AND METHODS

TISSUE ARRAY

In tissue microarrays, one specific protein antibody or DNA probe is hybridised to several tumour samples. These tissue arrays are generally produced in house, since tumour tissues are not commercially available. The arrays are used to easily and cost-efficiently explore primarily protein expression but also copy number levels of genes in a quantity of tissue samples.

In **Paper IV**, tissue microarrays were used to evaluate some of the findings from **Papers II** and **III** in a large set of new breast tumours. Four antibodies targeting proteins representing four different genes were tested for their significance in 5-year survivors and deceased patients. In brief, the tissue microarray slides were deparaffinised and autoclaved for at least one hour. The immunohistochemical staining was performed in an automated immunostainer. The microarray slides were incubated with the different antibodies, at a dilution of 1:300 for ADIPOR1, 1:500 for ADORA1, 1:1000 for BTG2; and 1:40 for CD46. The antibodies were visualised by the EnVision K5007 or LSAB K5007 visualisation system and then, the slides were washed in water, dehydrated and mounted. A pathologist evaluated the protein expression.

Quantitative Real Time PCR

Quantitative Real Time PCR (QPCR) is a technique that amplifies and simultaneously quantifies specific DNA or RNA sequences in a semi-quantitative fashion. By using gene specific primers and light emitting probes, the start quantity of DNA or converted RNA is measured during a PCR reaction, simply by measuring the amount of PCR cycles that are needed to reach a particular amount of DNA.

QPCR was used in **Paper III** to validate the differences in expression levels of fourteen genes that were differentially expressed in survivor tumours compared to the tumours from deceased patients in the EA study. We used the same RNA as in the EA experiment for all tumours but four, due to lack of access to material. For each tumour, cDNA was synthesised from total RNA. Commercially available validated TaqMan® Gene Expression Assays were used on triplicates of the samples and thermal cycling was performed with an initiation step at 95°C for 10 minutes, followed by 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. In each assay, a 2-fold dilution series of five samples (1:2, 1:4, 1:8, 1:16, 1:32) was used to be able to quantify the expression levels of the genes of interest according to the standard-curve

method. All samples were normalised to the geometric mean of two endogenous controls; *PPIA* and *PTER*.

Statistics

Generally, the two-tailed Student's t-test was used to evaluate the difference in number of chromosomal aberrations between survivors and deceased patients in **Paper I** as well as difference in gene expression for each gene between 10-year survivors and deceased patients in **Paper III**. In the gene expression analysis, we used a cut-off value of $P < 0.001$. We used this low P -value instead of correction for multiple testing, in order to avoid elimination of true positive genes. This means that the gene list we developed could have a interference of approximately 16 false positives. Nevertheless, when evaluated in an independent tumour set, the gene-list classified the independent samples well. A one-tailed Student's t-test was used to determine the difference in gene expression between the survival groups in the QPCR analysis in **Paper III**. The P -values for differences in frequency of each chromosomal aberration between the survival groups were calculated using the two-tailed Fisher's exact test in both **Papers I** and **II**. In addition, the two-tailed Fisher's exact test was used to evaluate the significance of differential protein expression between 5-year survivors and deceased patients in **Paper IV**. Kaplan Meier survival curves were produced in the SPSS version 16 software. P -values for the differences in survival between samples with or without the detected molecular characteristic (CNA, protein expression) were calculated using the Breslow-Wilcoxon test [50].

In both array studies, in **Papers II** and **III**, the first steps of data analysis were performed in BASE (BioArray Software Environment), and further information about statistics is available in the individual papers [51, 52].

RESULTS AND DISCUSSION

In this thesis, four papers are included. In the first two papers, we searched for genomic prognostic biomarkers. Then we proceeded by examining gene expression in relation to survival in the following paper, and in the last paper we studied protein expression in relation to long-term survival. Thus, we started with DNA, transcribed further to RNA and finally translated into protein, like in the living cell.

Genomic level

In **Papers I and II** we aimed to identify DNA copy number alterations (CNAs) that differed in frequency between 10-year survivors and deceased patients. We used metaphase CGH on a tumour set consisting of 57 primary node-negative breast tumours in **Paper I**, and continued with high-resolution array CGH in **Paper II** using 39 of the samples from **Paper I** plus two additional tumours. In both studies, a number of CNAs exhibited statistical significance. Gains at 4q, 5q31-5qter, 6q12-6q16, and 12q14-12q22 and losses of 17p, 18p and Xq were associated with decreased survival in **Paper I**. In addition, losses of four regions at 8p and Xp were associated with decreased survival in **Paper II** (Figure 4), and gains of two regions at 1q were more common in the tumours from 10-year survivors in **Paper II**, and one of the 1q regions region was also validated as a marker of 10-year survival in an independent dataset [53].

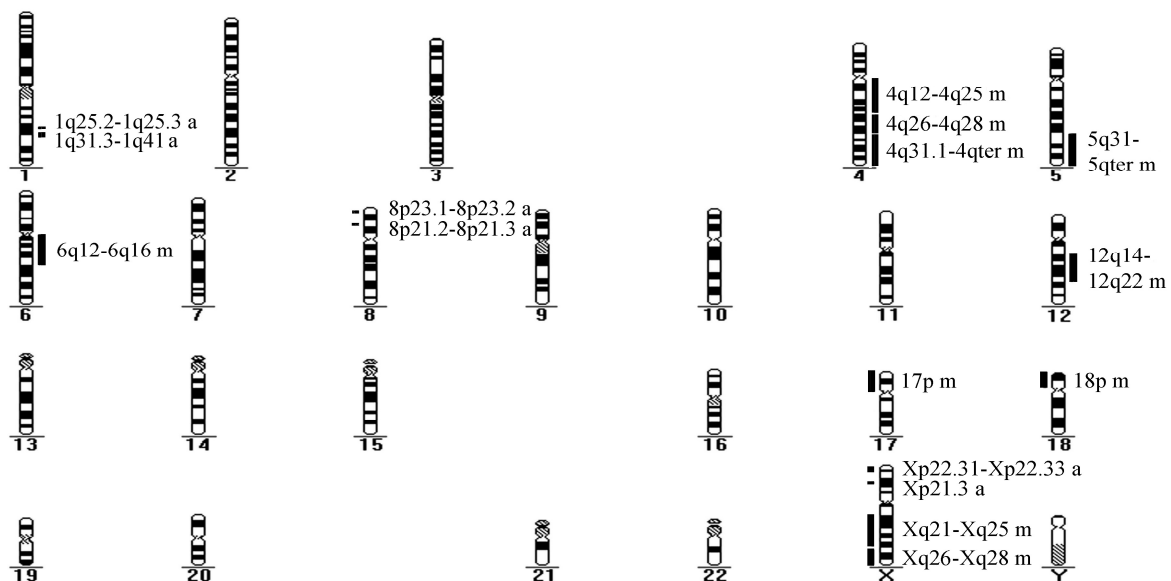


Figure 4. The significant CNAs in the aCGH and mCGH studies. Black bars to the left of the chromosome represent gains and to the right losses. CNAs from the aCGH study are marked with a, CNAs from the mCGH study are marked with m.

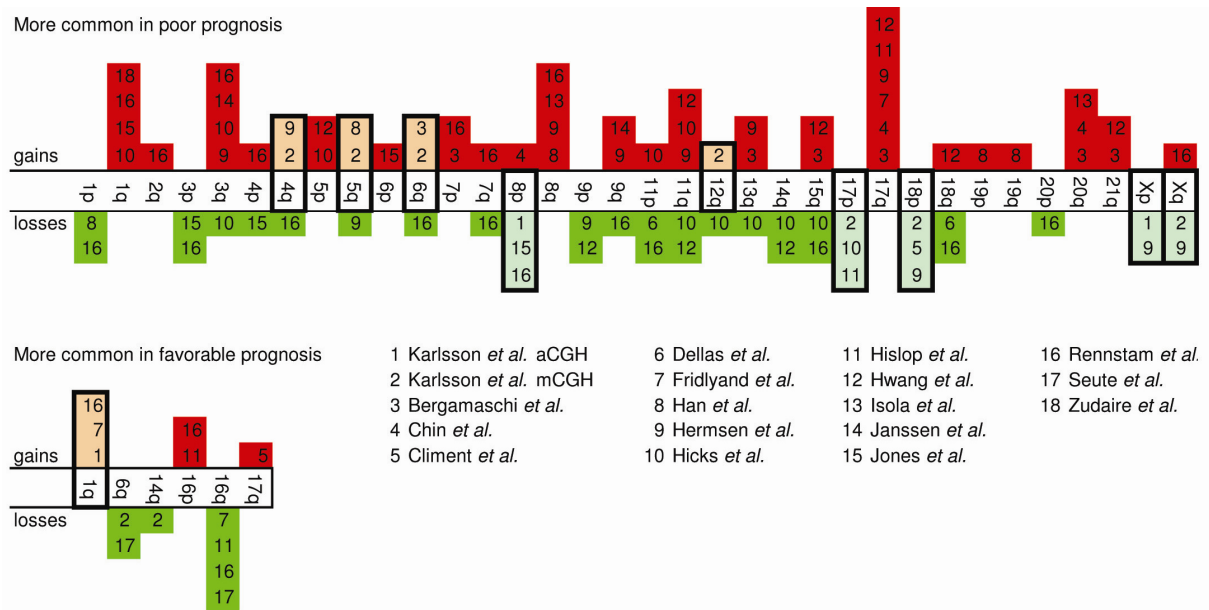


Figure 5. Compilation of prognostic CNAs revealed in 18 CGH studies of genetic alterations in association with clinical outcome. Red bars represent gains, and green bars represent losses. The CNAs detected in any of our two studies are highlighted with lighter red or lighter green.

Several studies have previously searched for CNAs with prognostic qualities in breast cancer using both mCGH [54-63] and aCGH [53, 64-68], of which some of the mCGH studies explored tumours exclusively from node-negative patients [54-57]. Various CNAs have been suggested as prognostic markers, illustrated in Figure 5. Most of the CNAs associated with prognosis were more common in tumours from patients with poor outcome, but some CNAs were associated with a favourable prognosis. Furthermore, the most commonly detected CNA with prognostic value is a gain at 17q, which has been correlated with poor outcome, with a minimal region of overlap at 17q12. This region was gained in approximately 20% of the samples in both of our CGH studies and did not differ between tumours from 10-year survivors and deceased patients. Generally, the concordance between the studies is low, indicating that finding CNAs that serve as prognostic markers is relatively difficult. The low concordance could be due to differences in study design, how different tumour materials are selected and quality of hybridisations. Anyhow, if successful, it would be an advantage working with the stable and uncomplicated DNA as compared with the more unstable RNA or the much more diverse proteins.

The mCGH study in **Paper I** revealed that gains at 4q, 5q31-5qter, 6q12-6q16, and 12q14-12q22 and losses of 17p, 18p and Xq were significantly more common in tumours from deceased patients than in tumours from 10-year survivors. All of these CNAs, with the exception of 12q14-12q22, have been implicated to have prognostic value for breast cancer patients in at least one previous study [55, 59, 64, 66, 67], as shown in Figure 5. In the aCGH study in **Paper II**, CNAs in six other regions showed

RESULTS AND DISCUSSION

statistical significance between tumours from deceased patients and 10-year survivors. Losses at 8p21.2-8p21.3, 8p23.1-8p23.2, Xp21.3 and Xp22.31-Xp22.33 were more common in tumours from deceased patients. In prior studies, both losses at 8p [60, 61], and loss of chromosome X [55], have been detected as a sign of poor clinical outcome, which is in concordance with this study. Interestingly, gains in two regions at 1q were significantly more common in the survivor tumours. Tumours with gains on chromosome 1q in combination with loss of 16p have in previous studies been suggested to represent a group of patients with better prognosis [61, 65]. The 1q chromosome arm was frequently altered in the entire material in our aCGH study, and has also, in contrast to this study, been implicated as an indicator of poor outcome in breast cancer [60, 61, 63, 67], which makes the interpretation of this CNA somewhat difficult. Since the 1q region is one of the most frequent genetic alterations in breast cancer, it is possible that different studies randomly get different impact of this CNA, due to diverse sample setup. However, were we able to verify the difference we detected at 1q31.3-1q41 in an independent tumour material [53].

Table 3. The CNAs showing significant differences in frequency between tumours from deceased patients and 10-year survivors in the mCGH and aCGH experiments. The *P*-values were calculated using both Fischer's exact test and Breslow-Wilcoxon calculation. A, the CNAs that attained statistical significance between the survivor groups in the mCGH study. *P*-values were calculated for the data from the aCGH study as well, in order to evaluate whether these CNAs had prognostic impact using the aCGH method. Each sample was designated to have the specific CNA if the CNA were detected in any of the clones within the region; hence, there might be regions with lower *P*-values within the regions. B, the CNAs that attained statistical significance between the survivor groups in the aCGH study. *P*-values were calculated for the data from the mCGH study as well, in order to evaluate whether these CNAs had prognostic impact using the mCGH method. The four regions at 8p and Xp were only represented by 8p21-8pter and Xp in the mCGH study and thereby are these four regions only represented by values for two regions.

mCGH regions region		Metaphase CGH				Array CGH			
event	<i>P</i> Fisher's	<i>P</i> Breslow	10-year surv. (%)	Dead (%)	<i>P</i> Fisher's	<i>P</i> Breslow	10-year surv. (%)	Dead (%)	
4q12-4q25	gain	0.0020	0.00050	11	5	0.73	0.26	23	32
4q26-4q28	gain	0.031	0.014	14	41	0.70	0.98	23	16
4q31.1-4qter	gain	0.027	0.0044	3	14	0.76	0.56	41	47
5q31-5qter	gain	0.019	0.0079	0	18	1.00	0.79	36	32
6q12-6q16	gain	0.035	0.013	9	32	1.00	0.96	27	26
12q14-12q22	gain	0.021	0.018	11	41	0.74	0.81	27	37
17p	loss	0.047	0.014	54	82	1.00	0.47	45	47
18p	loss	0.014	0.0025	14	45	0.49	0.18	23	37
Xq21-Xq25	loss	0.019	0.000001	0	18	0.11	0.015	23	47
Xq26-Xqter	loss	0.0062	0.000001	0	23	0.12	0.0012	9	32

aCGH regions region		Metaphase CGH				Array CGH			
event	<i>P</i> Fisher's	<i>P</i> Breslow	10-year surv.	Dead (%)	<i>P</i> Fisher's	<i>P</i> Breslow	10-year surv.	Dead (%)	
1q25.2-1q25.3	gain	0.78	0.83	69	73	0.029	0.018	43	26
1q31.3-1q41	gain	1.00	0.78	57	55	0.037	0.028	86	47
8p21.2-8p21.3	loss	0.42	0.15	46	59	0.026	0.0012	9	42
8p23.1-8p23.2	loss					0.037	0.00021	23	63
Xp21.3	loss					0.0022	0.00051	0	37
Xp22.31-Xp22.33	loss	0.70	0.13	11	18	0.026	0.00067	14	42

The CNAs detected in the mCGH study were not statistically significant in the array study and vice versa using Fisher's exact test, although the studies were performed on partly the same material. However, when comparing the survival rates of patients with tumours with or without the CNA, the regions on Xq detected in the mCGH study were of statistical significance in the aCGH material as well using a Breslow-Wilcoxon test (Table 3). Some of the other mCGH CNAs (4q12-4q25, 18p) showed differences between 10-year survivors and deceased patients in the aCGH study, although the differences were not statistically significant (Table 3). The four regions at 8p and Xq detected in the aCGH study were only represented by two regions in the mCGH study, and these did show differences between survival groups in the mCGH study, although not statistically significant (Table 3). All CNAs detected in one of the studies were also detected in the other although did not differ significantly between 10-year survivors and deceased patients as seen in Table 3. The discrepancy between the two studies could possibly be explained by the high resolution of aCGH that allows specific regions and distinct breakpoints to be detected. In metaphase CGH, each chromosome arm was divided into only one to three sub-regions before evaluation, generating large CNAs that sometimes in reality would be relatively small, with non-specific breakpoints. In general, it is difficult to identify CNAs with prognostic value in that can be verified in independent breast cancer materials. However, the CNAs can still be interesting for further investigation, both as prognostic markers themselves, but also as a way to find specific genes to explore further.

Gene expression level

In **Paper III**, we wanted to identify a set of genes whose expression could predict long-term survival in node-negative breast cancer patients. We used expression microarrays (EAs) and found that a set of 51 genes could predict 10-year survival with great certainty in our tumour set (Figure 6a). (The specific 51 genes is found in **Paper III** [19].) None of the tumours from deceased patients was classified to belong to the favourable prognosis group and only five survivor tumours were misclassified into the poor prognosis group, which results in an accuracy of 89%. Furthermore, since none of the tumours classified with a favourable prognosis came from a deceased patient, this classifier could assist in the selection of patients that do not require further treatment. It is preferable to provide post-surgical treatment to more patients than needed, compared to not treating patients that would have benefited from further treatment. An attempt to evaluate the 51 gene list in a tumour

RESULTS AND DISCUSSION

material from a previous EA study by van't Veer and colleagues was performed [26]. This study consisted of 78 tumours from node-negative patients whose disease relapsed or not within five years from diagnosis. The list of 51 genes generated good results in this material as well (Figure 6b), with an accuracy of 74%. Most of the misclassifications were in the poor prognosis group, and only five tumours classified in the favourable prognosis group were from patients whose disease relapsed within five years of diagnosis. In the data from an EA study by Wang *et al.* [27], the results of classification were moderate to poor, probably due to the absence of 28 of the 51 genes in the Wang data set. Many of the genes in the list of 51 genes have previously been implicated in cancer, such as the *BCAT1*, *CCNB1IP1*, *CUL7*, *E2F2*, *GGH*, *GIT2*, *NEIL1*, *SALL4*, *SERPINB9* and *TM4SF5* genes [69-80].

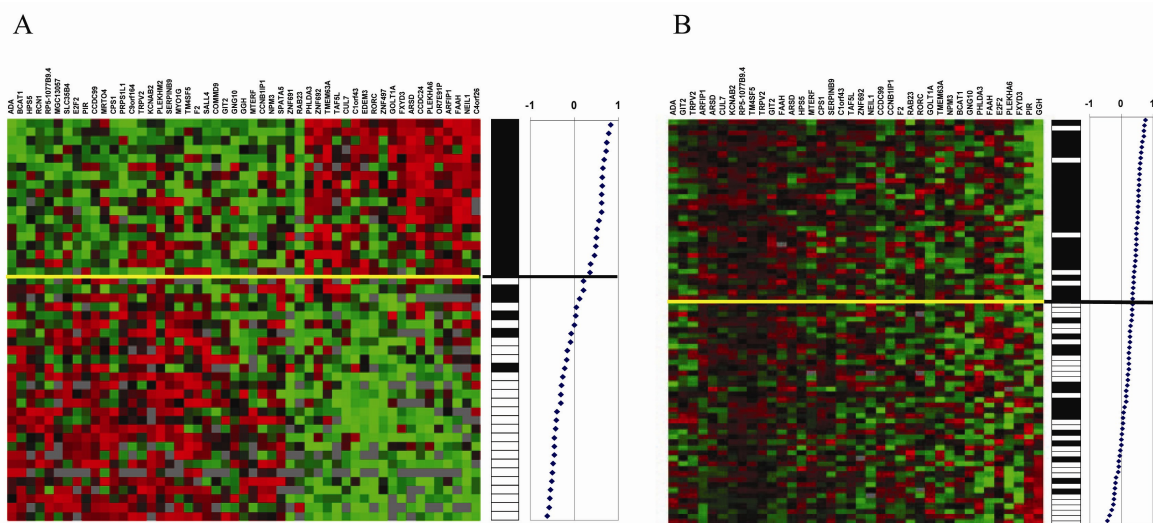


Figure 6. Correlation-based classification using the list of 51 genes. A, Classification of our tumours using the 51 genes shows 89% accuracy and no tumours from deceased patients were in the favourable prognosis group. B, Classification tumours analysed by van't Veer *et al.* using our gene list shows 74% accuracy and only 5 tumours from deceased patients in the favourable prognosis group. In A, black bars represent 10-year survivors while white bars represent patients that died within ten years from diagnosis. In B, black bars represent patients that were metastasis free for five years, while white bars represent patients that developed metastasis within five years. Plots to the right show the correlation between each tumour's expression profile and the favourable prognosis profile.

One approach to study gene expression in breast tumours is to use gene expression profiles to cluster the tumours into at least four molecular subgroups; Luminal A, Luminal B, basal and HER-2 positive [11, 28-31]. The subgroups differ in tumour behaviour and survival rate, and this way of exploring breast tumours has a probable clinical use, and seems rather robust. However, we chose to explore survival specifically, independent of molecular subgroup, since survival by definition is a central factor in breast cancer that can easily be brought to clinical use. The outcome approach has been extensively utilised in breast cancer [12-18, 20-27]. Of these investigations, a few have addressed exclusively node-negative samples [25-27]. In general, few genes are involved in several of the produced gene-lists, and

none of the genes in our list is included in any of the other lists suggested for node-negative patients [25-27]. van't Veer *et al.* performed the most renowned study in this area, where they have found that the expression signature of 70 genes, called "MammaPrint", could predict recurrence free survival [26]. This 70 gene set has been verified in several studies [81-84], and is presently used in a clinical trial involving 6 000 breast cancer patients [33]. Interestingly, when using this 70 gene set to predict outcome in our tumour material, approximately 70% of the tumours were correctly classified, and even though this is a quite good result, our gene set was slightly better in classifying van't Veer's tumour set, then their gene set was in classifying our tumour material. This is worth consideration since even if the gene set generated by van't Veer *et al.* might work sufficiently well in the clinic, it might still not be the most efficient gene set available. In addition, Paik *et al.* have done a QPCR study that identified a set of 21 genes, "Oncotype DX RS", where the expression can predict recurrence in tamoxifen treated, oestrogen receptor positive, node-negative breast tumours [24]. This expression profile has been validated [85] and is presently in a clinical trial where it is used to assist the choice of treatment [32]. None of these genes was present in our 51-gene list.

Our list of 51 identified genes could predict clinical outcome in our material with great certainty. It could predict clinical outcome in van't Veer's material as well, but not in Wang's material, probably due to the low number of genes found in Wang's material. Overall, our gene set worked similarly well in classifying van't Veer's material as their gene set on our material, slightly better considering the number of deceased patients/patients with recurrent disease in the favourable prognosis groups. Furthermore, the list of 51 genes might contain specific genes interesting for clinical outcome in breast cancer as well as being a good prognostic gene-set. Additional studies using larger sets of tumours are needed to define the significance of these genes during the genesis of lymph node-negative breast tumours.

Protein expression level

In **Paper IV**, we wanted to analyse the expression of four proteins in association with patient survival (ADIPOR1, ADORA1, BTG2 and CD46) since we found the corresponding genes to be differ significantly between 10-year survivors and deceased patients, both on copy number level in **Paper II** and gene expression level in **Paper III**.

RESULTS AND DISCUSSION

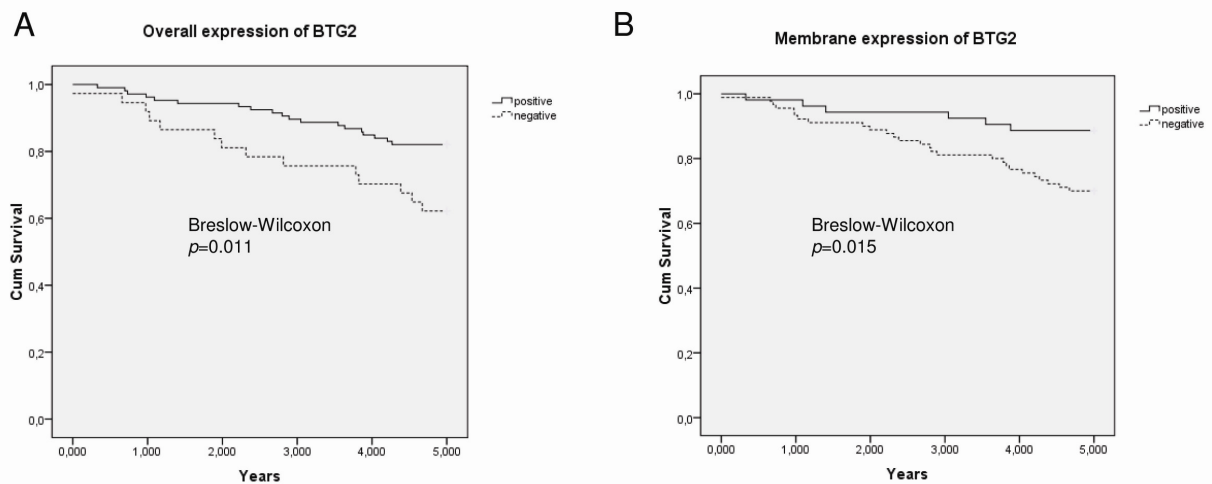


Figure 7. Kaplan-Meier survival curves illustrating the effect of BTG2 expression. A, shows the difference in survival between patients with tumours that revealed overall BTG2 expression and patients whose tumours did not, whereas B, shows the difference in survival between patients with tumours that revealed cell membrane specific BTG2 expression and patients whose tumours did not. The solid line represents patients whose tumours expressed BTG2 and the dashed line represents patients whose tumours did not. The *P*-values for the difference between the curves were calculated using a Breslow-Wilcoxon test.

The major finding was that the BTG2 protein was expressed significantly more frequently in tumours from 5-year survivors compared with tumours from deceased patients. The *P*-values for differential expression between the survival groups were below 0.05 but above 0.001 in the expression microarray study, and the genes are thereby not included in the list of 51 genes. BTG2 protein expression was detected both in the cytoplasm, and in the cellular membrane, and the overall expression of BTG2 differed significantly between survivors and deceased patients ($P=0.026$) using the Fisher's exact test, although the significance for specific membrane expression was even stronger ($P=0.013$) (Figure 7, Table 4). *P*-values were also calculated using the Breslow-Wilcoxon test and then overall expression of BTG2 showed stronger significance ($P=0.011$ versus $P=0.015$).

None of the other three analysed proteins (ADIPOR1, ADORA1 and CD46) revealed a statistically significant impact on overall survival (Table 4). Within the group of node-negative patients 55% of the deceased patients expressed BTG2 and 22% displayed membrane specific expression, compared to 81% and 51%, respectively in the 5-year survival group. However, these differences were not statistically significant (overall expression; $P=0.10$, membrane specific expression; $P=0.16$). This was probably due to the low number of deceased patients within the group of node-negative patients, only ten of which one showed large tissue loss on the slide and were thereby not possible to analyse.

Table 4. Differences in protein expression in tumours from 5-year survivors and deceased patients. *P*-values were calculated using a two-tailed Fisher's exact test. The samples designated as not available had few tumour cells, large tissue loss or affluence of necrotic tissue.

Protein	deceased patients		5 year survivors		not available	<i>P</i> -value 5-year survival
	positive (%)	negative (%)	positive (%)	negative (%)		
AdipoR1	26	74	17	83	3	0.29
Adora1	30	70	23	77	8	0.47
BTG2	61	39	82	18	6	0.026*
-membrane only	19	81	44	56	6	0.013*
-cytoplasm only	52	48	68	32	6	0.14
CD46	16	84	14	86	4	0.77

In the present investigation, we analysed 5-year survival instead of 10-year survival, simply because the tumours were collected between 2001 and 2002 and 10-year survival was consequently unattainable. The four proteins were selected based on genes that differed between 10-year survivors and deceased patients on both gene expression and DNA copy number levels. If a threshold of $P < 0.05$ was applied to the EA data in **Paper III**, 27 genes located within the CNAs detected in the aCGH study in **Paper II** were identified. On the basis of their involvement in cancer and the availability of commercial antibodies, the *ADIPOR1*, *ADORA1*, *BTG2* and *CD46* genes were selected among the 27 previously identified genes, to further investigate the of association of protein expression levels to patient survival.

The BTG2 protein, that was significantly more frequently expressed in tumours from 5-year survivors is a known tumour suppressor gene [86-88], that is directly regulated by p53 and involved in the p53-mediated response to DNA damage [89]. BTG2 is involved in cell cycle arrest in the transition from G1 to S phase [90, 91]. In addition, BTG2 can regulate G2/M cell cycle arrest independent of p53 [92, 93]. Down-regulation of BTG2 has been detected in several cancer types such as prostate cancer, breast cancer and gliomas [94-96]. In contrast to previous reports, 78% of the breast cancer samples showed moderate to strong expression of BTG2 in the majority of tumour cells in this study. Nevertheless, BTG2 was significantly down regulated in tumours from deceased patients compared with tumours from 5-year survivors, both in overall expression and cell membrane specific expression. This finding suggests that; high total BTG2 or specific cell membrane expression may contribute to an increase in survival. One previous study analysed BTG2 protein expression and correlated decreased nucleus expression to a more aggressive phenotype of breast cancer, even though they did not detect a significant difference in survival [97]. The discrepancy could be due to the use of different BTG2 antibodies since Kawakubo et al. detected staining predominantly in the nucleus. However, our finding supports

RESULTS AND DISCUSSION

the theory that down-regulation of BTG2 contributes to a more malignant behaviour of the cells. Moreover, BTG2 may act as a prognostic marker in breast cancer.

CONCLUDING REMARKS

In these four papers, we found that specific cytogenetic alterations, expression of particular genes, and protein expression of *BTG2* differed between long-term survivors and patients that died from breast cancer. Since most of the findings could be validated in external materials, we believe that these types of studies, scrutinising molecular characteristics of tumours removed from patients with long-term follow up, could contribute to the discovery of novel prognostic markers that could help improve the clinical prognosis of patients and thereby improve breast cancer treatment.

Differences in copy number alterations on the genomic level were identified between 10-year survivors and deceased patients. Using metaphase CGH, we found that gains at 4q, 5q31-5qter, 6q12-6q16, and 12q14-12q22 and losses of 17p, 18p and Xq were significantly more common in tumours from deceased patients. By using the higher resolution array CGH, different copy number alterations were correlated to 10-year survival; losses at 8p21.2-8p21.3, 8p23.1-8p23.2, Xp21.3 and Xp22.31-Xp22.33 were more common among deceased patients, and gains at 1q25.2-1q25.3 and 1q31.3-1q41 were associated with increased 10-year survival. Copy number alterations at specific chromosome regions detected as significant for patient survival, indicate that genes located within these regions can also be altered in expression, which could influence the aggressiveness of the tumour.

Using expression arrays, a list of 51 genes was established that could predict 10-year survival with great certainty in our patient material. The 51-gene list could also predict recurrence free survival in a data set from an external breast cancer study with high accuracy, indicating that this list may offer a good potential in the clinic to predict survival as well as indicate node-negative breast cancer patients that would not benefit from post surgical treatment.

The protein expression levels of the *BTG2* gene differed in expression between 5-year survivors and deceased patients. *BTG2* is a tumour suppressor gene but this is the first study correlating *BTG2* protein expression to patient survival in breast cancer. The *BTG2* gene also differed between 10-year survivors and deceased patients in copy number and in gene expression in the array CGH and expression microarray study, and thereby, the expression of *BTG2* may be a potential biomarker of long-term survival in breast cancer in general.

FUTURE PERSPECTIVES

Breast cancer is a very common disease, affecting over a million women every year. The current prognostic factors used to select treatment for breast cancer patients are, however, insufficient, and therefore, there is a great need of additional prognostic and predictive markers in order to further tailor and optimise treatment for the individual breast cancer patient.

Our work in this thesis provides some possible biomarkers that in the future may facilitate the classification of breast cancer patients into risk groups. Although, there is a lot of work to be done prior to clinical use; copy number alterations, gene expression profile and level of BTG2 expression needs extensive further validation. Therefore, it would be of great interest to further evaluate whether any of the CNAs revealed, the gene expression profile we generated as well as the expression of BTG2 has impact on patient survival in independent cohorts of breast cancer tumours. In addition, the list of 51 genes contains genes whose protein expression would be interesting to study in correlation to patient survival. Another fascinating study design would be to analyse whether the number of BTG2 gene copies in tumours are related to patient survival.

SAMMANFATTNING PÅ SVENSKA

Bröstcancer är den vanligaste cancerformen bland kvinnor som drabbar ca 7 000 kvinnor i Sverige per år [2]. I världen drabbas drygt 1 miljon kvinnor per år och ca 400 000 kvinnor avlider varje år till följd av bröstcancer [1]. Överlevnadsfrekvensen har ökat dramatiskt de senaste åren till följd av tidigare upptäckt och nya behandlingsmetoder [5]. I Sverige är femårsöverlevnaden ungefär 86% [2], medan den totala överlevnadssiffran för bröstcancer är ca 73% [1]. En stor andel bröstcancerpatienter idag överbehandlas och vissa underbehandlas. Detta beror på att de markörer som används idag är otillräckliga för att bedöma vilka som är högrisk- respektive lågrisk-patienter. För att finna biomarkörer som skulle kunna underlätta riskbedömningen av bröstcancerpatienter, har vi undersökt genetiska förändringar, genuttryck samt proteinuttryck i tumörer från patienter som har överlevt fem eller tio år och patienter som avlidit till följd av bröstcancer.

Vi identifierade skillnader mellan de två patientgrupperna på cytogenetisk nivå. Med hjälp av metafase-CGH fann vi att ökning av genetiskt material på kromosom 4q, 5q31-5qter, 6q12-6q16, och 12q14-12q22 liksom förluster av 17p, 18p and Xq var signifikant vanligare i tumörer från avlidna patienter. När vi använde den mer högupplösta metoden array-CGH var förlust av 8p21.2-8p21.3, 8p23.1-8p23.2, Xp21.3 and Xp22.31-Xp22.33 vanligare i tumörer från avlidna patienter, medan ökning av 1q25.2-1q25.3 and 1q31.3-1q41 vanligare hos 10-årsöverlevare.

På genuttrycksnivå fann vi att uttrycket av 51 gener kunde klassa våra patienter med hög säkerhet i två prognos grupper; god prognos och sämre prognos. Med hjälp av dessa geners uttryck kunde vi även med goda resultat klassa ett material från en tidigare studie [26]. Denna genlista skulle kunna användas i kliniken, dels som prognostiskt verktyg och dels som hjälpmedel för att avgöra vilka patienter som behöver vidare behandling.

När vi undersökte resultaten från array-CGH och genuttrycksförsöken såg vi att 27 gener skilde sig signifikant på både genetisk nivå och genuttrycksnivå. Vi undersökte uttrycket av fyra av dessa proteiner i ett nytt och större tumörmaterial och fann att BTG2 uttrycks mer frekvent hos patienter som överlevt i minst fem år efter diagnos jämfört med patienter som avlidit inom fem år. BTG2 har tidigare beskrivits som en tumörsuppressor [86-88], men inte tidigare som prognostisk markör. Vi anser att BTG2 kan vara en lovande prognostisk markör, dock krävs ytterligare bekräftande undersökningar innan den tas i kliniskt bruk.

ACKNOWLEDGEMENTS

ACKNOWLEDGEMENTS

This study has been performed at the Department of Oncology, University of Gothenburg, Sweden and was supported by King Gustav V Jubilee Clinic Cancer Research Foundation and the Swedish Cancer Society.

I would like to thank:

My supervisor **Khalil** for introducing me to the interesting field of molecular breast cancer, for sharing your knowledge and for always being there for me. Thank you for your support, for providing important contacts and for helping me through the difficulties that sometimes occur during the work of a thesis.

Per, my co-supervisor, for your enthusiasm, interesting discussions and for sharing your valuable knowledge in the clinical breast cancer field.

My co-worker **Ulla**, you have been a wonderful lab-partner and friend through my time in the lab. I know no one more efficient and professional in the lab than you.

Anna and **Toshima**, my group-members for lovely travel-company and scientific discussions. Anna, you are a great lab-partner, and Toshima, thank you for your excellent proofreading of many of my texts.

My co-authors; **Björn**, for providing your invaluable knowledge in statistics, **Frida** for helping us with the Q-PCR in Paper III, **Kristina** for helping us with the IHC in Paper IV, **Aniko** for finding time to competently evaluate the IHC, **Karin** and **Donal** for providing the TMAs in Paper IV.

Mårten Fernö and **Åke Borgs group** for rewarding collaboration.

Lovisa, you are an amazing co-worker, always caring and supporting and in addition, you are a lot of fun, thank you for always being there for me.

Karin E, **Nina**, **Karin M** and **Karolina**, thanks to you I have the great advantage of meeting my friends as I go to work. You have made it a lot easier for me to get up in the mornings.

Ghita for evaluating the imprints of the tumours used in Paper I-III as well as your helpfulness at any time in the lab.

ACKNOWLEDGEMENTS

Lisbet for helping producing the metaphase preparations and for being warm and caring.

Ragnar for your enthusiasm and interest, and for valuable discussions and support.

The staff at Onkologlab; **Kecke, Kristina, Ingela, Ingegerd, Helena, May, Kerstin, Åse, Jörgen, Giggi, Elisabeth, Arne, György** and **Claes** for your endless help, support and friendship.

Sara and **Bozza**, without the two of you I am not sure that I would have begun to feel at home so easily here in Gothenburg. **Sara**, thank you for proofreading my thesis, you are a wonderful and reliable friend who always support me and knows exactly what I need to hear.

Anna, Tex, Peppe, Tobbe, Lisa and **Sara-Johanna**, thank you for being such amazing friends and for bringing lots of joy and excitement into my life.

The rest of my friends, I am happy to have you all in my life.

Bissan, Björn, Agneta, Jenny, Måns, Petter, Åsa, Teodor, Anton and **Herman**, thank you for being such a supportive and lovely family in law.

Mom, Dad and **Malin**, you have always relied in me making me believe in myself. You are the best family anyone could ever wish for. I love you.

Milla, my amazing daughter. Every day with you is an adventure. I love you more than anything in the world.

And last, **Rickard**, for standing by my side through all those years. Thank you for your love, calmness and stability, which seem to be exactly what I need. Thank you for all your support, making this thesis possible. You are the love of my life.

REFERENCES

1. Parkin DM, Bray F, Ferlay J, Pisani P: **Global cancer statistics, 2002.** *CA Cancer J Clin* 2005, **55**(2):74-108.
2. Socialstyrelsen, Cancerfonden: **Cancer i siffror 2005**, vol. 2006. Stockholm: Socialstyrelsen, Cancerfonden; 2005.
3. Irigaray P, Newby JA, Clapp R, Hardell L, Howard V, Montagnier L, Epstein S, Belpomme D: **Lifestyle-related factors and environmental agents causing cancer: an overview.** *Biomed Pharmacother* 2007, **61**(10):640-658.
4. Hanahan D, Weinberg RA: **The hallmarks of cancer.** *Cell* 2000, **100**(1):57-70.
5. Berry DA, Cronin KA, Plevritis SK, Fryback DG, Clarke L, Zelen M, Mandelblatt JS, Yakovlev AY, Habbema JD, Feuer EJ: **Effect of screening and adjuvant therapy on mortality from breast cancer.** *N Engl J Med* 2005, **353**(17):1784-1792.
6. Goldhirsch A, Wood WC, Gelber RD, Coates AS, Thurlimann B, Senn HJ: **Progress and promise: highlights of the international expert consensus on the primary therapy of early breast cancer 2007.** *Ann Oncol* 2007, **18**(7):1133-1144.
7. Carter P, Presta L, Gorman CM, Ridgway JB, Henner D, Wong WL, Rowland AM, Kotts C, Carver ME, Shepard HM: **Humanization of an anti-p185HER2 antibody for human cancer therapy.** *Proc Natl Acad Sci U S A* 1992, **89**(10):4285-4289.
8. Michaelson JS, Silverstein M, Sgroi D, Cheongsiatmoy JA, Taghian A, Powell S, Hughes K, Comegno A, Tanabe KK, Smith B: **The effect of tumor size and lymph node status on breast carcinoma lethality.** *Cancer* 2003, **98**(10):2133-2143.
9. Weinberg RA: **Mechanisms of malignant progression.** *Carcinogenesis* 2008, **29**(6):1092-1095.
10. Mass RD, Press MF, Anderson S, Cobleigh MA, Vogel CL, Dybdal N, Leiberman G, Slamon DJ: **Evaluation of clinical outcomes according to HER2 detection by fluorescence in situ hybridization in women with metastatic breast cancer treated with trastuzumab.** *Clin Breast Cancer* 2005, **6**(3):240-246.
11. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS *et al*: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proc Natl Acad Sci U S A* 2001, **98**(19):10869-10874.
12. Feng Y, Sun B, Li X, Zhang L, Niu Y, Xiao C, Ning L, Fang Z, Wang Y, Cheng J *et al*: **Differentially expressed genes between primary cancer and paired lymph node metastases predict clinical outcome of node-positive breast cancer patients.** *Breast Cancer Res Treat* 2006.
13. Gruvberger-Saal SK, Eden P, Ringner M, Baldetorp B, Chebil G, Borg A, Ferno M, Peterson C, Meltzer PS: **Predicting continuous values of prognostic markers in breast cancer from microarray gene expression profiles.** *Mol Cancer Ther* 2004, **3**(2):161-168.
14. Huang E, Cheng SH, Dressman H, Pittman J, Tsou MH, Horng CF, Bild A, Iversen ES, Liao M, Chen CM *et al*: **Gene expression predictors of breast cancer outcomes.** *Lancet* 2003, **361**(9369):1590-1596.
15. Jenssen TK, Kuo WP, Stokke T, Hovig E: **Associations between gene expressions in breast cancer and patient survival.** *Hum Genet* 2002, **111**(4-5):411-420.
16. Jones C, Mackay A, Grigoriadis A, Cossu A, Reis-Filho JS, Fulford L, Dexter T, Davies S, Bulmer K, Ford E *et al*: **Expression profiling of purified normal human luminal and myoepithelial breast cells: identification of novel prognostic markers for breast cancer.** *Cancer Res* 2004, **64**(9):3037-3045.
17. Korkola JE, Blaveri E, DeVries S, Moore DH, 2nd, Hwang ES, Chen YY, Estep AL, Chew KL, Jensen RH, Waldman FM: **Identification of a robust gene signature that predicts breast cancer outcome in independent data sets.** *BMC Cancer* 2007, **7**:61.
18. Li LF, Xu XJ, Zhao Y, Liu ZB, Shen ZZ, Jin WR, Shao ZM: **Integrated gene expression profile predicts prognosis of breast cancer patients.** *Breast Cancer Res Treat* 2008.
19. Karlsson E, Delle U, Danielsson A, Olsson B, Abel F, Karlsson P, Helou K: **Gene expression variation to predict 10-year survival in lymph-node-negative breast cancer.** *BMC Cancer* 2008, **8**(1):254.

20. Miller LD, Smeds J, George J, Vega VB, Vergara L, Ploner A, Pawitan Y, Hall P, Klaar S, Liu ET *et al*: **An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival.** *Proc Natl Acad Sci U S A* 2005, **102**(38):13550-13555.
21. Naderi A, Teschendorff AE, Barbosa-Morais NL, Pinder SE, Green AR, Powe DG, Robertson JF, Aparicio S, Ellis IO, Brenton JD *et al*: **A gene-expression signature to predict survival in breast cancer across independent data sets.** *Oncogene* 2006.
22. Nuyten DS, Kreike B, Hart AA, Chi JT, Sneddon JB, Wessels LF, Peterse HJ, Bartelink H, Brown PO, Chang HY *et al*: **Predicting a local recurrence after breast-conserving therapy by gene expression profiling.** *Breast Cancer Res* 2006, **8**(5):R62.
23. Onda M, Emi M, Nagai H, Nagahata T, Tsumagari K, Fujimoto T, Akiyama F, Sakamoto G, Makita M, Kasumi F *et al*: **Gene expression patterns as marker for 5-year postoperative prognosis of primary breast cancers.** *J Cancer Res Clin Oncol* 2004, **130**(9):537-545.
24. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T *et al*: **A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer.** *N Engl J Med* 2004, **351**(27):2817-2826.
25. Tsumagari K, Chijiwa K, Nagai H, Makita M, Kasumi F, Akiyama F, Sakamoto G, Miki Y: **Postoperative prognosis of node-negative breast cancers predicted by gene-expression profiling on a cDNA microarray of 25,344 genes.** *Breast Cancer* 2005, **12**(3):166-177.
26. van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT *et al*: **Gene expression profiling predicts clinical outcome of breast cancer.** *Nature* 2002, **415**(6871):530-536.
27. Wang Y, Klijn JG, Zhang Y, Sieuwerts AM, Look MP, Yang F, Talantov D, Timmermans M, Meijer-van Gelder ME, Yu J *et al*: **Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer.** *Lancet* 2005, **365**(9460):671-679.
28. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA *et al*: **Molecular portraits of human breast tumours.** *Nature* 2000, **406**(6797):747-752.
29. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S *et al*: **Repeated observation of breast tumor subtypes in independent gene expression data sets.** *Proc Natl Acad Sci U S A* 2003, **100**(14):8418-8423.
30. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, Livasy C, Carey LA, Reynolds E, Dressler L *et al*: **The molecular portraits of breast tumors are conserved across microarray platforms.** *BMC Genomics* 2006, **7**:96.
31. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, Martiat P, Fox SB, Harris AL, Liu ET: **Breast cancer classification and prognosis based on gene expression profiles from a population-based study.** *Proc Natl Acad Sci U S A* 2003, **100**(18):10393-10398.
32. Sparano JA, Paik S: **Development of the 21-gene assay and its application in clinical practice and clinical trials.** *J Clin Oncol* 2008, **26**(5):721-728.
33. Cardoso F, Van't Veer L, Rutgers E, Loi S, Mook S, Piccart-Gebhart MJ: **Clinical application of the 70-gene profile: the MINDACT trial.** *J Clin Oncol* 2008, **26**(5):729-735.
34. Gerdes J, Lemke H, Baisch H, Wacker HH, Schwab U, Stein H: **Cell cycle analysis of a cell proliferation-associated human nuclear antigen defined by the monoclonal antibody Ki-67.** *J Immunol* 1984, **133**(4):1710-1715.
35. de Azambuja E, Cardoso F, de Castro G, Jr., Colozza M, Mano MS, Durbecq V, Sotiriou C, Larsimont D, Piccart-Gebhart MJ, Paesmans M: **Ki-67 as prognostic marker in early breast cancer: a meta-analysis of published studies involving 12,155 patients.** *Br J Cancer* 2007, **96**(10):1504-1513.
36. Urruticoechea A, Smith IE, Dowsett M: **Proliferation marker Ki-67 in early breast cancer.** *J Clin Oncol* 2005, **23**(28):7212-7220.
37. Mignatti P, Rifkin DB: **Biology and biochemistry of proteinases in tumor invasion.** *Physiol Rev* 1993, **73**(1):161-195.
38. Look MP, van Putten WL, Duffy MJ, Harbeck N, Christensen IJ, Thomssen C, Kates R, Spyrtatos F, Ferno M, Eppenberger-Castori S *et al*: **Pooled analysis of prognostic impact of urokinase-type plasminogen activator and its inhibitor PAI-1 in 8377 breast cancer patients.** *J Natl Cancer Inst* 2002, **94**(2):116-128.

REFERENCES

39. Foekens JA, Schmitt M, van Putten WL, Peters HA, Kramer MD, Janicke F, Klijn JG: **Plasminogen activator inhibitor-1 and prognosis in primary breast cancer.** *J Clin Oncol* 1994, **12**(8):1648-1658.
40. Duffy MJ: **Urokinase plasminogen activator and its inhibitor, PAI-1, as prognostic markers in breast cancer: from pilot to level 1 evidence studies.** *Clin Chem* 2002, **48**(8):1194-1197.
41. Eppenberger U, Kueng W, Schlaeppli JM, Roesel JL, Benz C, Mueller H, Matter A, Zuber M, Luescher K, Litschgi M *et al*: **Markers of tumor angiogenesis and proteolysis independently define high- and low-risk subsets of node-negative breast cancer patients.** *J Clin Oncol* 1998, **16**(9):3129-3136.
42. Pappot H, Gardsvoll H, Romer J, Pedersen AN, Grondahl-Hansen J, Pyke C, Brunner N: **Plasminogen activator inhibitor type 1 in cancer: therapeutic and prognostic implications.** *Biol Chem Hoppe Seyler* 1995, **376**(5):259-267.
43. Kallioniemi OP, Kallioniemi A, Piper J, Isola J, Waldman FM, Gray JW, Pinkel D: **Optimizing comparative genomic hybridization for analysis of DNA sequence copy number changes in solid tumors.** *Genes Chromosomes Cancer* 1994, **10**(4):231-243.
44. Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D: **Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors.** *Science* 1992, **258**(5083):818-821.
45. Osterberg L, Levan K, Partheen K, Helou K, Horvath G: **Cytogenetic analysis of carboplatin resistance in early-stage epithelial ovarian carcinoma.** *Cancer Genet Cytogenet* 2005, **163**(2):144-150.
46. **SCIBLU Genomics Center** [http://www.lth.se/sciblu/services/dna_microarrays]
47. Jonsson G, Staaf J, Olsson E, Heidenblad M, Vallon-Christersson J, Osoegawa K, de Jong P, Oredsson S, Ringner M, Hoglund M *et al*: **High-resolution genomic profiles of breast cancer cell lines assessed by tiling BAC array comparative genomic hybridization.** *Genes Chromosomes Cancer* 2007, **46**(6):543-558.
48. **Swegene DNA Microarray Resource Centre** [<http://swegene.onk.lu.se>]
49. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, Ragg T: **The RIN: an RNA integrity number for assigning integrity values to RNA measurements.** *BMC Mol Biol* 2006, **7**:3.
50. Gehan EA: **A Generalized Wilcoxon Test for Comparing Arbitrarily Singly-Censored Samples.** *Biometrika* 1965, **52**:203-223.
51. Saal LH, Troein C, Vallon-Christersson J, Gruvberger S, Borg A, Peterson C: **BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data.** *Genome Biol* 2002, **3**(8):SOFTWARE0003.
52. **Internal BASE at Lund University** [<http://base.onk.lu.se>]
53. Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, Kuo WL, Lapuk A, Neve RM, Qian Z, Ryder T *et al*: **Genomic and transcriptional aberrations linked to breast cancer pathophysiology.** *Cancer Cell* 2006, **10**(6):529-541.
54. Dellas A, Torhorst J, Schultheiss E, Mihatsch MJ, Moch H: **DNA sequence losses on chromosomes 11p and 18q are associated with clinical outcome in lymph node-negative ductal breast cancer.** *Clin Cancer Res* 2002, **8**(5):1210-1216.
55. Hermsen MA, Baak JP, Meijer GA, Weiss JM, Walboomers JW, Snijders PJ, van Diest PJ: **Genetic analysis of 53 lymph node-negative breast carcinomas by CGH and relation to clinical, pathological, morphometric, and DNA cytometric prognostic factors.** *J Pathol* 1998, **186**(4):356-362.
56. Isola JJ, Kallioniemi OP, Chu LW, Fuqua SA, Hilsenbeck SG, Osborne CK, Waldman FM: **Genetic aberrations detected by comparative genomic hybridization predict outcome in node-negative breast cancer.** *Am J Pathol* 1995, **147**(4):905-911.
57. Janssen EA, Baak JP, Guervos MA, van Diest PJ, Jiwa M, Hermsen MA: **In lymph node-negative invasive breast carcinomas, specific chromosomal aberrations are strongly associated with high mitotic activity and predict outcome more accurately than grade, tumour diameter, and oestrogen receptor.** *J Pathol* 2003, **201**(4):555-561.
58. Climent J, Martinez-Climent JA, Blesa D, Garcia-Barchino MJ, Saez R, Sanchez-Izquierdo D, Azagra P, Lluch A, Garcia-Conde J: **Genomic loss of 18p predicts an adverse clinical outcome in patients with high-risk breast cancer.** *Clin Cancer Res* 2002, **8**(12):3863-3869.

59. Hislop RG, Pratt N, Stocks SC, Steel CM, Sales M, Goudie D, Robertson A, Thompson AM: **Karyotypic aberrations of chromosomes 16 and 17 are related to survival in patients with breast cancer.** *Br J Surg* 2002, **89**(12):1581-1586.
60. Jones C, Ford E, Gillett C, Ryder K, Merrett S, Reis-Filho JS, Fulford LG, Hanby A, Lakhani SR: **Molecular cytogenetic identification of subgroups of grade III invasive ductal breast carcinomas with different clinical outcomes.** *Clin Cancer Res* 2004, **10**(18 Pt 1):5988-5997.
61. Rennstam K, Ahlstedt-Soini M, Baldetorp B, Bendahl PO, Borg A, Karhu R, Tanner M, Tirkkonen M, Isola J: **Patterns of chromosomal imbalances defines subgroups of breast cancer with distinct clinical features and prognosis. A study of 305 tumors by comparative genomic hybridization.** *Cancer Res* 2003, **63**(24):8861-8868.
62. Seute A, Sinn HP, Schlenk RF, Emig R, Wallwiener D, Grischke EM, Hohaus S, Dohner H, Haas R, Bentz M: **Clinical relevance of genomic aberrations in homogeneously treated high-risk stage II/III breast cancer patients.** *Int J Cancer* 2001, **93**(1):80-84.
63. Zudaire I, Otero MD, Caballero C, Valenti C, Martinez-Penuela JM, Isola J, Calasanz MJ: **Genomic imbalances detected by comparative genomic hybridization are prognostic markers in invasive ductal breast carcinomas.** *Histopathology* 2002, **40**(6):547-555.
64. Bergamaschi A, Kim YH, Wang P, Sorlie T, Hernandez-Boussard T, Lonning PE, Tibshirani R, Borresen-Dale AL, Pollack JR: **Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer.** *Genes Chromosomes Cancer* 2006, **45**(11):1033-1040.
65. Fridlyand J, Snijders AM, Ylstra B, Li H, Olshen A, Segraves R, Dairkee S, Tokuyasu T, Ljung BM, Jain AN *et al*: **Breast tumor copy number aberration phenotypes and genomic instability.** *BMC Cancer* 2006, **6**:96.
66. Han W, Han MR, Kang JJ, Bae JY, Lee JH, Bae YJ, Lee JE, Shin HJ, Hwang KT, Hwang SE *et al*: **Genomic alterations identified by array comparative genomic hybridization as prognostic markers in tamoxifen-treated estrogen receptor-positive breast cancer.** *BMC Cancer* 2006, **6**:92.
67. Hicks J, Krasnitz A, Lakshmi B, Navin NE, Riggs M, Leibu E, Esposito D, Alexander J, Troge J, Grubor V *et al*: **Novel patterns of genome rearrangement and their association with survival in breast cancer.** *Genome Res* 2006, **16**(12):1465-1479.
68. Hwang KT, Han W, Cho J, Lee JW, Ko E, Kim EK, Jung SY, Jeong EM, Bae JY, Kang JJ *et al*: **Genomic copy number alterations as predictive markers of systemic recurrence in breast cancer.** *Int J Cancer* 2008, **123**(8):1807-1815.
69. Yoshikawa R, Yanagi H, Shen CS, Fujiwara Y, Noda M, Yagyu T, Gega M, Oshima T, Yamamura T, Okamura H *et al*: **ECA39 is a novel distant metastasis-related biomarker in colorectal cancer.** *World J Gastroenterol* 2006, **12**(36):5884-5889.
70. Toby GG, Gherraby W, Coleman TR, Golemis EA: **A novel RING finger protein, human enhancer of invasion 10, alters mitotic progression through regulation of cyclin B levels.** *Mol Cell Biol* 2003, **23**(6):2109-2122.
71. Andrews P, He YJ, Xiong Y: **Cytoplasmic localized ubiquitin ligase cullin 7 binds to p53 and promotes cell growth by antagonizing p53 function.** *Oncogene* 2006, **25**(33):4534-4548.
72. Leone G, Sears R, Huang E, Rempel R, Nuckolls F, Park CH, Giangrande P, Wu L, Saavedra HI, Field SJ *et al*: **Myc requires distinct E2F activities to induce S phase and apoptosis.** *Mol Cell* 2001, **8**(1):105-113.
73. He P, Varticovski L, Bowman ED, Fukuoka J, Welsh JA, Miura K, Jen J, Gabrielson E, Brambilla E, Travis WD *et al*: **Identification of carboxypeptidase E and gamma-glutamyl hydrolase as biomarkers for pulmonary neuroendocrine tumors by cDNA microarray.** *Hum Pathol* 2004, **35**(10):1196-1209.
74. Frank SR, Adelstein MR, Hansen SH: **GIT2 represses Crk- and Rac1-regulated cell spreading and Cdc42-mediated focal adhesion turnover.** *Embo J* 2006, **25**(9):1848-1859.
75. Bandaru V, Sunkara S, Wallace SS, Bond JP: **A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to Escherichia coli endonuclease VIII.** *DNA Repair (Amst)* 2002, **1**(7):517-529.
76. Shinmura K, Tao H, Goto M, Igarashi H, Taniguchi T, Maekawa M, Takezaki T, Sugimura H: **Inactivating mutations of the human base excision repair gene NEIL1 in gastric cancer.** *Carcinogenesis* 2004, **25**(12):2311-2317.

REFERENCES

77. Ma Y, Cui W, Yang J, Qu J, Di C, Amin HM, Lai R, Ritz J, Krause DS, Chai L: **SALL4, a novel oncogene, is constitutively expressed in human acute myeloid leukemia (AML) and induces AML in transgenic mice.** *Blood* 2006, **108**(8):2726-2735.
78. ten Berge RL, Meijer CJ, Dukers DF, Kummer JA, Bladergroen BA, Vos W, Hack CE, Ossenkoppele GJ, Oudejans JJ: **Expression levels of apoptosis-related proteins predict clinical outcome in anaplastic large cell lymphoma.** *Blood* 2002, **99**(12):4540-4546.
79. van Houdt IS, Oudejans JJ, van den Eertwegh AJ, Baars A, Vos W, Bladergroen BA, Rimoldi D, Muris JJ, Hooijberg E, Gundy CM *et al*: **Expression of the apoptosis inhibitor protease inhibitor 9 predicts clinical outcome in vaccinated patients with stage III and IV melanoma.** *Clin Cancer Res* 2005, **11**(17):6400-6407.
80. Muller-Pillasch F, Wallrapp C, Lacher U, Friess H, Buchler M, Adler G, Gress TM: **Identification of a new tumour-associated antigen TM4SF5 and its expression in human cancer.** *Gene* 1998, **208**(1):25-30.
81. Bueno-de-Mesquita JM, Linn SC, Keijzer R, Wesseling J, Nuyten DS, van Krimpen C, Meijers C, de Graaf PW, Bos MM, Hart AA *et al*: **Validation of 70-gene prognosis signature in node-negative breast cancer.** *Breast Cancer Res Treat* 2008.
82. Buysse M, Loi S, van't Veer L, Viale G, Delorenzi M, Glas AM, d'Assignies MS, Bergh J, Lidereau R, Ellis P *et al*: **Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer.** *J Natl Cancer Inst* 2006, **98**(17):1183-1192.
83. Mook S, Schmidt MK, Viale G, Pruneri G, Eekhout I, Floore A, Glas AM, Bogaerts J, Cardoso F, Piccart-Gebhart MJ *et al*: **The 70-gene prognosis-signature predicts disease outcome in breast cancer patients with 1-3 positive lymph nodes in an independent validation study.** *Breast Cancer Res Treat* 2008.
84. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ *et al*: **A gene-expression signature as a predictor of survival in breast cancer.** *N Engl J Med* 2002, **347**(25):1999-2009.
85. Paik S, Tang G, Shak S, Kim C, Baker J, Kim W, Cronin M, Baehner FL, Watson D, Bryant J *et al*: **Gene expression and benefit of chemotherapy in women with node-negative, estrogen receptor-positive breast cancer.** *J Clin Oncol* 2006, **24**(23):3726-3734.
86. Boiko AD, Porteous S, Razorenova OV, Krivokrysenko VI, Williams BR, Gudkov AV: **A systematic search for downstream mediators of tumor suppressor function of p53 reveals a major role of BTG2 in suppression of Ras-induced transformation.** *Genes Dev* 2006, **20**(2):236-252.
87. Lim IK, Lee MS, Lee SH, Kim NK, Jou I, Seo JS, Park SC: **Differential expression of TIS21 and TIS1 genes in the various organs of Balb/c mice, thymic carcinoma tissues and human cancer cell lines.** *J Cancer Res Clin Oncol* 1995, **121**(5):279-284.
88. Elmore LW, Di X, Dumur C, Holt SE, Gewirtz DA: **Evasion of a single-step, chemotherapy-induced senescence in breast cancer cells: implications for treatment response.** *Clin Cancer Res* 2005, **11**(7):2637-2643.
89. Cortes U, Moyret-Lalle C, Falette N, Duriez C, Ghissassi FE, Barnas C, Morel AP, Hainaut P, Magaud JP, Puisieux A: **BTG gene expression in the p53-dependent and -independent cellular response to DNA damage.** *Mol Carcinog* 2000, **27**(2):57-64.
90. Guardavaccaro D, Corrente G, Covone F, Micheli L, D'Agnano I, Starace G, Caruso M, Tirone F: **Arrest of G(1)-S progression by the p53-inducible gene PC3 is Rb dependent and relies on the inhibition of cyclin D1 transcription.** *Mol Cell Biol* 2000, **20**(5):1797-1815.
91. Lim IK, Lee MS, Ryu MS, Park TJ, Fujiki H, Eguchi H, Paik WK: **Induction of growth inhibition of 293 cells by downregulation of the cyclin E and cyclin-dependent kinase 4 proteins due to overexpression of TIS21.** *Mol Carcinog* 1998, **23**(1):25-35.
92. Ryu MS, Lee MS, Hong JW, Hahn TR, Moon E, Lim IK: **TIS21/BTG2/PC3 is expressed through PKC-delta pathway and inhibits binding of cyclin B1-Cdc2 and its activity, independent of p53 expression.** *Exp Cell Res* 2004, **299**(1):159-170.
93. Hong JW, Ryu MS, Lim IK: **Phosphorylation of serine 147 of tis21/BTG2/pc3 by p-Erk1/2 induces Pin-1 binding in cytoplasm and cell death.** *J Biol Chem* 2005, **280**(22):21256-21263.
94. Ficazzola MA, Fraiman M, Gitlin J, Woo K, Melamed J, Rubin MA, Walden PD: **Antiproliferative B cell translocation gene 2 protein is down-regulated post-**

- transcriptionally as an early event in prostate carcinogenesis. *Carcinogenesis* 2001, **22**(8):1271-1279.
95. Kawakubo H, Carey JL, Brachtel E, Gupta V, Green JE, Walden PD, Maheswaran S: **Expression of the NF-kappaB-responsive gene BTG2 is aberrantly regulated in breast cancer.** *Oncogene* 2004, **23**(50):8310-8319.
96. Calzolari F, Appolloni I, Tutucci E, Caviglia S, Terrile M, Corte G, Malatesta P: **Tumor progression and oncogene addiction in a PDGF-B-induced model of gliomagenesis.** *Neoplasia* 2008, **10**(12):1373-1382, following 1382.
97. Kawakubo H, Brachtel E, Hayashida T, Yeo G, Kish J, Muzikansky A, Walden PD, Maheswaran S: **Loss of B-cell translocation gene-2 in estrogen receptor-positive breast carcinoma is associated with tumor grade and overexpression of cyclin d1 protein.** *Cancer Res* 2006, **66**(14):7075-7082.