Profiling Human Breast Tumor Biopsies: Gene Expression Changes associated with ERBB2 Status and Prognosis, and possible Implications for Molecular Breast Cancer Classification in the Clinic

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Patrick Urban
aus Basel (CH)

Basel, 2009

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. U. Aebi (Fakultätsverantwortlicher)
Prof. U. Eppenberger (Dissertationsleiter)
Prof. N Hynes (Korreferentin)

Basel, den 24. April 2007

Prof. H.-P. Hauri (Dekan)

Table of Contents

Introduction

Breast Cancer

Breast cancer is the most common malignancy in women in Western countries. The American Cancer Society estimates that 212,930 new cases of breast cancer were diagnosed and 40,870 died of breast cancer in the United States in 2005 [1]. With around 4,000 cases every year the incidence in Switzerland is comparable to the USA. In most Western countries, the incidence of breast cancer has increased steadily over the past few decades, but breast cancer mortality appears to be declining. This suggests a benefit from early detection (e.g. screening) and more effective treatment [2,3].

The etiology of the vast majority of breast cancer cases is unknown. However, numerous risk factors for the disease have been established such as female gender, age, family history of breast cancer, early menarche, late menopause, older age at first childbirth, prolonged hormone replacement therapy, previous exposure to therapeutic chest wall irradiation, benign proliferative breast disease, and genetic mutations such as the BRCA1 and BRCA2 genes. However, except for female gender and increasing patient age, these risk factors are associated with only a minority of breast cancers.

Breast cancer is commonly treated by various combinations of surgery, radiation therapy, chemotherapy, and hormone therapy. In order to select the appropriate treatment strategy accurate classification is required. Traditionally, breast cancer classification included histo-pathological and clinical parameters such as stage of the disease, age, menopausal status and grade of the primary tumor. In addition, a variety of genes involved in breast cancer biology were studied and proposed as biomarkers to improve the classification. However, only a few of them, such as hormone receptors and ERBB2 status are used today to classify breast cancer patients and make treatment decisions in clinical routine [4,5]. More recently, gene expression-based approaches suggested that they could be superior to classical classification systems [6-10]. In this context, microarrays have emerged as one of the key technology allowing to analyze ten thousands of different transcripts simultaneously and permit together with various bioinformatics techniques complex relationships in the data to be explored [11-19]. Furthermore, the assessment of phosphorylation sites and, thus, the activation status of receptors or other key proteins relevant to cancer may add important information [20,21].

Since breast cancer is a very heterogeneous disease, both, on the biological and clinical level, but variation in transcriptional programs accounts for much of this diversity, gene expression-based classification has the potential to provide an individualized output for each patient in terms of prognosis and therapy prediction. The former can influence the aggressiveness of treatment, or in the case of excellent prognosis, even help to completely spare patients from unnecessary adjuvant therapy. In contrast, predictive markers might enable to tailor the treatment strategy such as the putative targets or indicate resistances to specific types of treatment. Thus, personalized treatment based on individual molecular finger prints can potentially enhance the treatment efficacy and decrease the risk of side effects.

Microarrays and Breast Cancer

In the late nineties the first microarray experiments in breast cancer were published using cell lines, and normal and tumor breast tissues (see also "Stanford Breast Cancer Microarray Project": http://genome-www.stanford.edu/_breast_cancer/). Using 60 different cell lines from various tissues Ross et al. [22] described a consistent relationship between gene expression pattern and the tissue of origin. This allowed the authors to identify characteristic gene expression pattern for individual tissues but also to recognize outliers whose previous classification appeared incorrect. For example, some of the breast cancer cell lines clustered together with caner cell lines from the central nervous system, kidney or melanoma. Moreover, specific features of the gene expression patterns appeared to be related to physiological properties of the cell lines, such as their doubling time in culture, drug metabolism or the interferon response. Comparison of gene expression patterns in the cell lines to those observed in normal breast tissue or in breast tumor specimens revealed features of the expression patterns in the tumors that had recognizable counterparts in specific cell lines, reflecting the tumor, stromal and inflammatory components of the tumor tissue.

Similarly, Perou et al. [23] used microarrays and clustering to identify patterns of gene expression in human mammary epithelial cells growing in culture and in primary human breast tumors. Clusters of co-expressed genes identified through manipulations of mammary epithelial cells *in vitro* also showed consistent patterns of variation in expression among the breast tumor samples. By using immunohistochemistry with antibodies against proteins encoded by a particular gene in a cluster, the identity of the cell type within the tumor specimen that contributed the observed gene expression pattern could be determined. Clusters of genes with coherent expression patterns in cultured cells and in the breast tumors samples could be related to specific features of biological variation among the samples. One such cluster was found to correlate with variation in cell proliferation rates, another with IFN response. Additionally, the group identified clusters of genes expressed by stromal cells and lymphocytes in the breast tumors. These reports supported the feasibility and usefulness of studying variation in gene expression patterns in human cancers as a means to dissect and classify breast cancer.

Subsequent work focused on human breast tumors. Perou et al. [24] characterized variation in gene expression patterns in a set of 65 surgical specimens of human breast tumors from 42 different individuals, using home-made cDNA microarrays. These patterns provided a distinctive molecular portrait of each tumor. Moreover, tumors that were sampled twice (before and after a 16-week course of doxorubicin chemotherapy) as well as tumors that were paired with a lymph node metastasis from the same patient revealed gene expression patterns that were more similar to each other than either was to any other sample. Sets of co-expressed genes were identified for which variation in messenger RNA levels could be related to specific features of physiological variation. The tumors could be classified into subtypes distinguished by pervasive differences in their gene expression patterns.

Clustering of tumors showed two main groups: ER-positives and ER-negative tumors. These could be further divided into several subgroups and were named according to the predominantly expressed genes defining the cluster. The ER-positive group was designated "luminal", and the ER-negative was subdivided into "normal-like", "ERBB2" and "basal" subtypes. For example, the "luminal/basal" designation originated from the observation that two distinct types of epithelial cells are found in the human mammary gland: basal (and/or myoepithelial) cells and luminal epithelial cells. These two cell types can be distinguished immunohistochemically using antibodies to keratin 5/6 ("basal") and keratins 8/18 ("luminal"). The gene expression cluster characteristic "basal" and "luminal" subtypes contained these genes. Moreover, many ER-related genes defined the "luminal" subtype. Accordingly, "ERBB2" subtype was characterized by the expression of ERBB2 and other genes present on the "ERBB2-amplicon". However, not all ERBB2-positive tumors grouped into the "ERBB2" subtype.

In a follow-up study, Sorlie et al. [6] used gene expression patterns to correlate breast tumor characteristics with clinical outcome. A total of 85 microarray experiments representing 78 cancers, three fibroadenomas, and four normal breast tissues were analyzed by hierarchical clustering. As reported previously, the cancers could be classified into a "basal", "ERBB2" and "normal-like" groups based on variations in gene expression. A novel finding was that the previously characterized ER-positive "luminal" subtype could be divided into at least two subgroups ("luminal A" and "luminal B"), each with a distinctive expression profile. These subtypes proved to be robust by clustering using two different gene sets: first, a set of 456 genes previously selected to reflect intrinsic properties of the tumors ("intrinsic" gene set) and, second, a gene set that highly correlated with patient outcome. Subsequent survival analyses showed significantly different outcomes for the various groups, including a poor prognosis for the "basal", "ERBB2" and "normal-like" subtypes, and a good prognosis for the "luminal" subtype. Interestingly, there was further a significant difference in survival when considering the two ER-positive subtypes ("luminal A and B"), the latter showing similar prognosis to the ER-negative subtypes.

This "Stanford Classification" ("intrinsic" gene set) was later refined using a total of 115 malignant breast tumors and 534 "intrinsic" genes where the genes used for the classification were selected based on their similar expression levels between pairs of consecutive samples taken from the same tumor separated by 15 weeks of neoadjuvant treatment [25]. Cluster analyses of two published, independent data sets representing different patient cohorts from different laboratories, uncovered some of the same breast cancer subtypes, and were also associated with significant differences in clinical outcome. In the same study, the authors included further a group of tumors from BRCA1 carriers and found that this genotype predisposes to the "basal" tumor subtype. These results supported the idea that the microarray breast tumor subtypes represent biologically distinct disease entities.

Meanwhile, numerous microarray-based experiments were published investigating several aspects of breast cancer including gene expression changes associated with hereditary breast cancer (e.g. BRCA 1/2 status) [26], histological type [27,28], initiating oncogenic event [29], breast cancer progression [30], clinical status (ER or lymph node) [31,32], micro-anatomical location of the tumor cell within the tumor [33], metastasis and recurrence [6-8,10,34,35], treatment [36-39], primary tumor and metastasis [40], metastasis to specific organ sites [41],tumor grade [42], the impact of DNA amplification on gene expression changes [43], wound response [44,45],or ERBB2 overexpression [46,47]. However, despite the successful correlation of gene expression profiles with clinical and tumor biological features, subsequent biological interpretation, reproducibility or comparison often turned out to be non-trivial, sometimes complicating the translation into the clinical setting [48-50]. At least with respect to gene-expression-based outcome predictors consistency started to emerge [51].

An example of an microarray experiment is shown in Figure 1.

Microarrays and Bioinformatics

Microarrays are providing unexpected quantities of genome-wide data on gene-expression profiles. For example, an experiment with 30 samples in involving 10.000 genes and 10 different experimental conditions will produce $3x10^6$ pieces of primary information. Cross comparison of sample images can multiply this total by many times. Many computation tools are available to analyze the date, but the methods that are used can have profound influence on the interpretation of the results. The understanding of these computational methods and analyzing techniques is therefore required for optimal experimental design and meaningful data analysis [14,16,17,49,52]. These include methods and tools of mining and warehousing of biodata, image processing and data analysis software.

The simplest way to identify genes of potential interest is to search for those that are consistently either up or down regulated. However, identifying pattern of gene expression and grouping genes into classes might provide much greater insight into their biological function and relevance. To achieve this, there exists a large group of statistical methods. Supervised methods, for example, use already existing biological or clinical information to guide the clustering algorithm. In addition, before starting with the actual data analysis, the first step in every experiment starts with the careful selection of the array probes, usually from databases, and design of the experiment. Once the microarray slide is printed, hybridized, scanned and the image processed, normalization of the relative fluorescence intensities of the two channels is done. Normalization adjusts for differences in labeling and detection efficiencies for the fluorescent labels and for differences in the quantity of initial RNA from the samples examined in the assay. Again there are various methods to achieve this goal.
It has become increasingly clear, that there are many "good" approaches, and the application of various techniques will allow different aspects of the data to be explored. Cluster analysis, for example, does not give absolute answers. Instead, these are data-mining techniques that allow relationships in the data to be analyzed. Among the most promising and exciting applications, but also challenging, are those that classify human disease states such as cancer using patterns of gene expression signatures.

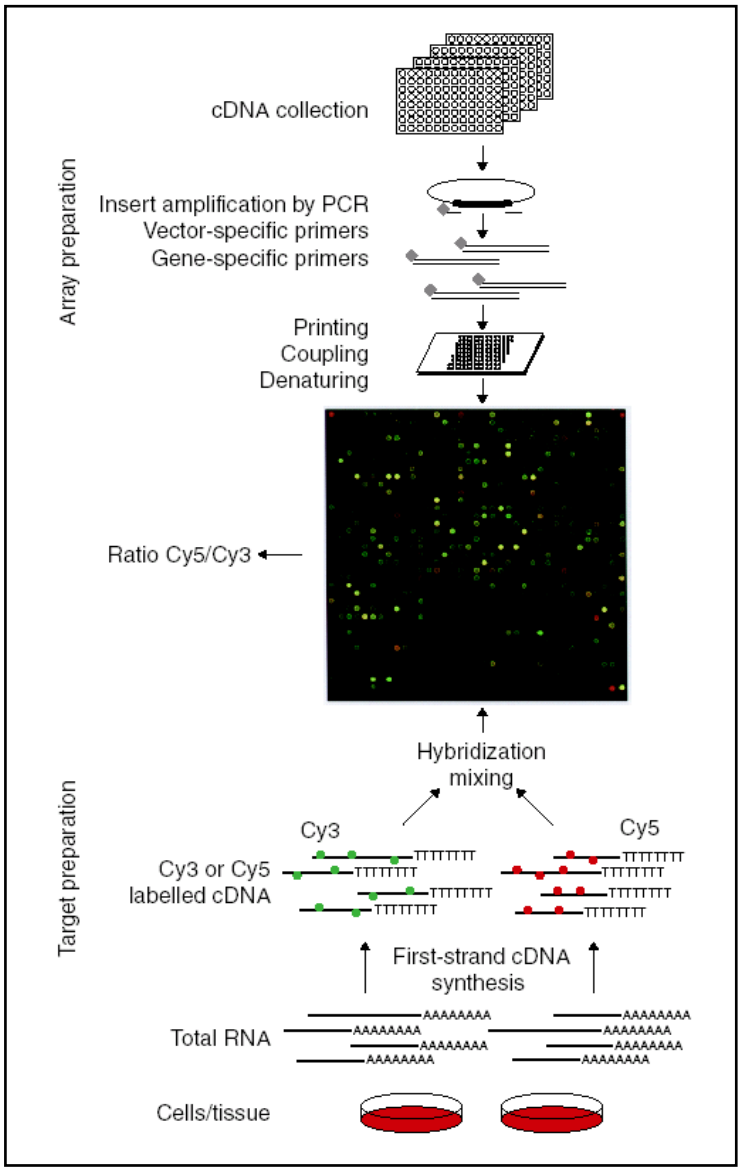


Figure 1. Schematic overview of probe array and target preparation for spotted cDNA microarrays [11]. The analysis of gene expression by microarrays includes several steps: 1) Array construction. cDNA (partial or complete) or chemically synthesized oligonucleotide sequences are spotted onto a slide. 2) Probe Preparation. RNA from cells or tissue is extracted, converted to cDNA and labeled. The use of different fluorescent dyes allows mRNAs from two different samples (usually probe and control) to be labeled in different colors. 3) Hybridization of probe to array. The two samples are mixed and will hybridize to complementary sequences through competitive binding on the array. Unbound material is washed away before scanning. 4) Scanning and detection. The hybridized array is scanned with the use of a confocal laser scanner that can detect each of the two fluorescent dyes. 5) Normalization and data analysis. The images produced during scanning for each dye are aligned by specialized software. The spot intensities are adjusted and then overlapped. Intensity, number of spots and background are determined and quantified. Controls, such as externally added sequences, reporter genes, or total fluorescence for each sample help to correct for differences in labeling and detections efficiency of the two fluorescent tags (normalization). Commonly, transcripts levels for each single gene are measured using the ratio of the two sample signals.

ERBB Receptors

The physiological and pathological roles of the ERBB family of receptors and ERBB2 in particular, with respect to normal and cancerous tissues as well as its implication for treatment strategies has been reviewed in [53-62] (Figure 2).
The ERBB family of growth factor receptors consists of 4 members: EGFR (ERBB1), ERBB2 (HER2/Neu), ERBB3 and ERBB4. Exposure of cells to ERBB receptor-specific ligands results in receptor homo-dimerization and/or hetero-dimerization, kinase activation, and auto- and cross-phosphorylation of cytoplasmic tyrosine residues. Various adaptor molecules bind to the phosphorylated receptors, mainly via SH2 (Src homology 2) and PTB (phosphotyrosine binding) domains, resulting in signal transduction initiation that ultimately regulates gene transcription [54,55].
Activated ERBB receptors stimulate many intracellular signaling pathways and, despite extensive overlap in the molecules that are recruited to the different active receptors, different ERBB family members preferentially modulate certain signaling pathways, owing to the ability of individual ERBBs to bind specific effector proteins. Two of the main pathways activated by the receptors are the mitogen-activated protein kinase (MAPK) and the phosphatidylinositol 3-kinase (PI3K)–AKT pathway [55,62]. Other important ERBB signaling pathways are the signal transducer and activator of transcription proteins (STAT's), often via EGFR activation, the SRC tyrosine kinase, protein kinase C (PKC), and mammalian target of rapamycin (mTOR), which is activated downstream of PI3K/AKT and other growth regulators.
It appears that the relative expression level of each ERBB family member, as well as ligand specificity, determines the nature of the dimerizations, and hence the repertoire of adaptors which bind to the activated receptors. This in turn determines the specificity and strength of downstream signaling. While ERBB2 has no known ligands, it is the preferred hetero-dimerization partner of other family members. Thus, ERBB2 overexpression is believed to enhance signaling from these receptors in response to binding of their specific ligands, but also independently through homo-dimerization and autophosphorylation.
Recently, Jones et al. [63] used protein microarrays comprising virtually every SH2 and PTB domain encoded in the human genome to measure the equilibrium dissociation constant of each domain for 61 peptides representing physiological sites of tyrosine phosphorylation on the four ERBB receptors. By slicing through the network at different affinity thresholds, the authors found surprising differences between the receptors. Most notably, EGFR and ERBB2 became markedly more promiscuous as when their concentration was raised, whereas ERBB3 did not. This, was proposed the authors, might contribute to the high oncogenic potential of EGFR and ERBB2 which are frequently overexpressed in many human cancers, including breast.

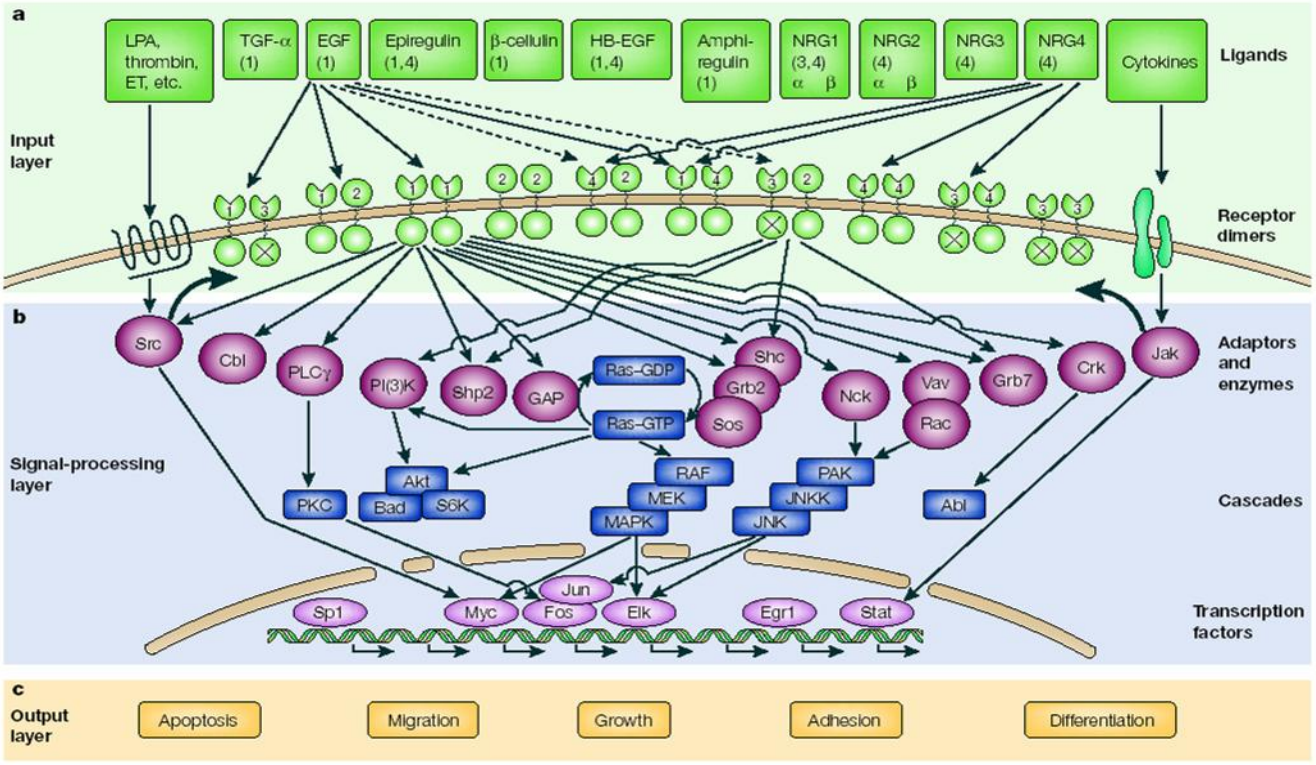


Figure 2. Overview of the ERBB2 signaling network (receptors, ligands and downstream pathways) [55].

ERBB2 and Breast Cancer

In human breast cancer ERBB2 (HER2/Neu) is overexpressed in 15 to 30% of cases, and is associated with poor prognosis and an increased likelihood of metastasis [20,64]. Similarly, also ERBB2 phosphorylation has been linked to more aggressive phenotype [21]. However, not all patients with ERBB2-positive tumors develop metastasis and despite intensive research efforts, the biological mechanisms underlying the oncogenicity of ERBB2 are still not fully understood.

Studies established that regulated ERBB2 signaling is required for normal development and homeostasis of the mammary gland. The four ERBB receptors and their multiple polypeptide ligands are differentially expressed during development of the mouse mammary gland [65]. Profiles suggest that EGFR and ERBB2/Neu are required during ductal morphogenesis, whereas ERBB3 and ERBB4 are preferentially expressed through alveolar morphogenesis and lactation. Moreover, gene knockouts established that EGFR and its ligand, amphiregulin (AR), are essential for ductal morphogenesis in the adolescent mouse and likely provide the required epithelial-stromal signal. In contrast, the phenotypes of transgenic mice expressing dominant negative ERBB2 and ERBB4 proteins suggest that these receptors act in different ways to promote or maintain alveolar differentiation. However, mouse models have also confirmed that overexpression or mutation of ERBB-network components, including ERBB2/Neu, EGFR and ERBB ligands contribute significantly to the initiation of human and rodent breast cancer.

Evidence for a causal link between ERBB2 expression in particular, and tumor progression has come from several experimental studies which showed that cells transfected with ERBB2 are more invasive *in vitro*, and metastatic *in vivo*, and ERBB2 expressing tumors arising in transgenic mice are also capable of metastasis. In addition, inhibition of the oncogene function is capable of reversing the malignant phenotype [53].

Cancer cell invasion and metastasis is a complex process with many steps involved, and ERBB2 was shown to interfere with many of these processes. On of the early changes is loss of normal tissue architecture and migratory capability. These phenomena are associated with alterations in the cell-cell and cell-ECM (extracellular matrix) interactions. There is compelling evidence that indeed alterations in this system can contribute to metastasis, and that in breast cancer, activation of ERBB2 is playing an important role in cell adhesion and signaling. For example, ERBB2 was found to inhibit transcription of the E-cadherin gene which is one of the major cell-cell adhesion molecules and is often reduced in breast cancer [66]. But also interactions between ERBB2 and b-catenin, CD44, ICAMs, integrins, and several other molecules involved in cell adhesion and motility has been described [53,67].

Another important process in cancer progression is proteolytic activity. Indeed, ERBB2 was shown to interact with expression of several proteases such as MMPs and uPA, their receptors or inhibitors [53,68-70]. Downstream mediators linking ERBB2 with increased proteolytic activity include the ETS family of transcription factors [71]. But proteolytic activity is not enough; tumor cell migration is a further important determinant of metastatic capacity. Overexpression of ERBB2 in breast carcinoma cell lines as well as HRG (heregulin) stimulation of non-invasive MCF-7 cell lines induces a migratory phenotype [72,73]. Spencer et al. [74] utilized carcinoma cells depleted of ERBB2, but not other ERBB receptor members, to specifically examine the role of ERBB2 in carcinoma cell migration and invasion. Cells stimulated with EGF-related peptides show increased invasion of the extracellular matrix, whereas cells devoid of functional ERBB2 receptors did not. Further, overexpression of ERBB2 in cells devoid of other ERBB receptor members was further sufficient to promote ERK activation and CAS/Crk coupling, leading to cell migration. Thus, ERBB2 appears to be a critical component in the migration/invasion machinery of carcinoma cells. Recently, a novel molecule, MEMO (mediator of ErbB2-driven cell motility) was described and seems to control cell migration by relaying extracellular chemotactic signals to the microtubule cytoskeleton after ERBB2 activation [67].

Tumor angiogenesis is yet another process relevant to tumor progression. The major inducers of angiogenesis are the vascular endothelial growth factors (VEGFs), although many other factors are known to play a role in angiogenesis [75]. The ERBB2 signaling pathway has been shown to impact angiogenesis at several levels, including the release of angiogenic factors, the response of endothelial cells to them, and interactions between tumor and endothelial cells during extravasation. Upregulation of VEGF transcription and protein production in cells lines occurred via a PI3K/AKT and the MAP/ERK and the MAPK/p38 pathways, and could be inhibited by p38 antagonists as well as monoclonal antibodies directed against ERBB2 [53,76,77]. Moreover, reports have described that ERBB2 signaling increased HIF1a protein synthesis in a rapamycin/mTOR-dependent manner providing a molecular basis for VEGF induction and tumor angiogenesis elicited by ERBB2 [78]. In a spontaneous metastasis model using human breast cancer cells lines stably transfected with constitutively active ERBB2 kinase, injected mice had increased metastasis incidence and tumor microvessel density [79]. Clinical validation in breast tumors that overexpress ERBB2 and which had higher VEGF expression, showed significantly higher p70S6K phosphorylation as well, and correlated with higher levels of AKT and mTOR phosphorylation. Additionally, patients with tumors having increased p70S6K phosphorylation showed a trend for worse disease-free survival and increased metastasis. More recently, also MMPs received extensive attention in this process, and which too, appear to be regulated – at least to some extend – via ERBB signaling [80].

Interestingly, Kao J et al. [81] addressed in a recent report the potential functional contribution of co-amplified genes since amplification rarely comprises only single genes. In breast cancer, the "ERBB2" amplification occurs within chromosome band 17q12, which contains the ERBB2 oncogene. Analysis of array-based comparative genomic hybridization and expression profiling data indicated that even the minimum region of recurrent amplification at 17q12 includes several other genes, including GRB7 and STARD3, which also exhibit elevated expression when amplified [46,47,82,83]. Western blot analysis confirmed overexpression of each in SKBR3 and BT474 cell lines which harbor the amplification. In these cell lines (but not in control MCF7 breast cancer cells lacking 17q12 amplification), targeted knockdown of ERBB2 expression using RNA interference (RNAi) resulted in decreased cell proliferation, decreased cell-cycle progression, and increased apoptosis. Notably, targeted knockdown of either GRB7 or STARD3 also lead to decreased cell proliferation and cell-cycle progression, albeit to a lesser extent compared with ERBB2 knockdown. Therefore, amplification and resultant overexpression of co-amplified genes could contribute to some of the characteristics observed in ERBB2-positive breast cancers.
Taken together, compelling evidence suggests that signaling via ERBB2 and other ERBB receptors can regulate many key processes in breast cancer, including proliferation, angiogenesis, invasion and metastasis.

ERBB Receptors and Therapy

The ERBB family has become a promising field for targeted therapy. Several antibodies directed against the extracellular domain of ERBB receptors and tyrosinekinase inhibitors (TKIs) are at advanced clinical development stages, or are already successfully used in the clinic [57,58,84-87]. The treatment of tumor cells with these agents affects and inhibits many of the intracellular pathways that are essential for cancer development and progression. The mechanisms of action in cancer patients, however, are often less clear. The same applies for putative mechanisms of resistances or selection of appropriate pharmacodynamic markers.
Trastuzumab (Herceptin ®, Genentech), for example, is a recombinant humanized monoclonal antibody directed against the extracellular domain of ERBB2, and is approved for therapy of ERBB2 overexpressing breast cancer. Various mechanisms have been observed or proposed *in vivo* and *in vitro* by which trastuzumab induces regression of ERBB2-overexpressing tumors [60]. Trastuzumab binding to ERBB2 was shown to lead to internalization and degradation of the ERBB2 receptor protein, and trastuzumab induced p27 (KIP1) levels and interaction with CDK2, thereby decreasing CDK2 activity. Trastuzumab was further shown to reduce signaling of the PI3K and MAPK pathways, promoting cell cycle arrest and apoptosis. Trastuzumab treatment could also reduce angiogenesis via decreased levels of VEGF and TSP1. Other data indicate that efficacy of trastuzumab could be related to its induction of immune response, e.g. via stimulation of natural killer cells and activation of anti-body dependent cellular cytotoxicity (ADCC). Besides, trastuzumab was shown to inhibit repair of DNA after treatment with cytotoxic chemotherapy.
The objective response rate of trastuzumab in metastatic breast cancer patient in phase II trials was rather low, ranging from 12 to 34% [85,88]. Thus, many tumors demonstrate primary *de novo* or intrinsic resistance to the drug. Combination with taxanes could increase response rates and survival [60]. However, the majority of patients who achieve an initial response to trastuzumab-based regimens develop resistance within 1 year. Proposed mechanisms of resistance to trastuzumab include the MUC4 protein that could block the inhibitory actions of the drug via direct binding with ERBB2, preventing interaction with its molecular target. Decreased interaction or gain-of-function could theoretically also result from mutations in the ERBB2 gene, and has been demonstrated for EGFR in lung cancer [89]. Also the existence of compensatory pathways and signaling aberrations downstream of ERBB2 have been discussed [57]. Furthermore, aberrant activation of other tyrosinekinases such as IGF1R or FGFR members occurs in various types of cancer. Loss of PTEN is yet another mechanism that has been associated with trastuzumab resistance [90].
To the situation, potential novel agents or combinations with trastuzumab could prove beneficial such as pan- or dual-specific ERBB tyrosinekinase inhibitors, pertuzumab (an antibody that in contrast to trastuzumab disrupts receptor heterodimerization), IFG1R inhibitors, mTOR inhibitors, inhibitors of angiogenesis and MMPs, but also optimized combinations with anti-hormonal (e.g. aromatase inhibitors) and classical chemotherapy regimen.
In this context, it is also interesting to note that toxicities reported with ERBB targeted therapies correlate with the normal, physiological functions of the respective targeted receptor. For example, ERBB2 plays an important role in the development of the heart [91]. Embryos that lack the receptor die due to improper formation of the ventricular trabeculea in the myocardium which is responsible for maintaining blood flow. Moreover, conditional ablation of ERBB2 in postnatal cardiac-muscle cell lineages revealed that ERBB2 is also essential in the adult heart for the prevention of cardiomyopathy; cardiac myocyte-targeted HER2 gene knockout in adult mice lead to impaired cardiac remodeling in response to stress [92]. Indeed, some breast cancer patients show cardiotoxicity when given trastuzumab, including cardiomyopathy, congestive heart failure and decreased left ventricular ejection fraction. This might even worsen if trastuzumab is combined with other cardio-toxic agents such as anthracyclines.

Aims of the Study

The main goals of the project were outlined as follows:

- To investigate the differences between ERBB2-positive and ERBB2-negative breast tumor samples on the gene expression level and characterize the molecular phenotype associated with ERBB2 status
- To investigate possible downstream effects associated with ERBB2 signaling
- To identify relevant subgroups or genes associated with outcome in ERBB2+ and ERBB2- tumors and explore putative clinical implications towards a molecular classification of breast cancer

For this purpose, collaboration was established between the University of Basel, Stiftung Tumorbank Basel (STB), OncoScore AG in Riehen, the Swiss Institute for Experimental Cancer Research (ISREC) and National Center of Competence in Research (NCCR) Molecular Oncology in Lausanne, and the Swiss Bioinformatics Institute in Lausanne. Furthermore, the Bioinformatics Core Facility (BCF) and the DNA Array Facility (DAFL) of the Center for Integrative Genomics (CIG) in Lausanne were involved.

Work started with cell lines and a set of human breast cancer biopsies to set up and optimize protocols, evaluate technology and perform feasibility study for using gene expression microarray for profiling human tumor biopsies. Different RNA extraction, amplification, labeling, hybridization and washing methods were evaluated. Various hardware and software was tested to optimize scanning, spot detection and normalization procedures. In parallel, quantitative real-time PCR (qrt-PCR) assays for a panel of over 60 cancer-related genes as well as quantitative immunoassays (ELISA and CLISA) for p-ERBB2 and p-AKT were developed. First test-cDNA microarrays were printed with 100 genes, and followed later by the first generation of 10K human arrays with over 10.000 genes.
For the main project a specific subset of primary breast cancer biopsies was selected for which banked fresh frozen tissue, clinical follow-up and histopathological data, prospective measurements of ER, PgR, uPA/PAI-1 and ERBB2 protein levels were available (Stiftung Tumorbank Basel). Cryo-sections were preformed for all samples including H&E slides, RNA extracted and amplified. Samples were analyzed by qrt-PCR and microarrays. Additionally, p-ERBB2 and p-AKT were measured by ELISA/CLISA. For the microarray analysis, a careful and extended pilot experiment was performed to guide the design. Subsequent analysis comprised several statistical and exploratory approaches, starting from normalization, cluster analysis and supervised methods to build classifiers and identify differentially expressed genes, and pathway mining. Validation was performed using qrt-PCR data, and was extended to bigger, independent sample collections including publicly available microarrays data sets.
Besides, comparison of a subset of breast tumor biopsies with a commercial microarray platform (Affymetrix U95 GeneChip; collaboration with Novartis), identification of differentially expressed genes in cells that were stimulated by Amiphiregulin or EGF, identification of differences between malignant and non-malignant breast tissues, and the assessment of paired core-biopsy versus surgical biopsy to evaluate possible gene alterations introduced by the sampling procedure were studied. Additionally, the roles of p-AKT and p-ERBB2 were also explored. Finally, clinical applications and feasibility assessment with respect to the use of molecular classification for routine diagnostics were evaluated.

Materials and Methods

Detailed methods are outlined in the attached publications section; additional protocols can be found in the Appendix. A brief summary of methods and materials is given below.

Microarrays

Spotted cDNA arrays were produced at the microarray core facility at ISREC (GeneMachine OmniGrid 3000, CA). An initial human test chip contained approximately 100 features (human cDNA and arabidobsis control genes) printed in triplicate for testing purposes. Later a human 10K chip with over 10.000 spots was produced containing the cDNA library from Incyte Genomics plus a selection of some 500 cancer related genes from a custom wish-list compiled at the institute from various research teams. Detailed specifications about chip design and content can be found on the DNA array facility Lausanne webpage: http://www.unil.ch/dafl/ page5509.html.
During the development and testing period several improvements were achieved due to the following changes: Exchange of the glass support from TeleChem (TeleChem International, Sunnyvale, CA) to Quantifoil (Quantifoil Micro Tools, Germany) leading to significantly lower background levels. The design of the chip was re-arranged and spike genes (*arabidobsis thaliana*) distributed all over the slide, particularly in every sub-array. *Arabidobsis* RNA was subsequently used during hybridizations and mixed in predefined concentrations and known ratios into the sample preparation allowing better control of the hybridizations, and make statements about the quality of the slides. Additionally, the spotting solution was changed reducing unspecific signals and the protocols for hybridizations and washings were improved (detailed protocols in appendix). All sequences on the chip were re-blasted and new, improved annotation files were generated in an automated fashion every few weeks. In addition, quality measures with respect to the spotted sequence were added, and routine quality controls analysis after each hybridization, scanning and normalization were defined and integrated into an automated slide processing web-tool which became part of each experiment. Some examples are shown in Figures 3 and 4.

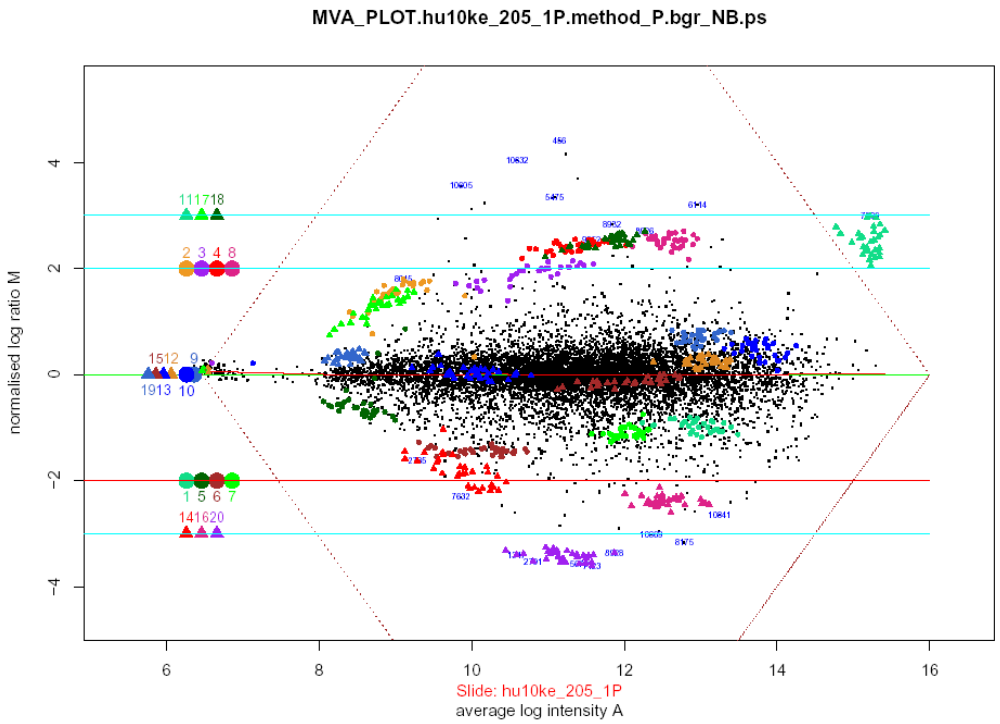


Figure 3. *"MVA plot" after normalization (print-tip specific LOWESS). Average log intensity versus normalized log ratio is plotted for a tumor sample hybridized against a reference (pool of all samples). Dotted lines indicate the detection limits of the scanner/experiment (saturation and background). Horizontal lines indicate 2 and 3-fold over or under expressed genes. Color-coded circles and triangles represent arabidopbsis spike genes. These genes were spiked into each experiment in known ratios and concentrations, and were used to control the labeling, hybridization, scanning and normalization procedures.*

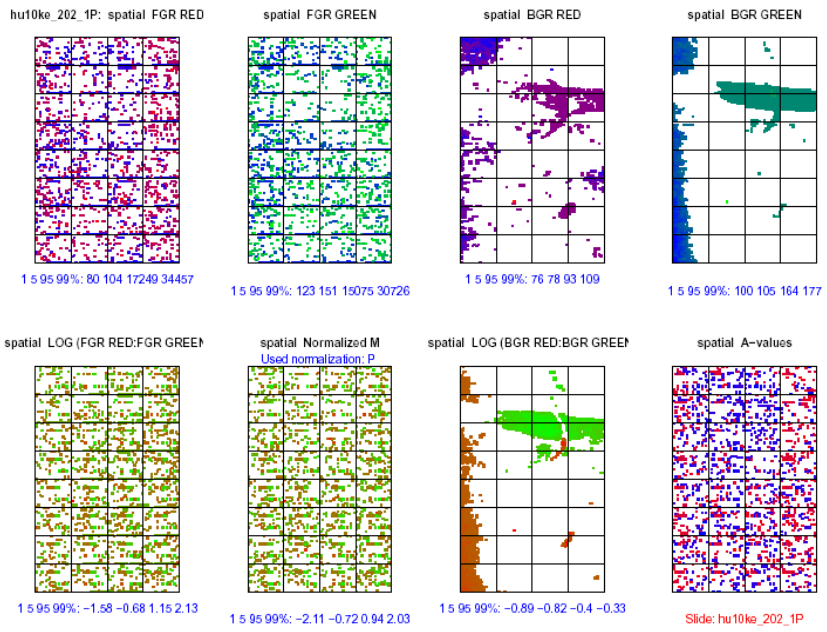


Figure 4. ***"****Spatial plots" of a microarray slide after normalization. Back- and foreground intensities are plotted for each channel (Cy3/target and Cy5/reference) separately and represented according to its special localization on the array. These were used together with several statistical criteria to assess the quality of slides/hybridization and identify potential issues.*

Cell Lines and Tissue Biopsies

Six different breast cancer cell lines (MCF-7, T-47D, ZR-75-1, BT-474, MDA-MB-231, SKBR-3) were cultured and grown in media, harvested, and pelleted. Aliquots were frozen and stored at -70°C. RNA was exacted and subsequently used for development of methods, validation of protocols, and to produce reference materials (e.g. RNA, protein extracts) for various experiments [93-95].

A first set of 39 frozen primary breast cancer fresh frozen biopsies was obtained from the STB (Stiftung Tumorbank Basel) for feasibility assessment of using archived tumor material for the purpose of the main project. RNA quantity, quality and histology (cryo-sections) were evaluated and first hybridizations performed. The feasibility assessment further included a comparison of two microarray platforms (data not presented).

For the main project, 100 primary breast cancer biopsies were selected from STB according to prospectively assessed ERBB2 protein expression levels by ELISA. Roughly, 50 % of samples over-expressed ERBB2 (>260ng/mg), and 50% had "intermediate" protein levels of ERBB2 (not over-expressed, 100-260ng/mg). For all patients complete clinical follow-up was available, including patient demographics, histo-pathological characteristics, treatment and survival data. Excluded were patients who received previous neoadjuvant treatment and patients with history of another cancer.

Cryo-sections were preformed for all biopsies using O.C.T (Tissue Tek, Electron Microscopy Sciences), stained with H&E and reviewed. Percentage of tumors cells, stromal component, inflammatory cells, fat and necrotic tissue were quantified. Biopsies with less than 50% tumor infiltration or less than 20% tumor cell content, or biopsies containing moderate to extensive necrosis were to be excluded or replaced for future experiments. Additional slides were prepared and stored for future laser-capture micro-dissection (LCM). Afterwards, biopsies were pulverized; part of the powder was used to extract RNA for microarray and qrt-PCR analysis, other parts powder were used to extract proteins/prepare cytosols and membrane fractions for ERBB2, p-ERBB2 and p-AKT ELISA/CLISA analysis.

Additional experiments showed that 2-3ug RNA per mm^3 tissue could be extracted from cryo-sections. Furthermore, the RNA quality from cryo-sections was generally good with no apparent degradation as well as tissue embedding by O.C.T. did not alter RNA extraction, amplification nor quality. Subsequently, qrt-PCR experiments were performed with RNA from cryo-sections after thorough pathological review.

RNA Extraction and Quality Assessment

Different RNA extraction methods were tested: (1) Trizol+RNeasy (RNAeasy Mini Kit, Qiagen; for details see appendix), (2) RNeasy alone, and (3) an automated extraction apparatus (ABI 6100 Nucleic Acid PrepStation, Applied Biosystems). Quality and quantity of the extracted RNA and the robustness of the method was evaluated by Bioanalyzer 2100 (RNA 6000 Nano LabChip-Kit; Agilent Technologies) and by O.D (18S:28S rRNA ratio, gel pattern/fragments).
Prior to extraction, biopsy material was pulverized [20]. There were no significant difference between the methods, however, the automated extraction device did not perform well when using fat containing tissues (e.g. breast cancer biopsies) - the wells and capillary system of the machine clogged.
The lower limit for extraction was 10.000 cells. Typical yields obtained for extractions from 50.000 cells were 30-50ng/ul and 180-250ng/ul for 250.000 cells, respectively. Measurements of the same samples by O.D. generally indicated higher yields by a factor of 1.5-2 as compared to the Bioanalyzer.
Subsequently, RNA extractions for all tumor biopsies in the main project were done with the RNeasy kit. No or insufficient RNA could be extracted in 7 out of 100 biopsies; 81 biopsies had RNA concentrations $\geq$ 0.15ug/ul ($\geq$ 7.5ug total RNA, elution in 50ul). The extracted RNA demonstrated mostly good to very good RNA quality (18S:28S rRNA ratio >1.5 or >2.0 respectively, no signs of degradation). Examples of 12 RNA profiles from breast tumor biopsies are given in Figure 5.

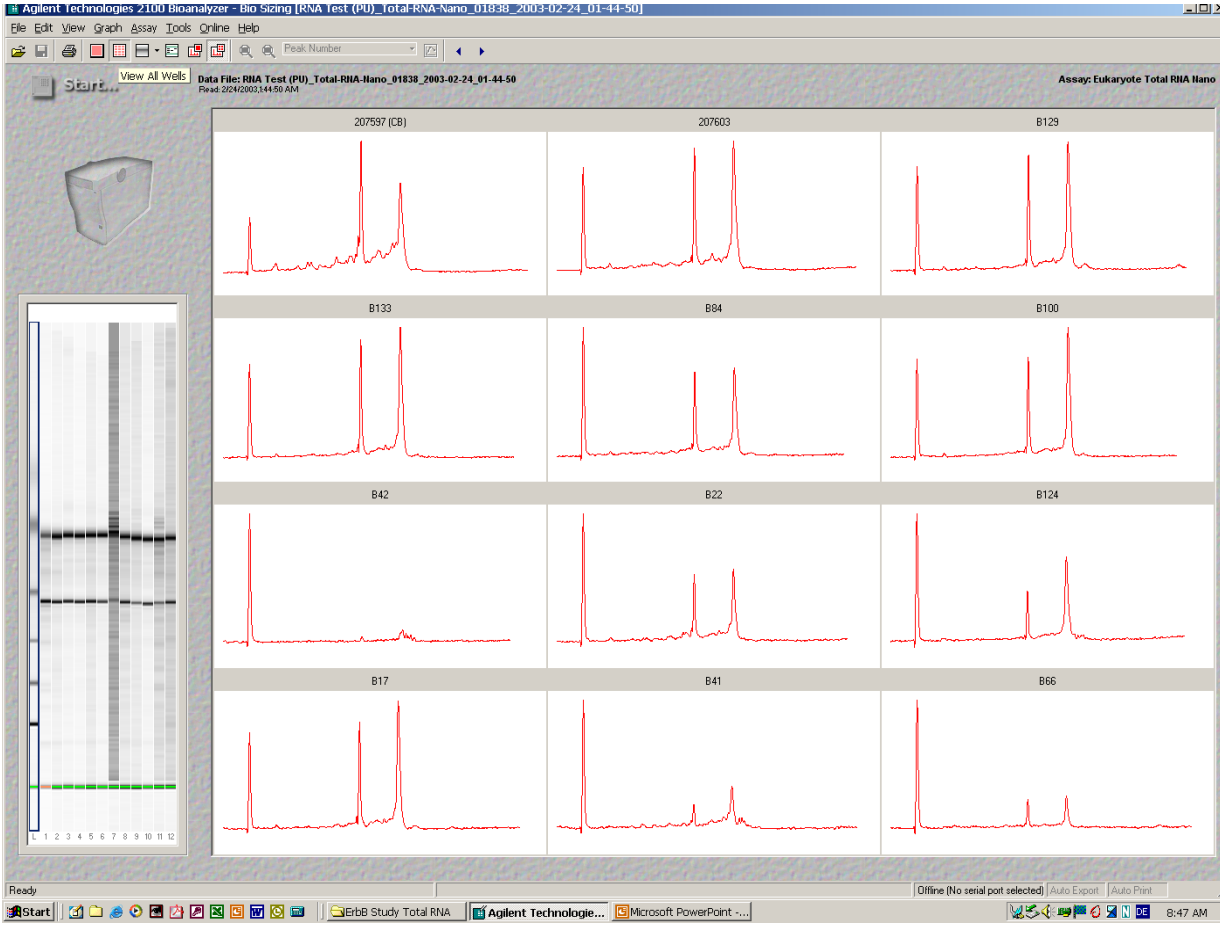


Figure 5. *Example of total RNA from 12 breast cancer samples analyzed after extraction on a Bioanalyzer. The left panel shows a "gel-like" picture; the right panel illustrates the abundance (y-axis) for various RNA sizes (x-axis) for each sample. The first peak corresponds to an internal marker, the subsequent two peaks correspond to 18S and 28S rRNAs. The first sample (top left) shows signs of degradation (decreased 28S peak and 28S:18S ratio, degradation fragments present).*

RNA Amplification and Labeling

RNA amplification was done with the Ambion amplification kit (MessageAmp II aRNA Amplification Kit, Ambion) based on the Eberwine method [96] (Figure 6). In brief, RNA is reverse transcribed using an oligo-dT primer with a T7 polymerase promoter followed by second strand synthesis. Subsequently, the double stranded cDNA is transcribed *in vitro* adding T7 polymerase. Applying this method, typically 2000-3000 fold amplification was achieved starting from 1 to 5ug total RNA. The mean size of the aRNA after amplification was around 1500 base pairs (bp's). Robustness of the procedure was assessed in repeated, independent amplifications of the same RNA demonstrating very high reproducibility.
Amplified RNA was labeled using the fluorescent dyes Cy3 and Cy5 (Cy3/5-dCTP) and 5ug aRNA in a reverse transcription step resulting in labeled cDNA. Out of the initial 100 biopsies, 92 were successfully amplified RNA (aRNA).

Detailed protocols for probe preparation can be found in the appendix.

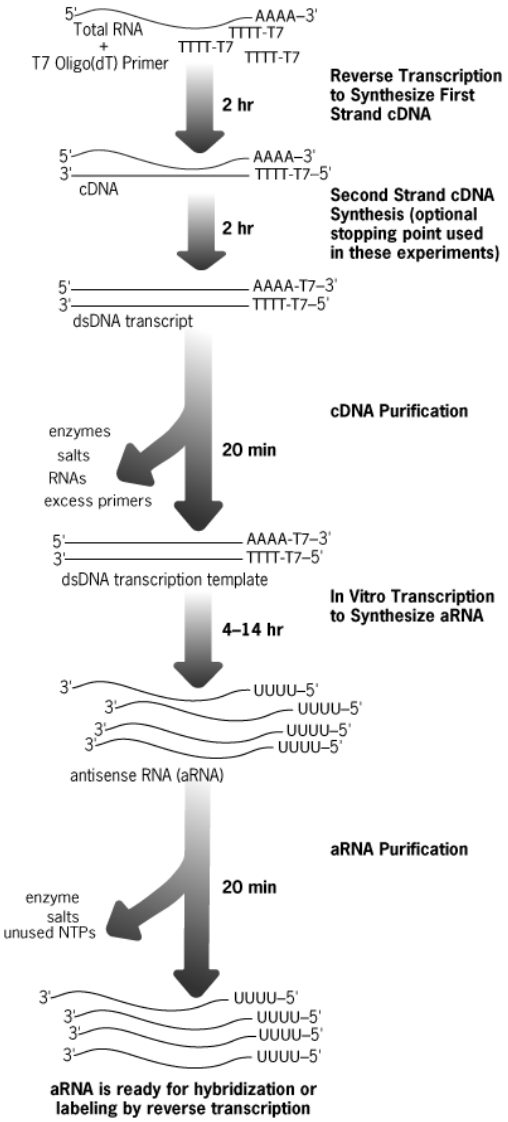


Figure 6. *Schematic overview of the RNA amplification procedure (MessageAmp II aRNA Amplification Kit, Ambion) based on the Eberwine method [96].*

Reference RNA

Pooled of RNA from all tumor biopsy samples was prepared and used as reference RNA since the goal of the project was mainly to measure differences between the samples rather than differences between different tissues or cell lines, and since microarrays measure small differences more accurately (limited dynamic range). Typically, 5ug of aRNA was required for one hybridization (5ug of pool (Cy3) and 5ug of target (Cy5)), and it was assumed that replicate hybridization would be performed (see "pilot experiment"). Thus, biopsies which had more than 21ug (equivalent to 4 hybridizations) of aRNA were used for the pool. If a biopsy had more than 21ug of aRNA, the excess RNA was used for the pool but not more than 30ug meaning that no biopsy contributed more than 30ug to the pool. The total amount of pooled aRNA generated was 1832ug (equivalent to 366 hybridizations).

Scanning and Feature Extraction

Two scanners were evaluated (Agilent, Agilent Technologies and ScanArray4000, Packard Bioscience). After extended testing, Agilent was selected for future experiments due to its ability to scan both channels (Cy3/Cy5) simultaneously leading to a better overlay of the two images and higher reproducibility in repetitive scans of the same image.
Numerous scanner settings and their influence on the final results were evaluated (e.g. saturation, photo-bleaching, detection of low expressed genes). In this context, repeated scans with different settings were applied in the main experiment (e.g. 50% and 100% laser power/PMT gain).
Two feature extraction software were compared: ScanAlyze (developed at the Eisen's Lab in Stanford, see webpage at: http://rana.lbl.gov/EisenSoftware.htm) and GenePix (Molecular Devices Corporation). In light of the "high throughput" aspect of the main project (more than 300 arrays to be scanned) and irregularities of the spot size/arrangement during array printing, GenePix was the preferred software owing to its automatic grid finding capability and adjustment of spot size (see example of a microarray image in the appendix).

Normalization and Clustering

Normalization is a procedure used to "standardize"/account for differences arising from labeling (incorporation efficiency), hybridization, scanning, the amount of starting material, etc., to allow for comparisons between experiments. Several methods can be applied. Together with the bioinformatics core facility some standard procedures and methods were established after extensive testing. Most of the work was automated by submitting the output files from the scanner directly to a web-based application, which would return the processed data including several quality indicators.
Normalization worked best with the print-tip specific LOWESS (locally linear fitting) correction. Both, Spot and the normalization procedure have packages that can be loaded into R statistical software (spot and sma package, www.r-project.org).
Cluster analysis was done using "Cluster" and "Treeview" software (available at http://rana.lbl.gov/EisenSoftware.htm) or directly in R statistical software. Generally, genes were filter according to variation, mean centered and normalized, and hierachically clustered (genes and arrays) using Spearman correlation as similarity metric and single linkage as clustering method.

qrt-PCR, p-ERBB2 and p-AKT

qrt-PCR assays were set up in collaboration with OncoScore AG and STB in Basel. Protein assays (sandwich immunoassays), if not commercially available, were set up in collaboration with Molekulare Tumorbiologie, University of Basel and STB. Details can be found in [93-95]. Eighty-nine of the 92 breast cancer biopsies for which RNA was amplified and used on microarray were successfully assessed by qrt-PCR using a panel of 60 cancer related genes; a list with all genes can be found in the appendix. Another 70 of these biopsies were analyzed for p-ERBB2 (ELISA), and 65 samples for p-AKT (CLISA).
Comparison between the methods revealed that there was generally high correlation between mRNA expression levels measured by microarray and qrt-PCR for most of the genes (see results section). There was further high correlation between mRNA and protein expression levels of ESR1 and ERBB2 (r^2=0.74 and 0.71 respectively; data not shown).

Statistical Analysis

Statistical analysis was carried out using "R" statistical software and bioconductor packages (www.r-project.org). Development of a classification algorithm included bootstrapping to select genes in a training set and rank them according to their association with a desired outcome variable based on t-statistics for categorical or Cox coefficients in case of survival data. Performance was evaluated using a gene score and ROC (receiver operator characteristics) by calculating the area under the curve (AUC) starting for the top ranked gene, and optionally by successively adding variables according to their rank to calculate new scores and AUC's. The number of variables (e.g. genes) in the final model was selected according to the performance of the AUC curve. The optimal cutoff for the gene scores as well as other variables was determined using CART (classification and regression trees) to obtain the best separation of the classes [97]. Finally, the same genes, score and cutoff were used to assess the performance on the left out samples (test set). Univariate and multivariate Cox proportional hazard model was used to assess association with survival. Kaplan-Meier method was used to plot survival curves and the logrank test to compared outcomes.

Results

Several experiments and studies were performed prior to starting the main experiment to develop and validate methods, and assess the feasibility to use human breast cancer biopsies in a high-through put gene expression profiling. Moreover, an extensive pilot experiment was performed in close collaboration with the Bioinformatics Core Facility (BCF) to assess relative importance of different sources of variability because the whole experiment/hybridization would take several weeks and array printing was done in batches of 100 slides. For more information on some of the results and conclusions from these studies please consult the appendix.
Most of the data from this work has been published; detailed data is presented and discussed there [93,94,98-100]. However, a summary of the key findings as well as supplementary data not previously published will be presented on the subsequent pages. Work started with exploratory and correlative analysis using microarray gene expression data from 89 primary breast cancer biopsies and recapitulated many of the previously established breast cancer microarray findings reported in the literature. Subsequent analysis focused on the characterization of the ERBB2 status on a molecular level and the identification of prognostic markers in various subsets of breast cancer patients, particularly in patient subgroups stratified by ERBB2 and ER status. In this context, also the value of "activated" ERBB2 as measured by phosphorylation of the receptor, and downstream signaling pathways (e.g. phosphorylated AKT) were explored. In addition, different methods to build a classifier to predict ERBB2 status, survival and other clinical parameters were assessed. Finally, the findings were validated in a larger collective of breast cancer patients assessed by quantitative real-time PCR, and extended and compared to independent breast cancer microarray data sets for which gene expression and clinical data were publicly accessible.

Evaluation of gene expression changes associated with ERBB2 status, survival, and a classifier for ERBB2 status

Gene expression levels detected by microarray correlated strongly with qrt-PCR results for several known prognostic as well as predictive breast cancer genes (see Figure 7). Similar results were observed for correlations with protein expression levels (ELISA) of ER and ERBB2, and to a much lesser extent also with uPA. Both, ER and ERBB2 expression levels showed bimodal distributions with approximately 50% of tumors being ERBB2+/- (selection bias of the sample set), and 75% and 25% being ER+/-, respectively (Figure 8).
Unsupervised clustering analysis of the 89 breast cancer biopsies showed that the dominant gene cluster separating the samples into the two main groups was driven by the "ER-signature" (ER, GATA3, LIV1, BCL2, GREB1, RERG, etc), and correlated well with the ER status of the tumor samples (Figure 9). The "ERBB2"-cluster constituted another important, although smaller cluster in terms of the number of genes, and was anti-correlated with ER (Figure 9). Many of the genes in the "ERBB2"-cluster mapped to chromosome 17q12-21 ("ERBB2 amplicon", ERBB2, GRB7, STARD3, PSMD3, etc; Figure 9). A third prominent cluster was mainly defined by genes belonging to the "basal-like" subgroup (KRT5, 14, 17, KIT, MMP7, etc; Figure 10). Other clusters were dominated by immune response and inflammation or stroma genes (Figure 11). Interestingly, these clusters showed also significant differences with respect to outcome (Figure 12).

Supervised analysis revealed that ERBB2 status was strongly associated with the underlying genetic changes, arising from the amplification of ERBB2 and it flanking genes. Indeed, these genes demonstrated also correlated expression pattern. However, not all genes from the "ERBB2-amplicon" showed simultaneous co-expression, suggesting that such variation might arise from different amplification patterns at DNA level, reflecting the changes observed at mRNA level (Figure 13).
Survival analysis underlined the important role of ERBB2 and the amplicon, respectively (Table 1). Among the top genes associated with disease-free survival (DFS) mapped several to chromosome 17q12-21 ("ERBB2-amplicon" region). However, survival analysis in subsets of patients with ERBB2+ and ERBB2- tumors revealed striking differences with respect to the genes associated with patient outcome (Table 2, see also Figures 19 and 20). The gene which correlated best with metastasis-free survival in ERBB2 samples was uPA, and was putatively shown to be regulated by ERBB2 in the literature [53,68-70]. This finding was later followed-up and validated in a bigger set of breast cancer biopsies assessed by qrt-PCR, and in two independent microarray breast cancer data set, which were publicly accessible (Amsterdam [34] and Rotterdam [8] microarray data sets).

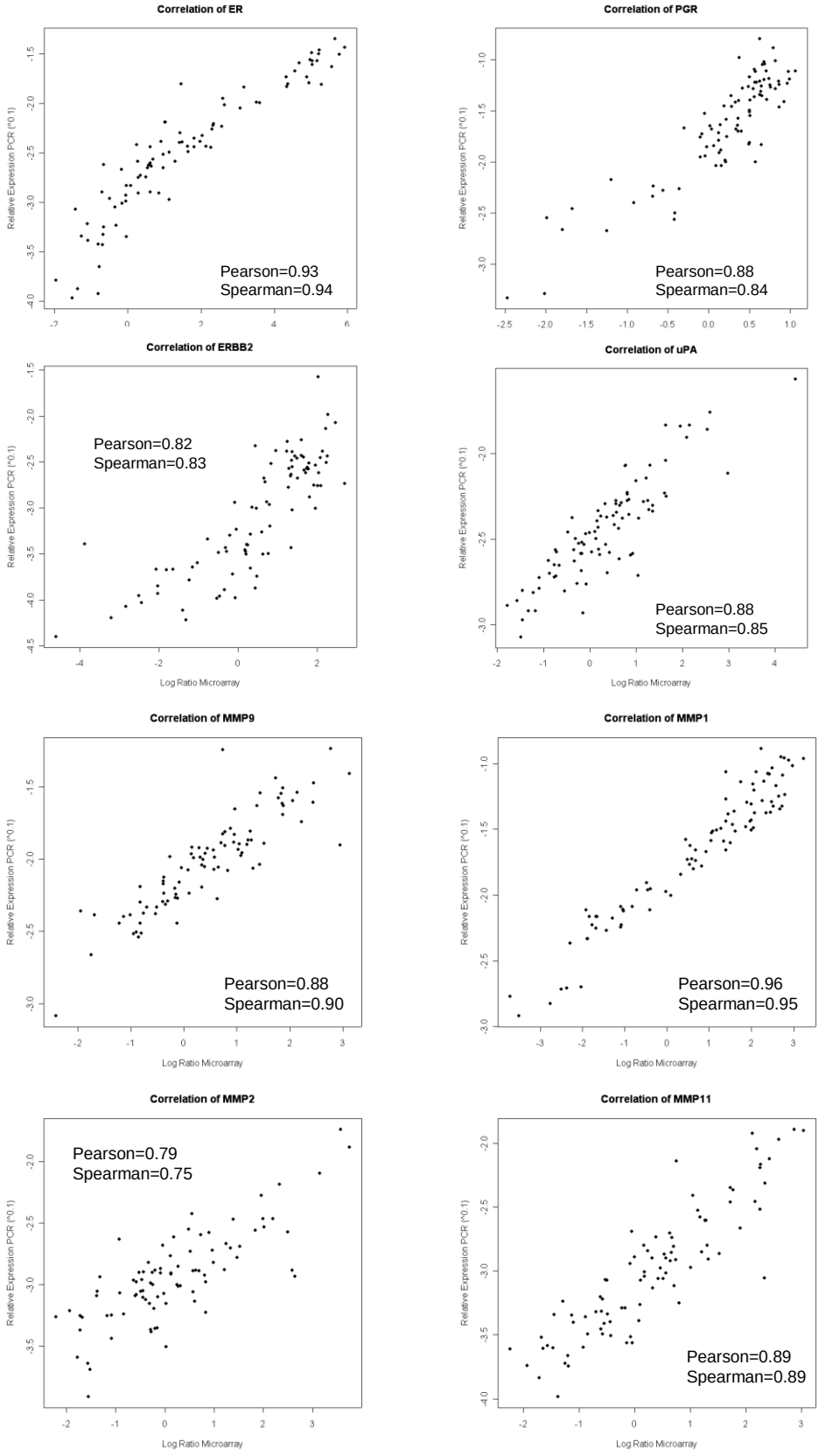
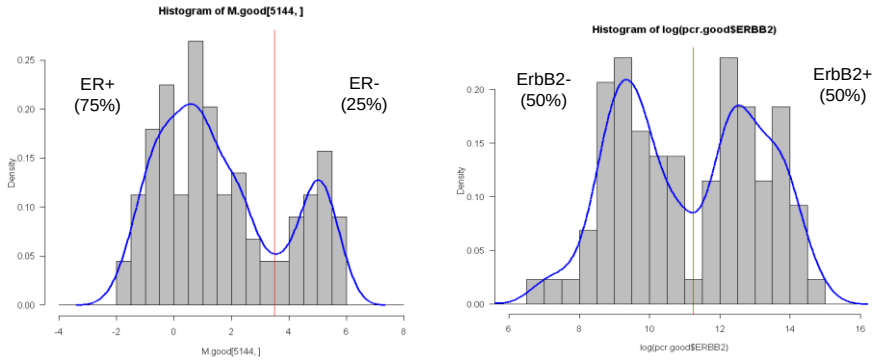


Figure 7. *Correlation between mRNA expression levels measured by microarray and qrt-PCR for known prognostic and predictive breast cancer genes (ER, PGR, ERBB2 and uPA in (top 4 panels) and MMPs (bottom 4 pannels)).*

Distribution of ER and ERBB2



Status	ER-	ER+	Total
ERBB2-	3	42	44
ERBB2+	19	25	45
Total	22	67	89

Figure 8. *Distribution of ER and ERBB2 (mRNA level) in the 89 breast cancer biopsies as measured by microarray.*

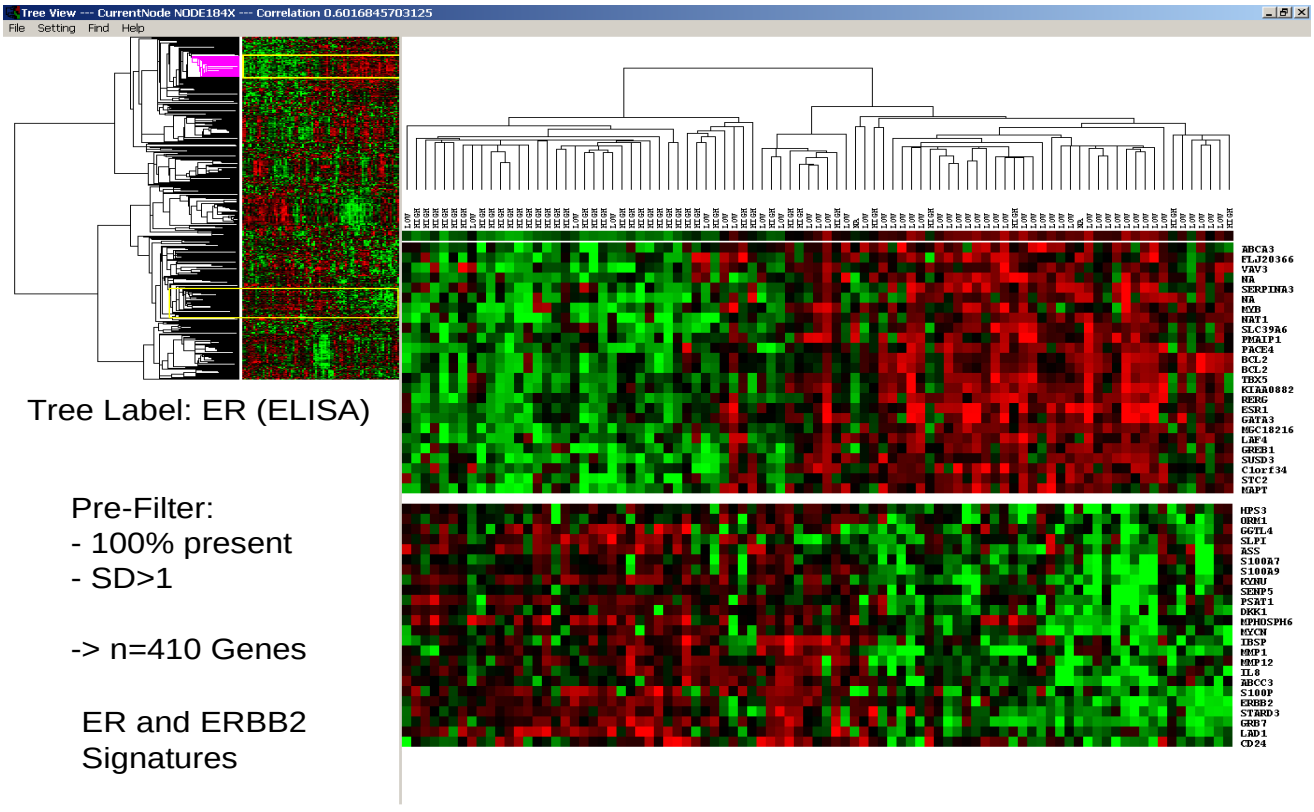


Figure 9. *Unsupervised clustering of 89 breast cancer biopsies (horizontal tree) based on 410 genes (vertical tree). Genes were filtered using standard deviation for the difference in genes expression among the samples. Clustering of samples is dominated by the "ER"-cluster (ER, GATA3, LIV1, BCL2, GREB1, RERG, etc) and correlates with the ER status based on ELISA measurement (horizontal tree labels). The "ERBB2"-cluster is anti-correlated with ER and contains several 17q12-21 genes ("ERBB2 amplicon", ERBB2, GRB7, STARD3, PSMD3, etc.).*

Figure 10. *“Basal-like” cluster with expression of basal keratins (KRT5, 14, 17), KIT and MMP7.*

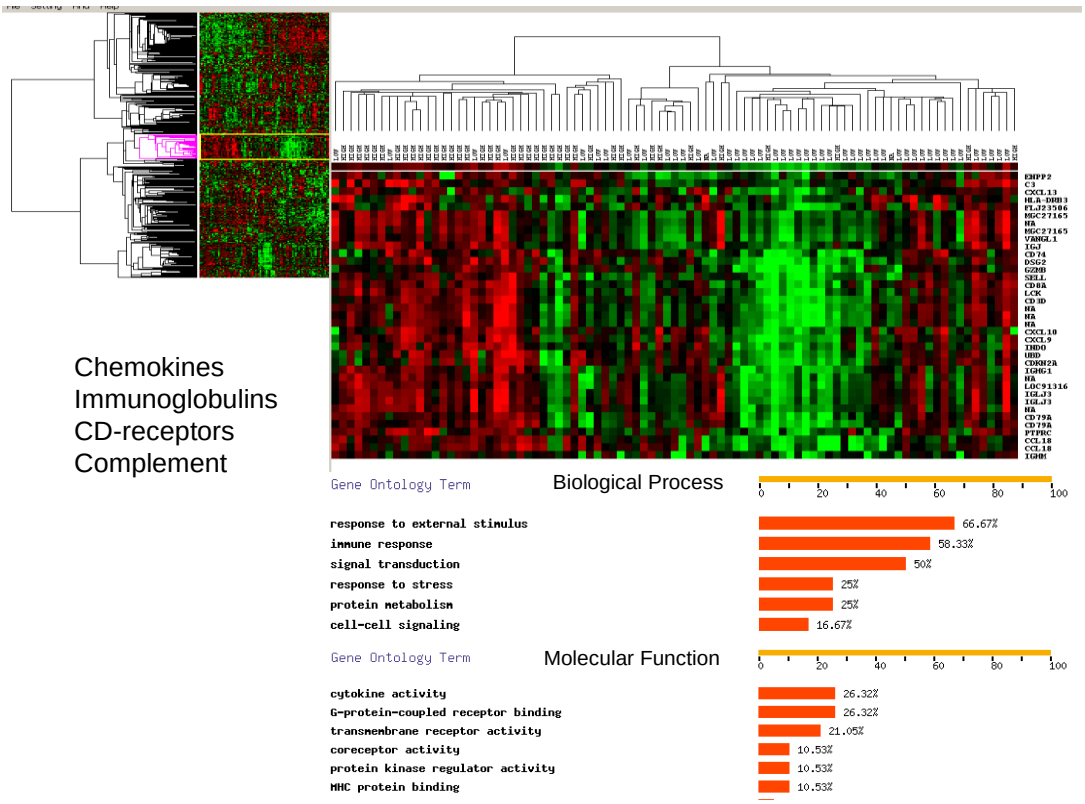


Figure 11. *Example of another prominent cluster associated with inflammatory processes and/or immune response.*

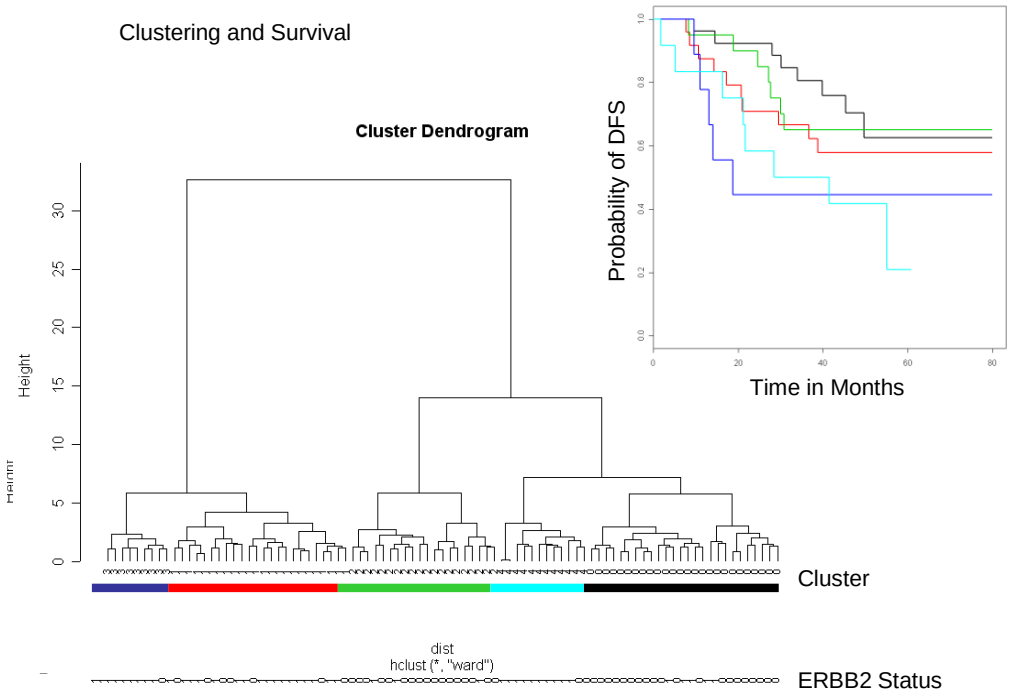


Figure 12. *Unsupervised clustering: Clusters show differences in survival (DFS) as well as in ERBB2 status.*

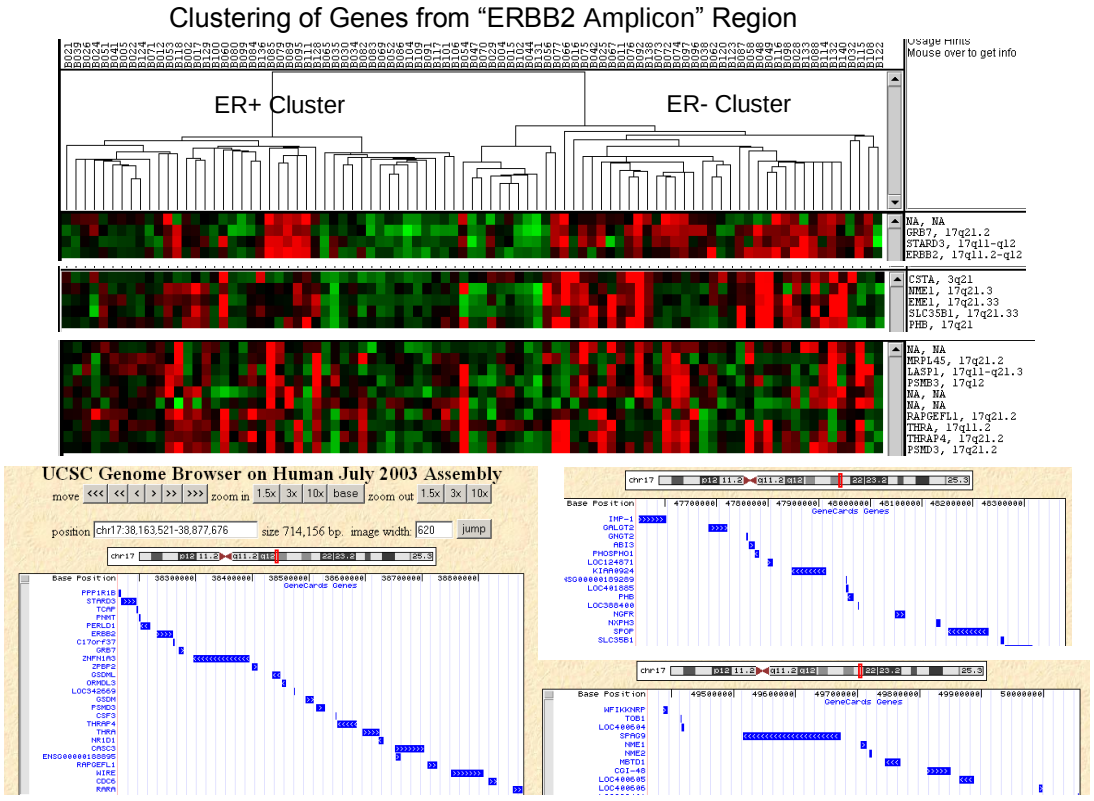


Figure 13. *Genes from the "ERBB2-amplicon"(17q12-21) show co-expression pattern and are typically overexpressed in ER- tumors. However, not all "ERBB2-amplicon" genes show simultaneous co-expression suggesting that different DNA amplification patterns might reflect the expression changes observed at RNA level.*

Clone	ChromLoc	Gene	p.Cox
IC1911640	17q21.2	MRPL45	0.00653
IC1818527	18q11.2	LAMA3	0.0089
IC1968268	17q21.2	PSMD3	0.00912
IC1737833	17q12	PSMB3	0.01027
IC2056158	10q24	PLAU	0.01115
IC1821420	17q11-q12	STARD3	0.01221
IC293495	NA	SERPINA3	0.01231
IC1701725	12q24.1	OAS1	0.01351
IC5047895	10q11.2	DKK1	0.01428
IC2799253	5q31	NME5	0.01456
IC4116386	6q25.1	ESR1	0.01555
IC1711151	14q21-q24	HIF1A	0.01699
IC1240890	6p21.3	HLA-DQA1	0.01735
IC2989812	1q21	ANXA9	0.01884
IC2288855	17q21.1	MAPT	0.02186
RG52741	16p13.3	ABCA3	0.02280
IC1505977	4q32-q33	GRIA2	0.02419
IC2059176	17q11-q21.3	LASP1	0.02944
IC863708	16q12.2	TNRC9	0.03106
IC630995	3q21-q25	AGTR1	0.03305
IC2537863	1p31.1	C1orf29	0.03399
RG502151	17q25	SLC16A3	0.03843
IC1858050	16q22.1	CDH1	0.04043
IC856900	1p34-p12	CYP4B1	0.04349
IC1646294	10p15	GATA3	0.04396
IC1637576	17q21.2	THRAP4	0.04466
IC2823476	5q35.2	STC2	0.04863
IC1480159	19q13.3	KLK6	0.05061
IC2622181	6q12	ME1	0.05370
IC1842009	7q31	SLC26A3	0.05410
RG2542529	17q11.2-q12	ERBB2	0.05891

Table 1. *List of top genes associated with survival in the overall population (n=500 top microarray genes were analyzed; filter based on standard deviation; Cox regression analysis, association with DFS)*

Clone	ChromLoc	Gene	p.Cox
IC2056158	10q24	PLAU	0.00144
IC2506867	4p14	UCHL1	0.00345
IC2537863	1p31.1	C1orf29	0.00469
IC4855492	6q21-q22	COL10A1	0.00509
IC2986240	6q21-q22	COL10A1	0.00617
IC1577614	5q11.2	FST	0.00658
IC1782172	8q22-q23	SDC2	0.00718
IC1701725	12q24.1	OAS1	0.00816
IC2636634	19p13.1	COMP	0.00840
IC1258790	13q33	ITGBL1	0.01051
IC2900277	14q32	IFI27	0.01255
IC4079783	11p15.4	ADM	0.01519
IC1421929	16q22.1	CDH3	0.02012
IC1813269	16q13-q22	CES1	0.02586
IC5047895	10q11.2	DKK1	0.02595
RG378461	4q21-q25	SPP1	0.02950
IC978433	8q22.3	CTHRC1	0.03072
IC1405940	3q23-q24	PLOD2	0.03108
IC1240890	6p21.3	HLA-DQA	0.03330
IC1643186	4q11-q13	PDGFRA	0.03339
IC1506256	9q34.1	CRAT	0.03883
IC1215596	10q25-q26	IFIT1	0.04265
IC2797546	11q22.3	MMP10	0.04274
IC980544	2p15-p13	SLC1A4	0.04281
IC2352645	4q13-q21	AREG	0.04970
IC3040858	1q21-q23	HIST2H2B	0.05176

Table 2. *Top genes associated with survival in the subgroup of ERBB2+ tumors only (n=500 top microarray genes; Cox analysis, DFS).*

Subsequently, classifiers were evaluated to classify samples according to a known clinical or biological parameter such as ERBB2 status. For this purpose, bootstrapping was used to select genes in a training set and rank them according to their association with a desired outcome variable based on t-statistics. The area under the curve (AUC) was calculated using a gene score and ROC (receiver operator characteristics) starting for the top ranked gene, and thereafter by successively adding genes according to their rank to calculate new scores and assess AUC. The number of genes in the final model was selected according to the performance of the AUC curve. Then, the optimal cutoff for the gene score with the selected number of genes was determined using CART (classification and regression trees) to obtain the best separation of the classes. Finally, the same genes, score and cutoff were used to assess the performance on the left out samples (test set, Figure 14).
For ERBB2 status, this classifier showed almost 100% accuracy in the training and around 95% in the test set when all genes from the microarray gene set were used. This was mainly due to the fact, that obviously ERBB2 and other 17q12-21 genes were preferably selected. As a consequence, if the number of genes in the classifier was increased the performance started to decrease since "noise" was added - genes that do not discriminate as accurately ERBB2 status as do genes from the ERBB2-amplicon. Once all genes from the amplicon were removed, the classifier reached its best performance with around 20-30 genes in the model. The performance in the training set revealed typically less than 10% misclassification, and around 10-15 % misclassification in the test set (Figure 15). The genes discriminating ERBB2+/- samples contained many known cancer-related genes (e.g. MYCN, S100P, MMP1, CEACAM6, etc) as well as ER-related genes (ESR1, RERG, BCL2, GREB1, ERBB4, etc; Figure 22). As expected, the two main groups showed significant difference with respect to survival (Figure 16). Finally, a three-fold cross-validation was implemented to better control for gene selection and performance assessment of the classifier (Figure 17). A "combined" ROC-curve might be used to evaluate or define specific cutoff values which will meet desired test characteristics (e.g. sensitivity and specificity of the test).

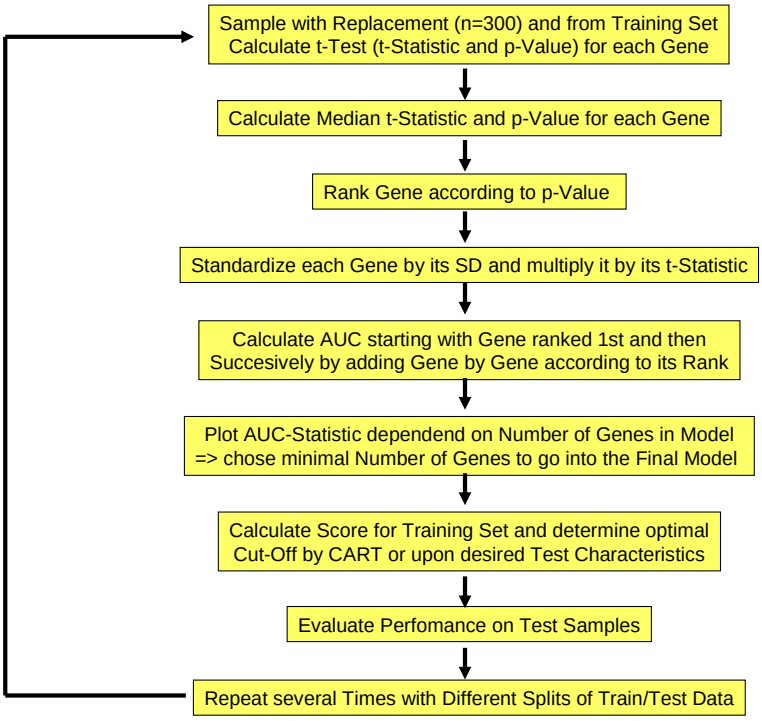


Figure 14. *Overview of classification algorithm. Bootstrapping was used to select genes in a training set and rank them according to their association with a desired outcome variable (class) based on t-statistics. The area under the curve (AUC) was calculated using a gene score and ROC (receiver operator characteristics) starting for the top ranked gene, and thereafter by successively adding genes according to their rank to calculate new scores and AUC's. The number of genes in the final model was selected according to the performance of the AUC curve. Then, the optimal cutoff for the gene score with the selected number of genes was determined using CART (classification and regression trees) to obtain the best separation of the classes. Finally, the same genes, score and cutoff were used to assess the performance on the left out samples (test set).*

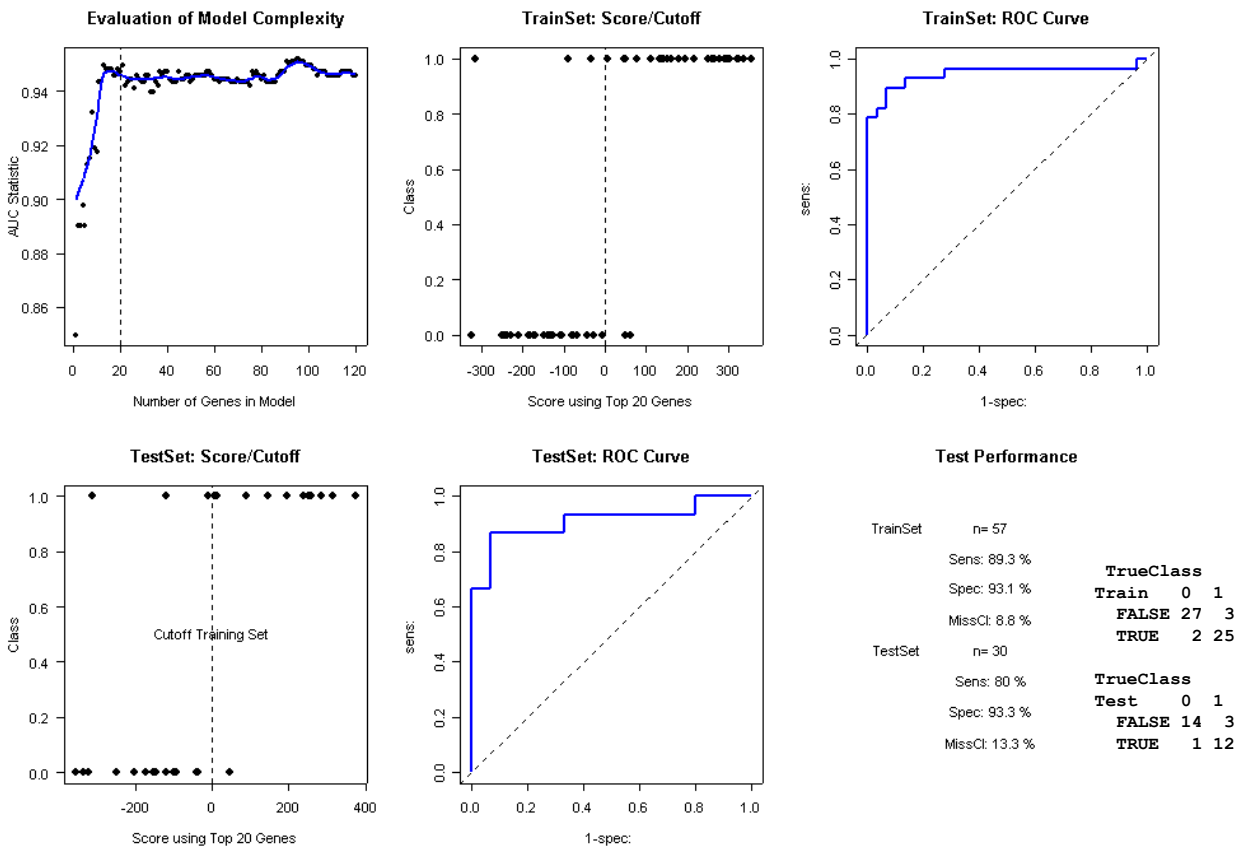


Figure 15. *Performance evaluation of a classifier for ERBB2 status; "ERBB2-amplicon" genes were removed. The performance of the classifier (AUC) demonstrated a plateau after including approximately 20 genes in the training set. Using an "optimal" cutoff for a 20-gene score by CART resulted in a misclassification rate of around 5 % in the training, and 13% in the test set.*

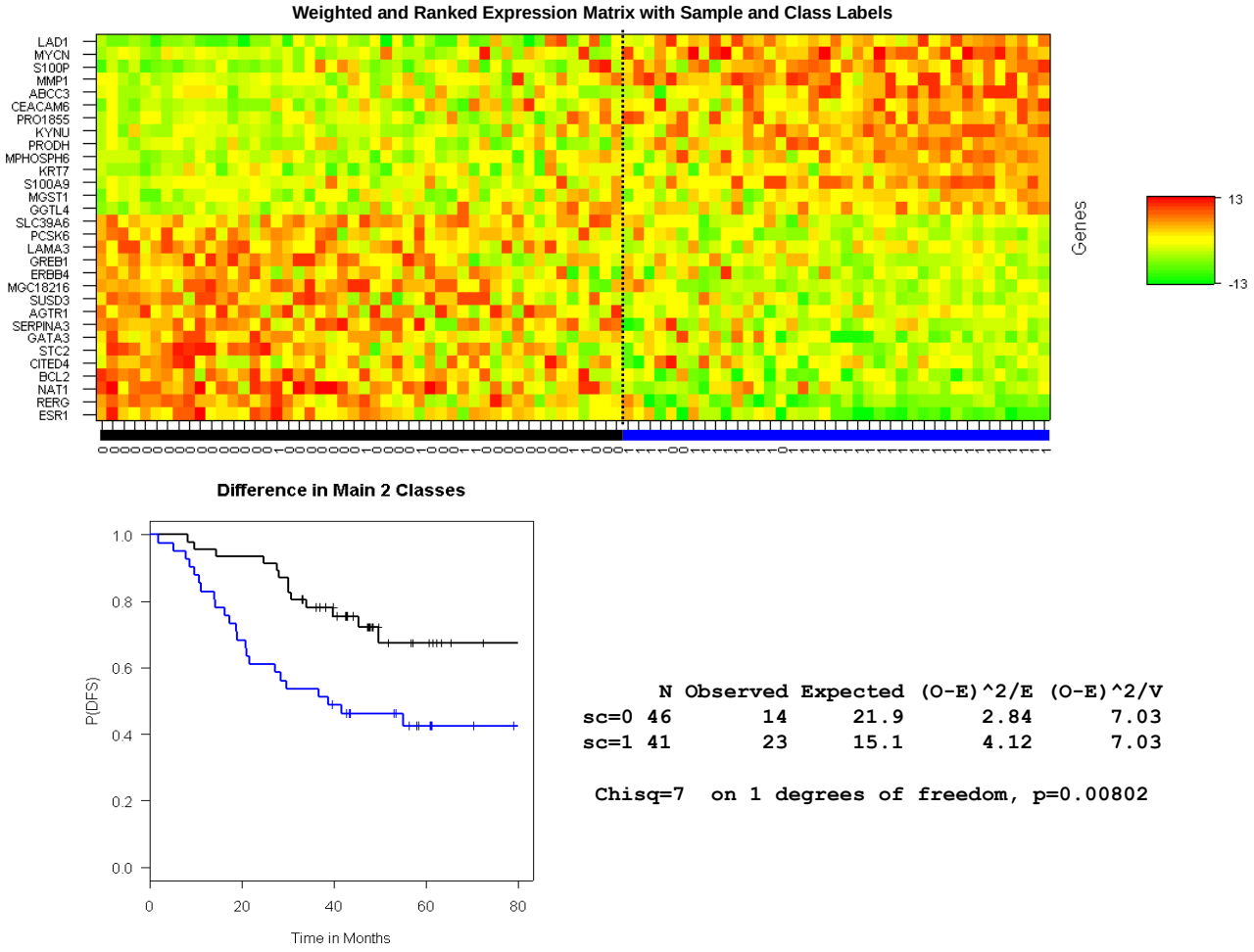


Figure 16. *Heatmap of the complete data set (n=89 samples) based on the genes and score calculated according to the classifier (above). Samples are aligned in horizontal, genes in vertical order and were ranked according to the t-statistic (genes) or score (samples), respectively. The actual (true) labels of the ERBB2 status are given below (0=ERBB2-, 1=ERBB2+). Genes from the "ERBB2-amplicon" (17q12-21) were excluded from the analysis. Moreover, the two main groups show significant differences in survival as shown in the Kaplan-Meier curve below. The genes discriminating ERBB2+/- samples contained many known cancer-related genes (e.g. MYCN, S100P, MMP1, CEACAM6, etc) as well as ER-related genes (ESR1, RERG, BCL2, GREB1, ERBB4, etc.).*

Example ERBB2: 3-fold crossvalidation, top 400 genes, n.boot=200, 17q11-21 genes eliminated

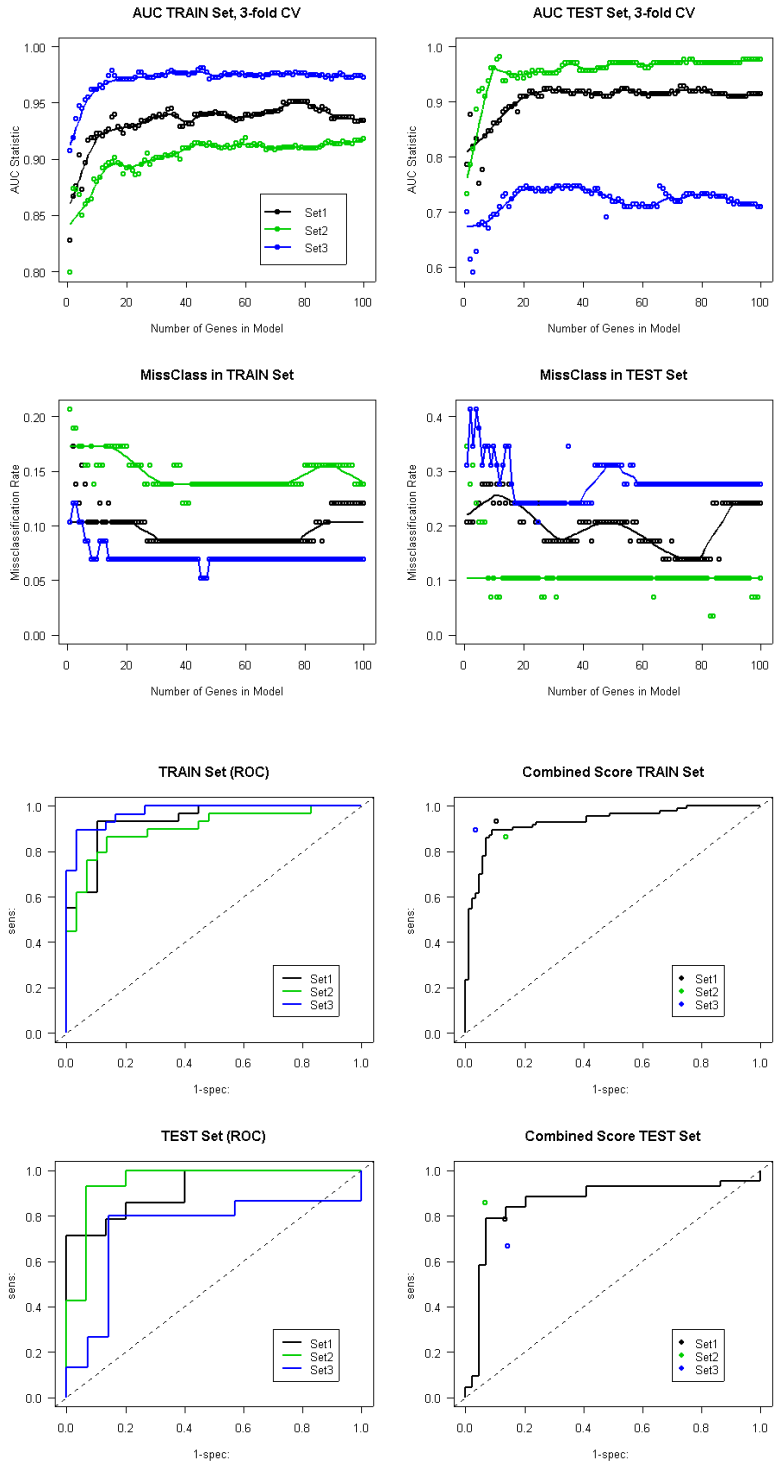


Figure 17. *Example of the performance of the classifier in the test and training set using ERBB2 status as class variable and 3-fold cross-validation (the three subsets of data are colored in black, green and blue). All genes from the "ERBB2-amplicon" were excluded. The AUC statistic reached typically a plateau when 20 or more genes were included in the classifier. The misclassification rate was lowest when using around 30 or more genes. The combined ROC curves for the test and training set might be used to select a specific cutoff for the test according to the desired or predefined requirements for the classifier (sensitivity versus 1-specificity are plotted in these curves). In this example, the test might be selected to have >90% specificity with around 80% sensitivity.*

The same classification algorithm was further tested for its ability to classify tumors according to the nodal status of the patient. With around 40 genes the sensitivity reached 80% and specificity over 90% (Figure 18). However, when the same genes and score were applied to the test set, the misclassification rate increased to almost 50%. Thus, the classifier did not accurately predict the nodal status of a patient based on the gene expression profile of the primary tumor. Finally, similar analysis of genes separating histological types (ductal versus lobular) showed that e-cadherin was the most prominent gene. This finding has been described in the literature before. Loss of e-cadherin expression might further account for the typical "indian files" pattern of lobular breast cancers, where individual tumor cells rather than clusters or tubular structures prevail.

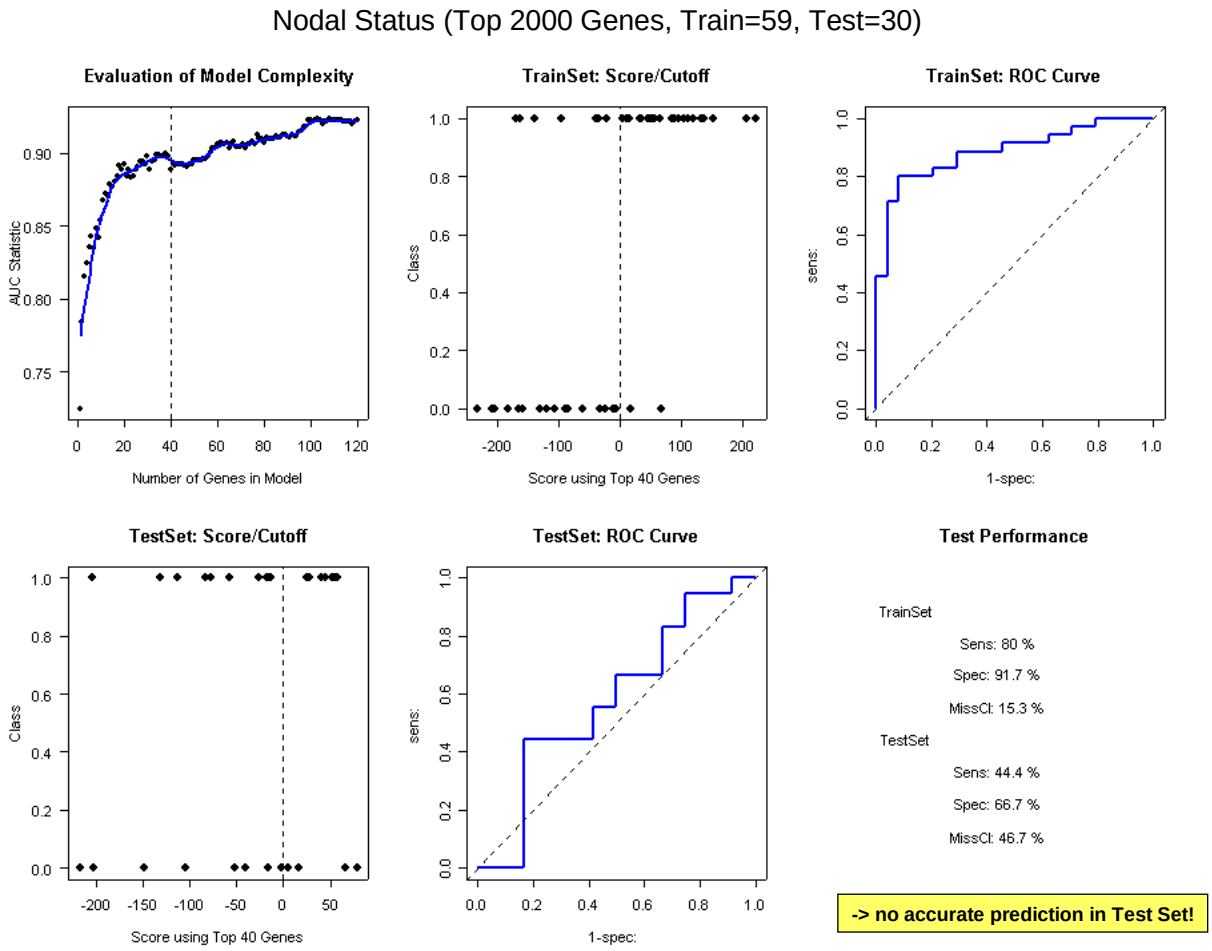


Figure 18. *Example of the same classifier used above to classify tumors according to the nodal status. As shown in the training set, by using around 40 genes the sensitivity reached 80%, specificity >90%. However, when the same genes and score were applied to the test set the misclassification rate increased to almost 50%. Thus, the classifier could not accurately predict the nodal status of a patient based on the gene expression profile of the primary tumor.*

Evaluation and validation of mRNA expression markers capable of identifying ERBB2+ breast cancer patients associated with distant metastasis and reduced survival (follow-up of microarray findings)

The prognostic value of 60 biomarkers detected in 317 primary breast cancer patients (STB data set) was assessed by qrt-PCR individually in ERBB2+/- patients by univariate Cox regression revealing five candidates (uPA, MMP11, uPAR, MMP1, and MMP3) significantly associated with metastasis-free survival in ERBB2+ samples (MFS, Figure 19 and 20) [100]. All five genes encode proteases and their levels of mRNA expression correlated strongly with one another. In contrast, none of these genes showed significant prognostic value in ERBB2- patients. Survival in this subgroup of patients was mainly driven by proliferation genes (E2F1, TOP2A, Survivin, TYMS, TK1, Cyclin D, etc.) [100]. Possible confounding by treatment, ER status or proliferation was excluded after stratified Cox analysis. The prognostic value of the most significant gene (uPA) was subsequently successfully validated in independent microarray data sets (Amsterdam [34] and Rotterdam [8] data sets).

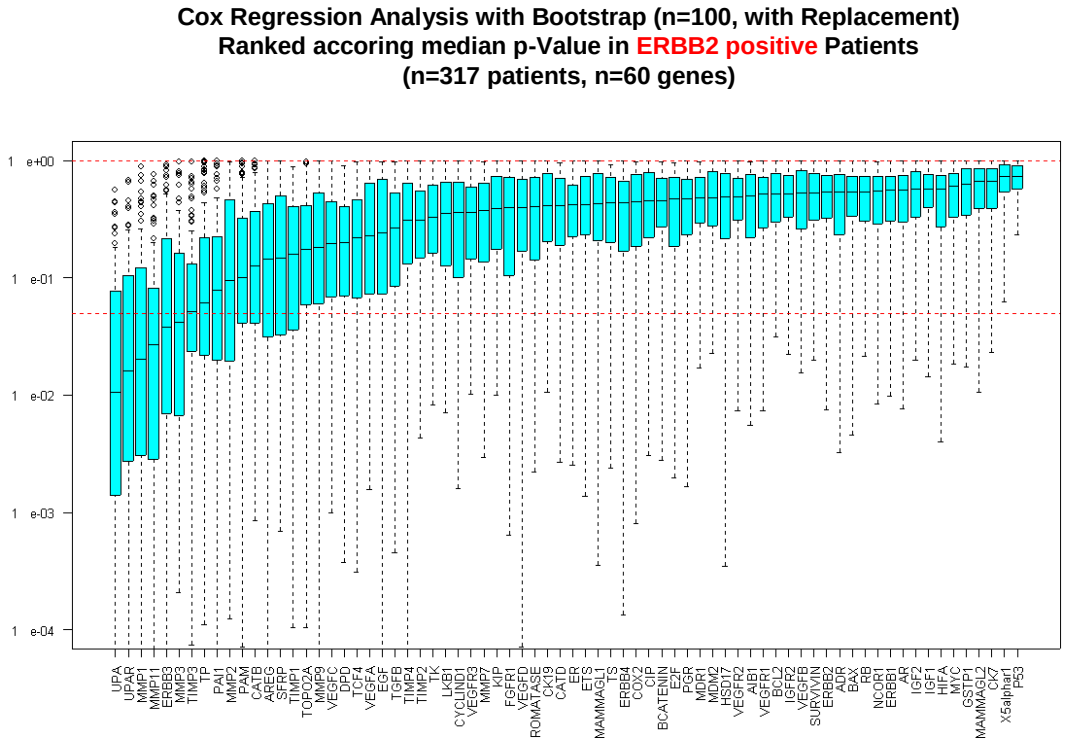


Figure 19. *Association of 60 selected cancer-associated genes detected by art-PCR with survival (MFS) in a subset of 82 ERBB2+ primary breast cancer samples. Proteases (uPA, uPAR, MMP1, MMP11, MMP3, TIMP3, etc) were among the most significant genes in this subgroup of patients.*

Multivariate Cox analysis included nodal status, grade, tumor size, age, hormone receptor status and proliferation. In all three study sets uPA demonstrated independent prognostic value and was significantly associated with MFS in ERBB2+ breast patients. Besides uPA, nodal status retained independent prognostic value in both the STB and Amsterdam study set, respectively. Of note, uPA proved not to be significantly associated with MFS in ERBB2- patients in any of the three study sets. Cutoff values for uPA status were chosen to identify ERBB2-positive cases with poor prognosis (5-year MFS of 50% or less) after evaluation of the 5-year MFS as a function of uPA mRNA expression. The cutoff values determined by this method revealed comparable cutoffs for uPA among the three data sets. Subsequently, ERBB2+ and ERBB2- patients were dichotomized by uPA status and Kaplan-Meier analyses performed. Within each study set, uPA status proved to be a strong prognostic factor for the development of distant metastasis but only among ERBB2-positive breast cancer patients (hazard ratios: STB study 3.1 [95% CI 1.3-7.4]; Amsterdam study 3.5 [1.5-8.1]; and Rotterdam study 2.8 [1.1-7.0]; all $p<0.02$; Figure 21). The prognostic value of uPA overexpression was even more pronounced for overall survival among ERBB2+ Amsterdam patients (HR=4.6 [1.9-11.5]; $p<0.001$). A search for alternative cutoff values for uPA failed to identify any level of

uPA mRNA expression significantly associated with MFS among ErbB2- breast cancer patients in any of the three study sets [100].

ER expression was significantly lower in ERBB2+ as compared to ERBB2- tumors. However, no significant difference was found with respect to ER status, nodal status, tumor size, age or treatment group between the dichotomized ERBB2+ /uPA- and ERBB2+/uPA- patients in any of the three study sets. In addition, we investigated the 70-genes prognostic signature [34] which classified almost all ERBB2+ patients into the poor prognosis group and revealed no difference in regard of the uPA status among ERBB2+ patients.

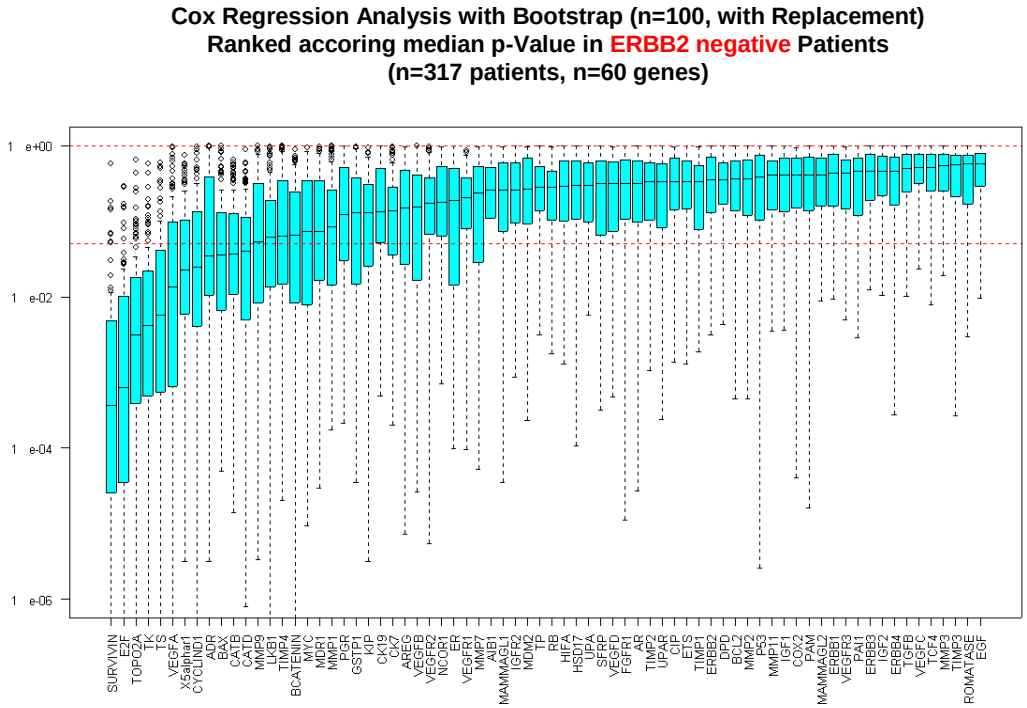


Figure 20. *Association of 60 selected cancer-associated genes detected by art-PCR with survival (MFS) in a subset of 245 ERBB2- primary breast cancer samples. Proliferation genes (Survivin, E2F1, TOP2A, TK, TYMS, Cyclin D, etc) are among is the most significant feature in this subgroup of patients.*

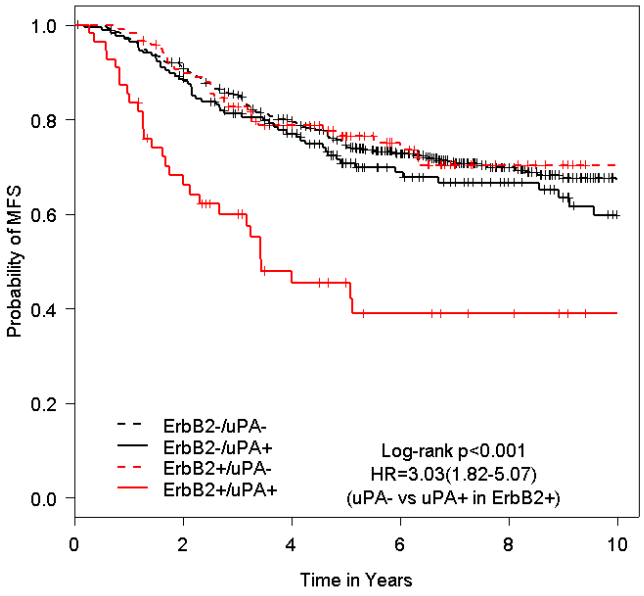


Figure 21. *Kaplan-Meier analysis (MFS) of ERBB2 combined with uPA status (meta-analysis of three independent breast cancer data sets comprising 898 samples; cutoff for uPA was 25th percentile, for ERBB2 based on mixture model) [100].*

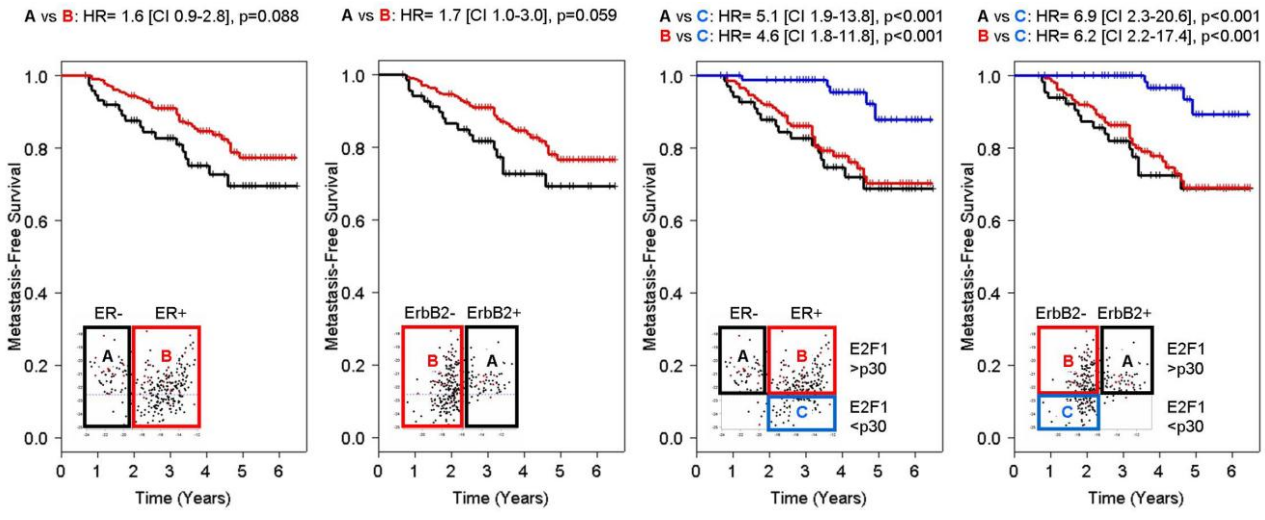
Evaluation of proliferation as a key biological process relevant to cancer: associations with different breast cancer phenotypes and survival

Proliferation was selected as an important hallmark of cancer and determinant of cancer outcome and evaluated in different subsets of patients [101,102]. Increased tumor cell proliferation is frequently accompanied with cell matrix remodeling and neoangiogenesis which together with a fast growth rate are the biological basis of an aggressive tumor phenotype. We focused on E2F1, BIRC5, TK1, TYMS and TOP2A as surrogate markers for proliferation [99].

The prognostic value of proliferation genes was first evaluated in our data set of 317 breast cancer patients by univariate Cox regression: all five markers revealed very similar prognostic value and correlated significantly with distant metastasis-free survival (MFS). However, subsequent analysis in subgroups of ER+/ER- or ERBB2+/- patients revealed that proliferation was significantly associated with MFS in ER+ and ERBB2-, but not in ER- or ERBB2+ patients (Figure 22) [99]. ER- and ERBB2+ tumors expressed almost exclusively high proliferation levels while ER+ or ERBB2- patients contained a subgroup of tumors with low proliferation levels. Analogous findings were made in independent microarray data sets (Amsterdam and Rotterdam data sets). However, the difference was less pronounced in the Rotterdam data set which contained a smaller population of low proliferating ER+ or ERBB2- tumors.

The prognostic value proliferation was further assessed in multivariate analysis retaining independent prognostic value in contrast to ER or ERBB2 status. Moreover, we included a previously established 70-gene prognostic signature which was successfully used to discriminate patients with good and poor prognosis [7,34]: 70-gene signature and proliferation showed similar prognostic value and were together with tumor size, age and nodal status independent prognostic markers. Subsequently, we used proliferation, ER and ERBB2 status for Kaplan-Meier analysis (Figure 28). In all three study populations, patients with low proliferation correlated with a significantly better prognosis (MFS). However, while the 5-year MFS of patients with tumors expressing low proliferation levels was very good for patients in the STB (88%) Amsterdam (94%) data sets, it was less favorable in the Rotterdam data set (74%). Moreover, the hazard ratio was 4.3 ($p<0.001$) and 5.1 ($p<0.001$) in the former two data sets, but only 1.6 ($p=0.041$) in the latter one [99]. Since the proportion of low and intermediate grade tumors was much smaller in the Rotterdam data set, the cutoff (30th percentile) was probably too high to identify the subgroup of low-proliferating tumors. In contrast, lowering the cutoff yielded very comparable results. Kaplan-Meier analysis of ER+/- and ERBB2+/- patients revealed no difference once the subgroup of low proliferating tumors was removed (Figure 22).

Finally, comparison of proliferation and a previously published 70-gene prognostic signature [7] revealed that patients assigned to the good-prognosis group by gene signature typically expressed low proliferation levels and were ER+ (Figure 23) [99]. In addition, there was good correlation between proliferation and the correlation coefficients used by the authors to classify patients into the good- and poor-prognosis groups. E2F1 and 70-gene prognostic signature and E2F1 showed also very similar outcomes in Kaplan-Meier analysis (Figure 23) [99]. Tumors classified by both into the low-risk group were at lowest risk to develop metastasis. Similar results were obtained with respect to the prognostic profiles defined by the intrinsic subtypes gene classification, the wound response signature and the Oncotype DX™ recurrence score (Figure 23).



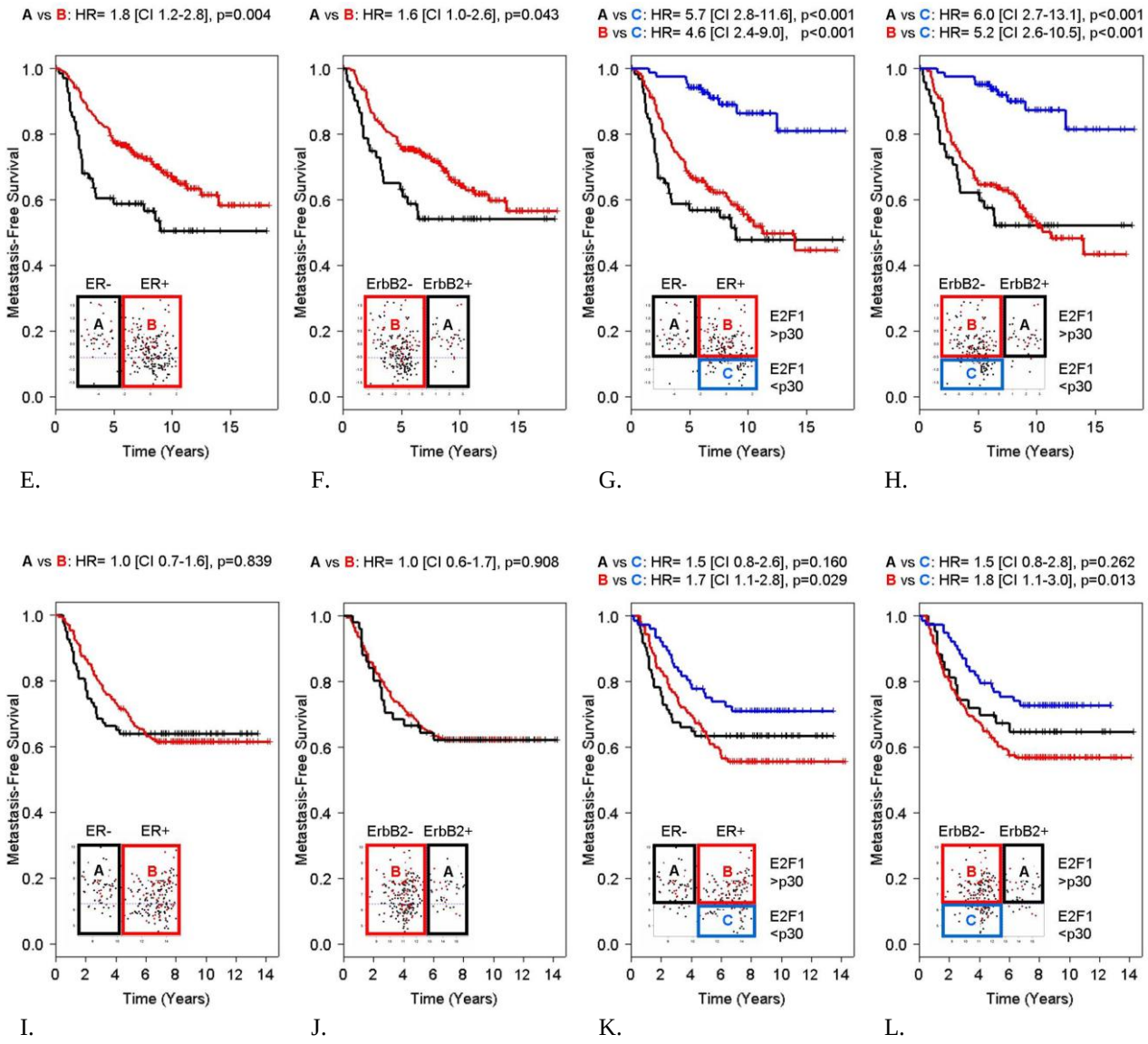
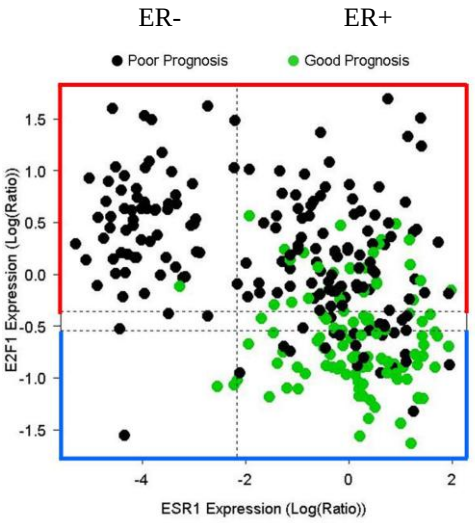
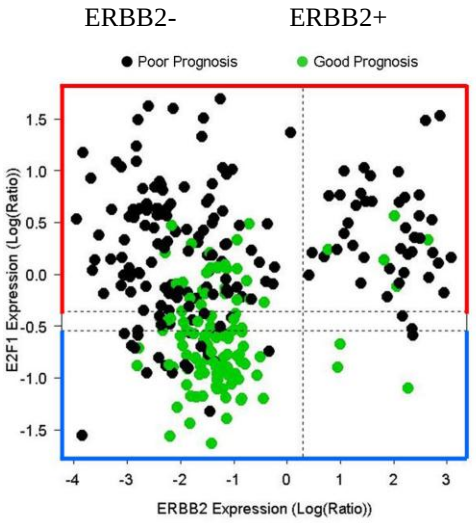


Figure 22. *Kaplan-Meier Analysis (MFS). Top: STB data set, middle: Amsterdam data set, bottom: Rotterdam data set. ER status (Fig. A, E, I), ERBB2 status (Fig. B, F, J), combination of ER status and E2F1 (Fig. C, G, K), combination of ERBB2 status and E2F1 (Fig. D, H, L). Pictograms correspond to the scatter plots shown in Figure 1. Color code indicates the patient population analyzed in the corresponding survival curve* [99].

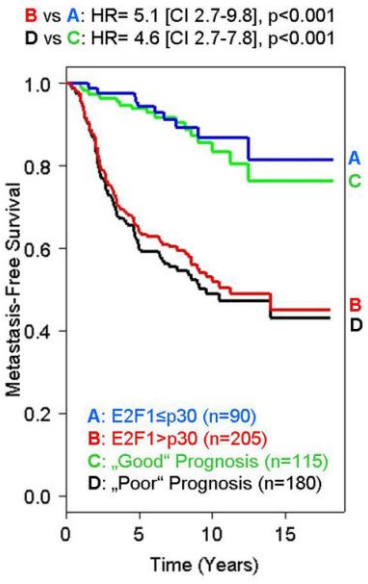
Figure 23 (next page). *Comparison of E2F1 and 70-gene signature in the Amsterdam data set: Scatter plots of estrogen receptor (ESR1) and ERBB2 versus E2F1 expression levels (Fig. A and B). Horizontal lines indicate the 30th and 39th percentile for E2F1, respectively.*
Kaplan-Meier analysis (MFS) of the Amsterdam data as reported by Fan et al. [51]*: E2F1 (30th percentile cutoff) versus 70-gene signature (Fig. C), luminal subtypes (Fig. D), recurrence score (Fig. E) and wound response signature (Fig. F)* [99].


ER-
ER+
Poor Prognosis
Good Prognosis
E2F1 Expression (Log(Ratio))
ESR1 Expression (Log(Ratio))

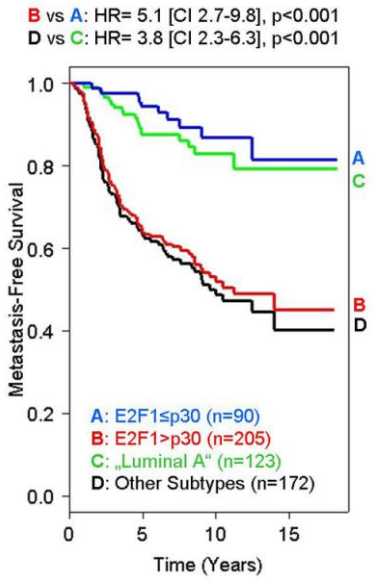
A.


ERBB2-
ERBB2+
Poor Prognosis
Good Prognosis
E2F1 Expression (Log(Ratio))
ERBB2 Expression (Log(Ratio))

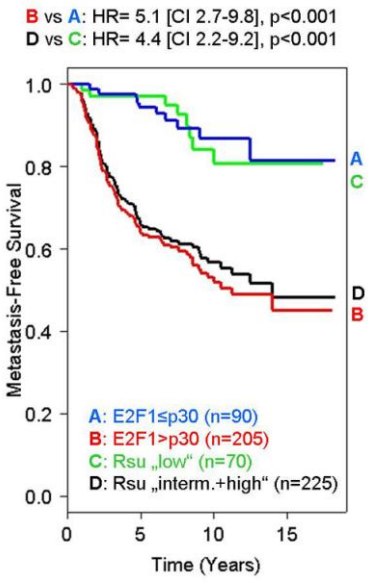
B.


B vs A: HR= 5.1 [CI 2.7-9.8], p<0.001
D vs C: HR= 4.6 [CI 2.7-7.8], p<0.001
Metastasis-Free Survival
Time (Years)
A: E2F1≤p30 (n=90)
B: E2F1>p30 (n=205)
C: „Good" Prognosis (n=115)
D: „Poor" Prognosis (n=180)

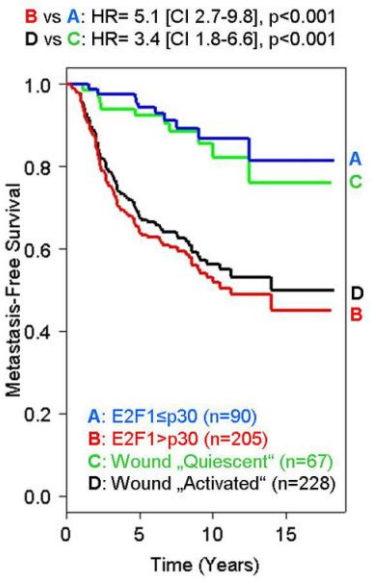
C.


B vs A: HR= 5.1 [CI 2.7-9.8], p<0.001
D vs C: HR= 3.8 [CI 2.3-6.3], p<0.001
Metastasis-Free Survival
Time (Years)
A: E2F1≤p30 (n=90)
B: E2F1>p30 (n=205)
C: „Luminal A" (n=123)
D: Other Subtypes (n=172)

D.


B vs A: HR= 5.1 [CI 2.7-9.8], p<0.001
D vs C: HR= 4.4 [CI 2.2-9.2], p<0.001
Metastasis-Free Survival
Time (Years)
A: E2F1≤p30 (n=90)
B: E2F1>p30 (n=205)
C: Rsu „low" (n=70)
D: Rsu „interm.+high" (n=225)

E.


B vs A: HR= 5.1 [CI 2.7-9.8], p<0.001
D vs C: HR= 3.4 [CI 1.8-6.6], p<0.001
Metastasis-Free Survival
Time (Years)
A: E2F1≤p30 (n=90)
B: E2F1>p30 (n=205)
C: Wound „Quiescent" (n=67)
D: Wound „Activated" (n=228)

F.

Evaluation of gene expression profiles detected in core biopsies and corresponding surgical biopsy

Gene expression profiles of 22 consecutive breast cancer patients, for whom both core (CB) and surgical biopsy (SB) were available, were evaluated by qrt-PCR for a panel of over 60 cancer-related genes [98]. All core biopsies were obtained under sonographic guidance. After establishing the diagnosis of breast cancer, all patients underwent breast surgery with sentinel node lymphonodectomy or axillary lymphonodectomy if indicated.
Hierarchical clustering revealed that paired CB and SB generally clustered together. Interestingly, two separate CBs taken from the same patient were very similar to each other, however, differed from their surgical GEP. Subsequent analysis of paired CB and SB showed high correlation (r_s from 0.86 to 0.98, all $p<0.001$) [98]. Gene-by-gene analysis by means of paired *t*-tests showed significant differences between CB and SB for four genes: PAI-1, COX-2, uPAR and MMP1 demonstrated all increased levels in the surgical sample [98]. The increase in the expression of these genes was not related to the time-frame between CB and surgery. All other genes were very similarly expressed in paired CB / SB. In order to verify whether the higher expression levels of PAI-1 and COX-2 observed in SBs could have been induced by the preceding CB procedure, expression levels of the same genes were compared in an independent population of over 300 primary breast cancer patients. These samples were investigated with the same technique but had no CB taken prior to surgery. However, there was no significant difference between the expression levels of PAI-1 or COX-2 measured in surgical samples of the present and the reference study population. More importantly, no variation was observed for the remaining genes, particularly with respect to clinically used markers such as ER, PgR and ERBB2 between CB and SB [98].

p-Y1248-ERBB2 levels were measured with a two-site chemiluminescence-linked immunosorbent assay (CLISA) in 70 primary breast cancer patients [93]. Statistical analysis aimed at investigating relationship with known prognostic markers, survival, other ERBB family members measured on protein and mRNA level as well as microarray gene expression data.
p-ERBB2 correlated strongly with ERBB2 levels, both on RNA (r=0.62) and protein (r=0.53) level. When the median value of p-ERBB2 was taken as cutoff for p-ERBB2 status, 8 out of 30 (27%) ERBB2- and 27 out of 40 (68%) ERBB2+ tumours were considered p-ERBB2+ (p=0.001). No significant association was found between p-ERBB2 levels and nodal status. Further, p-ERBB2 expression levels correlated inversely with ER (r_s=-0.54) and PgR (r_s=-0.46) mRNA and protein expression levels (r_s=-0.67 and r_s=-0.45), and with EGFR (r_s=0.26) and ERBB4 (r_s=-0.47) mRNA levels. The negative correlation between p-ERBB2 and ERBB3 mRNA was statistically not significant [93].
Univariate Cox regression analysis revealed significant correlation DFS and OS. Interestingly, p-ERBB2 retained significant prognostic value in both, ERBB2- and ERBB2+ patients. Subsequently, Kaplan-Meier showed poorer outcome for p-ERBB2+ patients (p=0.004); this was more pronounced in nodal positive than nodal negative patients. Of note, p-ERBB2 status was significantly associated with DFS in the subgroup of ERBB2- patients. Further, p-ERBB2 was an independently associated with survival in multivariate analysis including ER, EGFR, ERBB2, pT, pN and age [93].
Gene expression analysis from microarray data revealed that p-ERBB2 correlated strongly with the "ERBB2 amplicon" genes (17q11-21, Table 3). Functional annotation showed strongest association with cell growth/cell cycle (Figure 24). Generally, there was strong overlap between genes associated with ERBB2 and p-ERBB2 status, however, also significant differences with respect to differentially expressed genes (Figure 25). No correlation was found between p-ERBB2 and p-AKT.

Spearman Correlation, pERBB2, n=Top 408 Genes

Spearman	Gene	ChromLoc	Description	
0.621072947	STARD3	17q11-q12	START domain containing 3	
0.548218287	ERBB2	17q11.2-q12	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2	
0.523130546	GRB7	17q21.2	growth factor receptor-bound protein 7	
0.44120062	LAD1	1q25.1-q32.3	ladinin 1	
0.439233821	MGST1	12p12.3-p12.1	microsomal glutathione S-transferase 1	
0.430266886	ABCC3	17q22	ATP-binding cassette, sub-family C (CFTR/MRP), member 3	
0.418729752	S100P	4p16	S100 calcium binding protein P	
0.407463252	EME1	17q21.33	essential meiotic endonuclease 1 homolog 1 (S. pombe)	
0.399504306	PSMD3	17q21.2	proteasome (prosome, macropain) 26S subunit, non-ATPase, 3	
0.394379084	KRTHB3	12q13	keratin, hair, basic, 3	
0.37967905	RGS16	1q25-q31	regulator of G-protein signalling 16	
0.374915503	CXCL10	4q21	chemokine (C-X-C motif) ligand 10	
0.359018757	DKK1	10q11.2	dickkopf homolog 1 (Xenopus laevis)	
0.336458627	COMP	19p13.1	cartilage oligomeric matrix protein	
0.321397667	KYNU	2q22.3	kynureninase (L-kynurenine hydrolase)	
0.303861791	PRODH	22q11.21	proline dehydrogenase (oxidase) 1	
0.293879909	MMP1	11q22.3	matrix metalloproteinase 1 (interstitial collagenase)	**Positive Correlation**
0.289961867	THRAP4	17q21.2	thyroid hormone receptor associated protein 4	
0.287320658	CKMT1	15q15	creatine kinase, mitochondrial 1 (ubiquitous)	
0.279565149	MMP13	11q22.3	matrix metalloproteinase 13 (collagenase 3)	
0.278667604	MAGEA2	Xq28	melanoma antigen, family A, 2	
0.275241492	SLC16A3	17q25	solute carrier family 16 (monocarboxylic acid transporters), member 3	
0.273401496	BIRC5	17q25	baculoviral IAP repeat-containing 5 (survivin)	
0.264555977	MGLL	3q21.3	monoglyceride lipase	
0.262954479	PHB	17q21	prohibitin	
0.261669378	RGS1	1q31	regulator of G-protein signalling 1	
0.260438319	MYCN	2p24.1	v-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian)	
0.259619089	CCL18	17q11.2	chemokine (C-C motif) ligand 18 (pulmonary and activation-regulated)	
0.258964359	AQP5	12q13	aquaporin 5	
…				
…				
-0.545236745	MAPT	17q21.1	microtubule-associated protein tau	
-0.542544518	MGC18216	15q26.3	hypothetical protein MGC18216	
-0.528608083	ESR1	6q25.1	estrogen receptor 1	
-0.505277889	STC2	5q35.2	stanniocalcin 2	
-0.466710255	AGTR1	3q21-q25	angiotensin II receptor, type 1	
-0.451391476	RERG	12p13.1	RAS-like, estrogen-regulated, growth-inhibitor	
-0.441568495	KIF5C	2q23.3	kinesin family member 5C	
-0.43962629	ALB	4q11-q13	albumin	**Negative Correlation**
-0.436178559	BCL2	18q21.33	B-cell CLL/lymphoma 2	
-0.420511596	LAF4	2q11.2-q12	lymphoid nuclear protein related to AF4	
-0.407230032	NAT1	8p23.1-p21.3	N-acetyltransferase 1 (arylamine N-acetyltransferase)	
-0.402698658	BCL2	18q21.33	B-cell CLL/lymphoma 2	
-0.402561979	ABCA3	16p13.3	ATP-binding cassette, sub-family A (ABC1), member 3	
-0.39309137	SCNN1B	16p12.2-p12.1	sodium channel, nonvoltage-gated 1, beta (Liddle syndrome)	
-0.385832366	FLJ20366	8q23.2	hypothetical protein FLJ20366	
-0.384577726	SERPINI1	3q26.2	serine (or cysteine) proteinase inhibitor, clade I (neuroserpin), member 1	
-0.381991221	COL4A6	Xq22	collagen, type IV, alpha 6	
-0.368563804	GLRB	4q31.3	glycine receptor, beta	
-0.363205242	GATA3	10p15	GATA binding protein 3	
-0.36075752	KCNE4	2q36.3	potassium voltage-gated channel, Isk-related family, member 4	
-0.356178203	CCNG1	5q32-q34	cyclin G1	

Table 3. Correlation of p-ERBB2 expression levels (CLISA) with genes expression levels measured by microarray. Genes from the "ERBB2-amplicon" are among the strongly correlating genes, revealing a similar pattern like ERBB2.

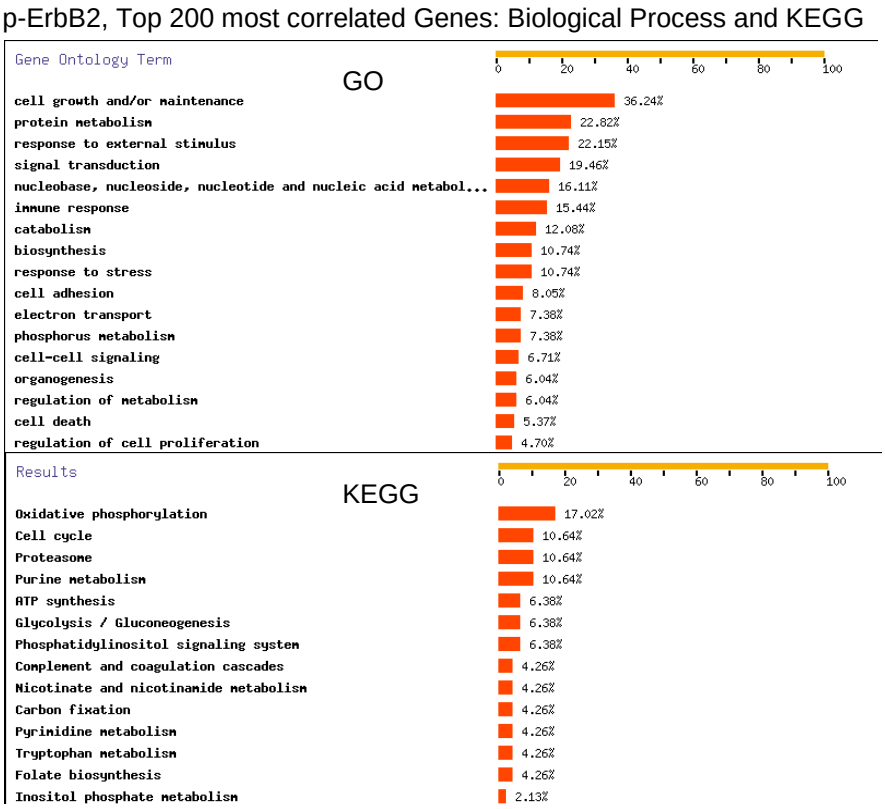


Figure 24. *Functional annotation of the most correlated genes with p-ERBB2 (GO: gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes).*

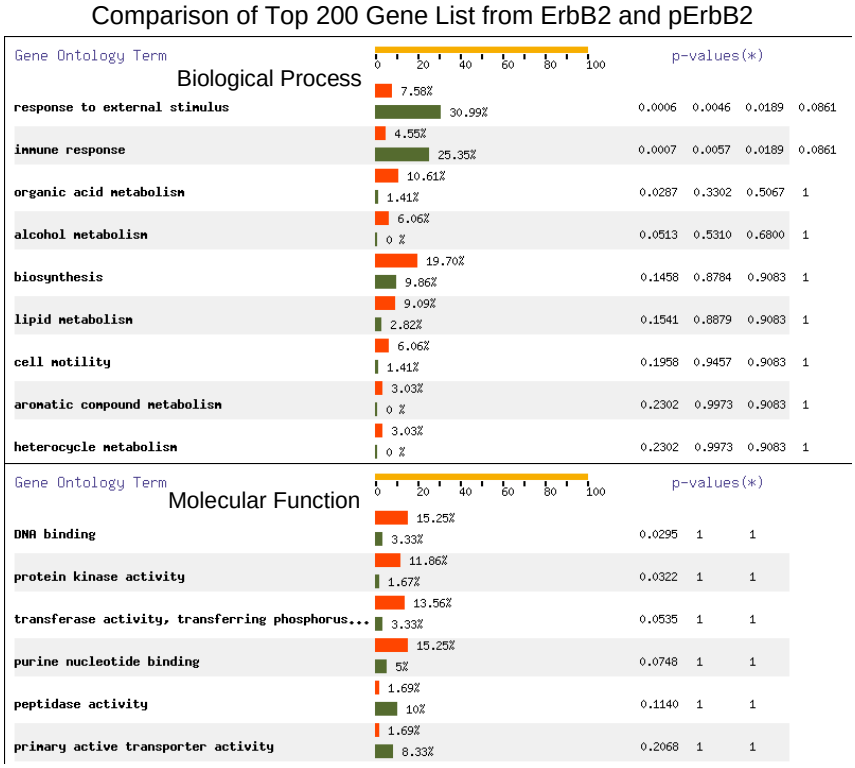


Figure 25. *Comparative annotation of the 200 top genes associated with ERBB2 and p-ERBB2, respectively.*

Quantitative expression levels of total phosphorylated p-S473-AKT (AKT1/2/3) were measured by means of a two-site chemoluminescence-linked immunosorbent assay (CLISA) on cytosol extracts from 156 primary breast cancer samples [94]. There was no association between p-AKT and nodal, ER, ERBB2 or p-ERBB2 status, and tumor size. Survival analysis revealed that only very high expression levels of P-Akt correlated ($p < 0.01$) with poor prognosis. Interestingly, the prognostic value of P-AKT expression was more pronounced in the subset of ERBB2+ patients ($p<0.001$) [94]. Subsequently, p-AKT expression levels were correlated with mRNA expression levels measured by qrt-PCR of more than 60 cancer related biomarkers involved processes such as proliferation, hormone dependency, apoptosis, angiogenesis, invasion and metastasis: association was found only with markers of proliferation (TYMS, TK1, survivin (BIRC5), TOP2A and E2F1, all $p<0.05$). In agreement with this observation, microarray analysis in a subset of 65 biopsies underlined the correlation of p-AKT with proliferation (Table 4, and Figures 26 and 37). A schematic overview about the signaling pathways downstream of AKT impacting proliferation is summarized in Figure 28.

Correlation (Spearman): pAkt with n=3839 Top Genes from Array

Negative Correlation:

Spearman	Gene	ChromLoc
-0.344668384	THRA	17q11.2
-0.31914751	PPARGC1A	4p15.1
-0.311787513	NDN	15q11.2-q12
-0.300725544	IGHM	14q32.33
-0.299163873	NUCB1	19q13.2-q13.4
-0.297477382	RAFTLIN	3p25.1
-0.294386832	IGF1	12q22-q23
-0.292595104	C1S	12p13
-0.29196375	MGC16044	12q13.11
-0.289853428	ITM2B	13q14.3
-0.289590963	PTGS1	9q32-q33.3
-0.285231143	HP	16q22.1
-0.285153016	RRAS	19q13.3-qter
-0.284495267	EBAF	1q42.1
-0.27863175	IGFBP7	4q12
-0.276138654	C1S	12p13
-0.272607212	HLA-DRB3	6p21.3
-0.270893911	IL2RB	22q13
-0.265024927	LOC387763	11p11.2
-0.25924454	CILP	15q22
-0.259032741	ALG3	3q27.3
-0.256341385	ADPRTL1	13q11
-0.256238626	ARGBP2	NA
-0.254614188	CDKN2A	9p21
-0.25346353	FLRT2	14q24-q32
…		

Positive Correlation:

Spearman	Gene	ChromLoc	Description
0.508204961	TYMS	18p11.32	thymidylate synthetase
0.4754085	HMGB2	4q31	high-mobility group box 2
0.384736698	KIF22	16p11.2	kinesin family member 22
0.362103958	CENPF	1q32-q41	centromere protein F, (mitosin)
0.344425276	ZNF354A	5q35.3	zinc finger protein 354A
0.338594473	RANBP17	5q34	RAN binding protein 17
0.328902674	RAD54L	1p32	RAD54-like (S. cerevisiae)
0.320870264	CDC2	10q21.1	cell division cycle 2, G1 to S and G2 to M
0.319418346	UHRF1	19p13.3	ubiquitin-like, containing PHD and RING finger
0.316854051	ZNF200	16p13.3	zinc finger protein 200
0.315328926	EZH2	7q35-q36	enhancer of zeste homolog 2 (Drosophila)
0.315020795	CBX1	17q	chromobox homolog 1 (HP1 beta homolog Drosophi
0.314501055	DDX11	12p11	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 1
0.314345757	CGI-01	1q24-q25.3	CGI-01 protein
0.310300072	FEN1	11q12	flap structure-specific endonuclease 1
0.308545454	TP53BP2	1q42.1	tumor protein p53 binding protein, 2
0.308525889	SFRS1	17q21.3-q22	splicing factor, arginine/serine-rich 1
0.306824588	KIF20A	5q31	kinesin family member 20A
0.305441749	LUC7L	16p13.3	LUC7-like (S. cerevisiae)
0.302252143	PCNA	20pter-p12	proliferating cell nuclear antigen
0.302187956	STMN1	1p36.1-p35	stathmin 1/oncoprotein 18
0.301859198	C16orf34	16p13.3	chromosome 16 open reading frame 34
0.298186725	CDC2	10q21.1	cell division cycle 2, G1 to S and G2 to M
0.296003347	MGC24665	16p13.2	hypothetical protein MGC24665
0.294398685	RAD51	15q15.1	RAD51 homolog (RecA homolog, E. coli) (S. cere
0.289153311	MAD2L1	4q27	MAD2 mitotic arrest deficient-like 1 (yeast)
0.285599516	CDCA3	12p13	cell division cycle associated 3
…			

-> many genes involved in cell cycle / proliferation (DNA Replication, Mitosis, Transcription, etc)

Table 4. *Correlation of p-AKT (CLISA) with genes expression levels of 3839 genes measured by microarray. Many genes involved in proliferation / cell cycle are among the most correlated genes with phosphorylated AKT.*

Example pAkt: Top 100 correlated Genes

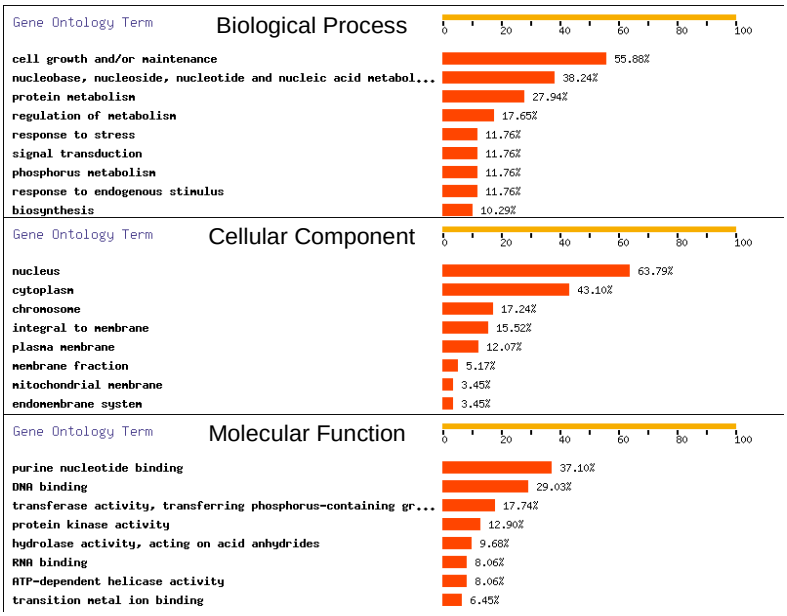


Figure 26. *Functional annotation of the most correlated genes (microarray) with p-AKT (gene ontology)*

Example: KEGG Pathways, Cell Cylce

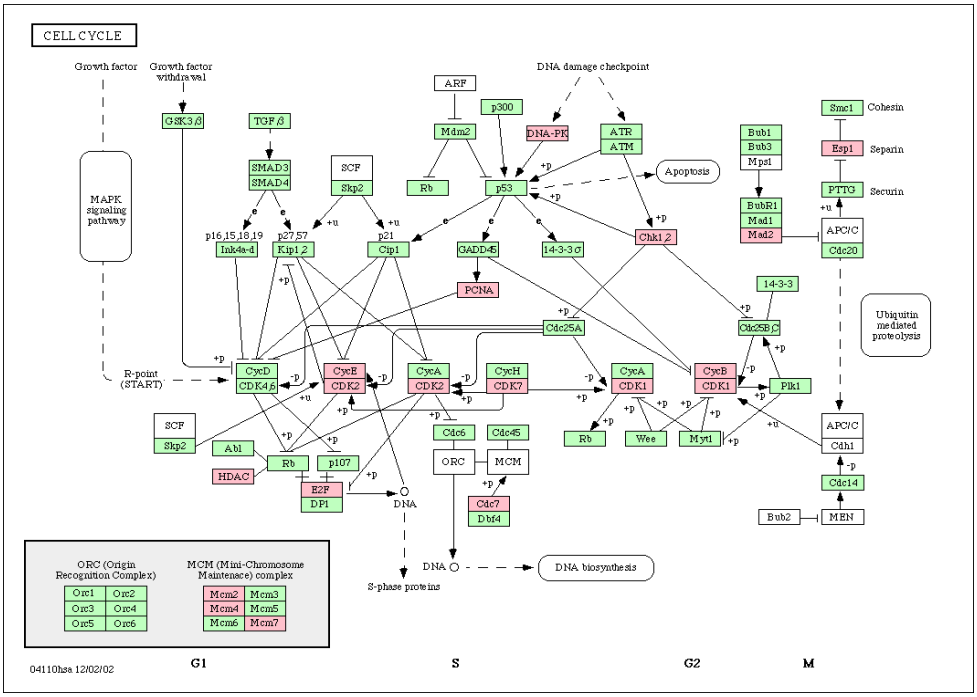


Figure 27. *Example of mapping the top 100 most correlated genes with p-AKT in KEGG pathways (Kyoto Encyclopedia of Genes and Genomes, Cell cycle). Genes underlined in red were among the top 100 genes which correlated most with p-AKT expression.*

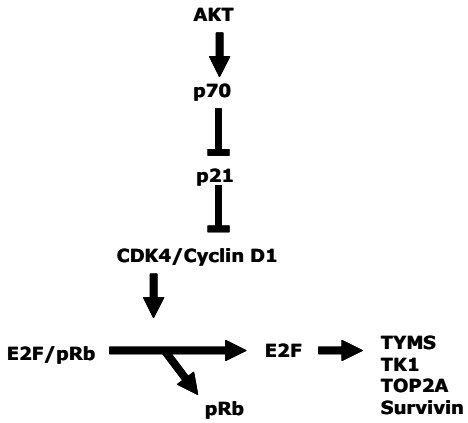


Figure 28*. Overview of AKT signaling to proliferation.*

Discussion

Depending on the stringency of the criteria used, several differentially expressed genes were identified when comparing ERBB2+ with ERBB2- tumor samples or that were associated with ERBB status. Interestingly, many of these mapped to chromosome 17q12-21 which harbors the "ERBB2-amplicon", indicating the underlying amplification is responsible for this finding. Indeed, parallel measurements of mRNA levels and gene copy number by comparative genomic hybridization (CGH) and gene expression microarrays revealed that variation in gene copy number contributes in a high degree to variation in gene expression in tumor cells [103]. Pollack *et al.* [103] found that DNA copy number influenced gene expression across a wide range of DNA copy number alterations (deletion, low-, mid- and high-level amplification) with 62% of highly amplified genes having moderately or highly elevated gene expression, and that overall, at least 12% of all the variation in gene expression among the breast tumors was directly attributable to underlying variation in gene copy number. In this context, also the systematic analysis of gene copy number and expression by CGH microarrays in breast cancer cell lines and tissues revealed co-amplification and simultaneous elevation of expression levels of several genes in addition to ERBB2 [82]. However, pattern analysis of deletions and amplification of the 17q12-21 region around the ERBB2 gene also demonstrated that different co-amplification patterns occur, where ERBB2 is not always co-amplified with other genes of this region and *vice versa* [83]. This was in line with our observations that genes from the ERBB2-amplicon did not always demonstrate simultaneous co-expression pattern of all genes from this region in the cluster analysis.

In addition to "ERBB2-amplicon" genes we identified several other differentially expressed genes associated with ERBB2 status, some of which might potentially be regulated by ERBB2 or arise through association with phenotypic features of the disease such as the cell of origin. Among these were genes related to ER status, genes involved in metastasis and invasion (e.g. proteases, S100P/A4), angiogenesis (e.g. VEGF, IL8, HIF1A), proliferation (BIRC5, TK1, CDC2, CDKN1A) and cell adhesion (e.g. CEACAM6, CDH1). Moreover, our results demonstrated that using a signature of around 30 genes could be used to accurately classify the samples according to their ERBB2 status, notably in the absence of genes from the "ERBB2 amplicon". Similar findings were reported in a microarray study by Bertucci *et al.* [47] who developed a gene-expression signature distinguishing ERBB2+/- breast tumors.

Gene expression studies of genes associated with ERBB2 status in mammary cell lines transfected with ERBB2 or cell lines harboring the "ERBB2-amplicon" demonstrated that there were many differentially expressed genes compared to normal cell lines in both cases, however, there was very little overlap between the cells transfected with ERBB2 only and cells with the whole amplicon [46]. Accordingly, there was very little overlap reported by the same authors when comparing the differentially expressed genes identified in cell lines with those differentially expressed between ERBB2+/- tumor samples. In fact, there were only two common differentially expressed genes between cell lines and tumors, one of them being ERBB2. Nevertheless, many of the differentially expressed genes mapped to similar pathways including proliferation, cell adhesion, invasion, etc., and showed some overlap with the genes identified in our breast tumor tissues. For example, S100P and S100A4 were among these genes and have been previously implicated in breast tumorigenesis and ERBB2 status [6,104,105]. Along these lines, White *et al.* [106] investigated cellular responses to ERBB2 overexpression in mammary cell lines at baseline and after stimulation with heregulin by microarrays. Among the constitutively higher or lower expressed genes, thus, not related to stimulation be heregulin were 21 up- and 27 down-regulated genes including S100P. Stimulation by heregulin led to increased MAPK signaling, increased ERBB2 (EGFR)-related autocrine signaling, and inhibition of basal IFN signaling. However, there was limited overlap with the differentially expressed genes observed in tumor tissues. Of note, this group further demonstrated that gene expression changes correlated well with protein expression levels of the same genes.

With respect to the gene expression changes associated with ERBB2 status observed in breast tumor biopsies, current data suggest that they reflect different aspects of the disease. These include the underlying genetic changes (eg. the "ERBB2 amplicon"), the cellular phenotype which itself appears to be combination of genetic changes and possibly the cell of origin, as well as the interplay between tumor tissue and stroma, vasculature and immune response. Clinically, they result in an aggressive phenotype through regulation of key processes relevant to cancer such as proliferation, invasion, angiogenesis and cell adhesion. Data from the literature demonstrated, that ERBB2 indeed can regulated and interfere with these processes, and inhibition of ERBB2 using a monoclonal antibody directed against the extracellular leads to prolonged survival.

With respect to molecular breast cancer classification, recent microarray studies demonstrated that breast cancers can be classified into molecularly distinct and clinically relevant subgroups based on variations in gene expression patterns. Sorlie *et al.* [6] used unsupervised methods and classified breast cancers into the ER-negative "basal" and "ERBB2-like" subtypes, and two ER-positive "luminal" subtypes (luminal A and B). Of note, not all ERBB2-positive tumors fall into the ERBB2-like category by microarray; few ERBB2+ samples grouped together with luminal B tumors and, thus, were ER+. Other research groups used supervised methods to identify gene signatures associated with prognosis [8,10,34]. Despite promising results and successful correlation of gene expression profiles with survival or treatment response, subsequent biological interpretation turned out to be non-trivial provoking some skepticism regarding reproducibility and consistency of the newly

proposed molecular classifiers [48,49,107,108]. However, recent data indicate that although little overlap was observed on the level of individual genes, the underlying biological processes and pathways appear to be similar [42,48,51].
Instead of correlating and mapping prognostically significant genes/signatures or molecular phenotypes to biological pathways we explored a "reverse supervised approach", and correlated known and well defined biological processes to prognostic signatures and molecular phenotypes. In addition, the analysis was extended to selected breast cancer subtypes on an individual basis allowing gaining insight into some of the common mechanisms linked with breast cancer prognosis, and point towards distinct molecular pathways being relevant according to the breast cancer subtype investigated. At the same time, we were able to shed some light on possible sources for discrepancies encountered among different prognostic genes signatures in the literature. For example, the 70- and 76-gene prognostic signatures published with the Amsterdam and Rotterdam data sets share only three genes [48,109], thus, it was speculated that differences in microarray technology or the presence of multiple signatures that all correlate with survival could be responsible for this observation [48,108]. However, our results demonstrate that there are substantial differences in tumor biology between the two study populations, in particular due to heterogeneity in sample populations. Thus, we suggest that true biological differences in the signatures might exist which are not related to technical problems or data variability.
In our analysis, proliferation emerged a key determinant of breast cancer prognosis and is particularly suitable to identify a subgroup of patients with favorable outcome. The prognostic value of established markers such as ER and ERBB2, in contrast, is strongly related to proliferation. This also applies for the molecular subtypes described by Sorlie *et al.* [6], the Amsterdam 70-gene signature [34] as well as a recently identified recurrence score [10] and the wound response signature [38,44]. Thus, proliferation seems to be one of the most "downstream" players of all of these markers in terms of prognosis. Moreover, proliferation and in particular E2F1 as a transcription factor also appear to be a good candidate markers for chemotherapy response: genes such as TK1, TOP2A and TYMS are regulated by E2F1 or associated with proliferation and are involved as direct targets or in metabolism of 5-FU and anthracycline-based therapy [39,110-113]. Indeed, recent studies demonstrated that high proliferating tumors responded better to chemotherapy [36,39]. Therefore, breast cancer classification should include accurate quantification of the proliferation status on a routine clinical basis. Current data suggest that patients with high proliferation levels then should receive systemic chemotherapy on top of endocrine treatment for ER+ and ERBB2-targeted therapy for ERBB2+ tumors. In contrast, tumors with low E2F1 levels will not require aggressive treatment. Since these tumors are ER+ endocrine therapy might be sufficient [10,48] (Figure 29).

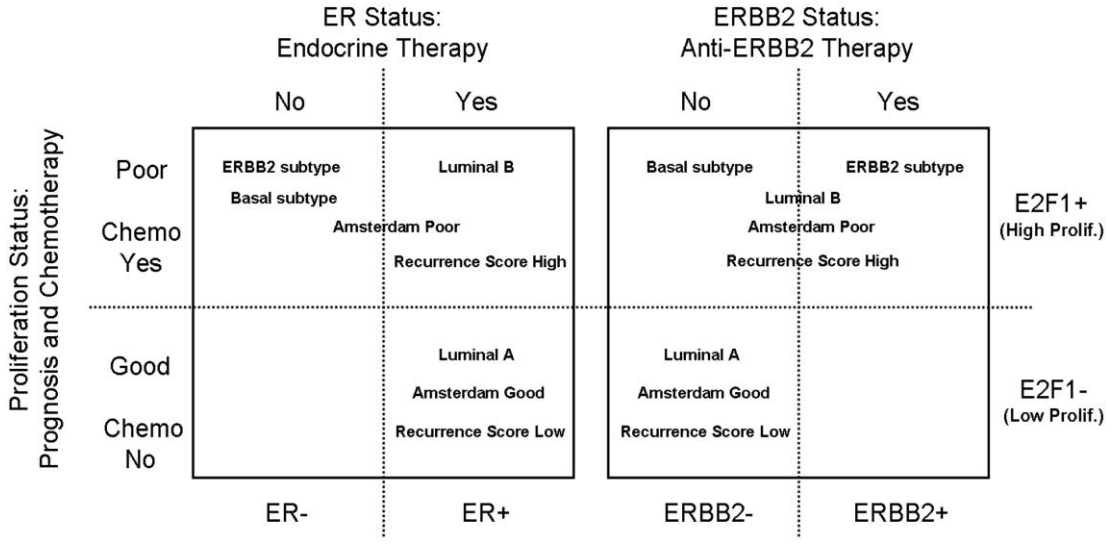


Figure 29. *Overview and tentative relationships between recently proposed molecular breast cancer classifiers and ER, ERBB2 and E2F1. Possible therapeutic implications are indicated. "Amsterdam Poor" and "Amsterdam good" (Amsterdam 70-gene signature [34]); "Luminal A", "Luminal B", "Basal subtype", and "ERBB2 subtype" (Stanford classification [6]); "Recurrence Score High" and "Recurrence Score Low" (recurrence score [10], Genomic Health/Oncotype DX ™).*

While consistency starts to emerge with respect to classification of low-risk patients efforts should continue to refine and further explore the high-proliferation, poor-prognosis subtypes. We did a first step into that direction and analyzed ERBB2+/- tumors separately. In our analysis proteases emerged as the most relevant genes with regard to progrnosis in ERBB2+ patients and further raised the question about a possible causal relationship. Indeed, there is evidence to support the view that ERBB2 can transcriptionally upregulate uPA expression and members of the MMP family [53,68,69]. This was in line with the observation that ERBB2+ tumors expressed significantly higher proteases levels as compared to ERBB2- tumors. Similar findings were reported for other cancers [114,115]. However, the mechanistic interaction linking ERBB2 with uPA upregulation and increased breast cancer metastatic potential is likely more complex. Moreover, ERBB2 overexpression occurs within tumor epithelium while uPA and other proteases are also expressed by tumor stroma [116,117], something further to consider when microdissection techniques are used for expression profiling and subsequent cancer classification. However, these results indicate that apparently only taken in conjunction, uPA and ERBB2 overexpression determine the aggressive breast cancer phenotype. This functional interaction would deserve further study at both cellular and clinical levels. Apparently, the proteolytic activity of uPA alone is insufficient to determine the most metastatic of breast cancer phenotypes and perhaps requires the proliferation and survival advantages provided by ERBB2 overexpression. Moreover, the role of other frequently co-amplified genes on the "ERBB2-amplicon" deserves further investigation.

These finding also provide some evidence to suggest that the heterogeneous nature of the breast cancer will likely require developing "individualized gene signatures" rather than one signature for all patients since the prognostic value of individual markers or gene signatures depends on the biological context of the tumor. Thus, patient stratification might be necessary to further optimize current molecular breast cancer classification. Moreover, the use of a "reverse supervised approach" in terms of assessing the contribution of individual biological processes (hallmarks of cancer) might constitute an interesting approach. Together these constitute valuable strategies towards a biology driven classification and molecular understanding of breast cancer, and facilitating the clinical interpretation with respect to the development of appropriate treatment strategies. In this context, ERBB2+ patients might be worthwhile to consider, in addition to standard therapy, for anti-uPA or MMP-targeted therapy.

Another aspect to be considered with respect to clinical application of gene expression-based molecular classification became apparent during the study of paired core and surgical biopsy from the same patient. Today ultrasound-guided core biopsy is a well established method to diagnose breast cancer. Thus, molecular profiles might potentially be performed on core biopsy material. However, the clinical utility of the information gained by CB needs to be representative for the whole tumor. Our study of paired core biopsy and surgical sample suggests that quantitative expression levels of 60 genes detected in CBs were highly comparable to their paired surgical samples. However, gene-by-gene analysis demonstrated higher expression level of PAI-1, COX-2, uPAR and MMP1 in the surgical specimen as compared to their paired CB. Tissue sampling by CB is known to cause local injury which in response will induce wound healing characterized by recruitment of inflammatory cells, stimulation of stromal and epithelial cell proliferation, cell migration, and increased angiogenesis. All four genes have been described to be essentially involved in these processes [118,119]. However, reparative processes associated with wound healing share also many parallels with cancer [120-122].While proteases, their inhibitors, cyto-/chemokines and growth factors are essential for the process of wound healing and tissue repair they play also central roles in cancer progression. For example, uPAR and PAI-1 are responsible for the degradation and remodeling of the extracellular matrix, angiogenesis, cell adhesion and migration which are necessary for tumor cell invasion and metastasis [123,124]. Both were further associated with poor clinical outcome in breast cancer [125-127]. Accordingly, a previously identified "wound-response signature" turned out to have prognostic value in several carcinomas including breast cancer [44].

Therefore, the observed alterations of individual genes might be attributable to a wound healing process induced by the core biopsy procedure. However, this might impact the clinical interpretation with respect to tumor aggressiveness and subsequent treatment decisions. Nevertheless, molecular classification of breast cancer samples is feasible and even in smallest amounts of tissue, and seems to outperform traditional classification systems. Therefore, the accurate quantification of at least the proliferation status as part of routine clinical breast cancer assessment should be encouraged.

For more detailed discussion of the results see attached publications [93,94,98-100].

A possible gene-expression-based approach towards a molecular breast cancer classification is outlined in Figures 30-35. Molecular analysis will include an algorithm to assess malignancy status, eg. whether the analyzed tumor sample is malignant or benign, followed by stratification according to ERBB2 status and individualized risk score calculation for each subgroup. Based on the prognosis and selected predictive genes, therapy will be tailored according to the specific expression changes for each patient.

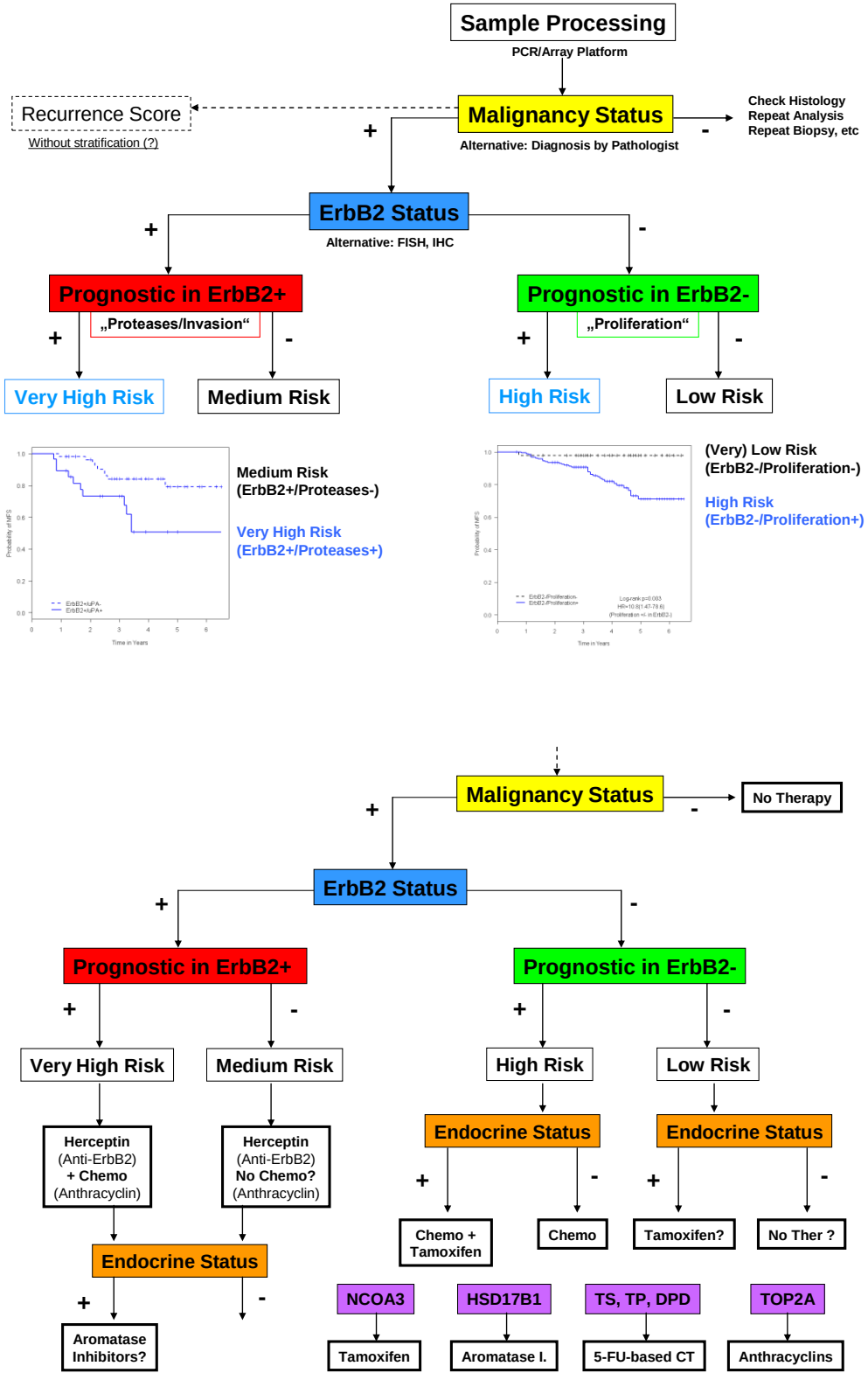


Figure 30. *Schematic for a putative classification algorithm for routine clinical diagnosis and risk assessment (above) with potential therapeutic implications (therapy prediction and selection, below). Malignancy status, prognosis, ERBB2 and ER status, and therapy prediction will be assessed using a PCR or microarray-based test with few, well selected genes.*

Cox in All (MFS)

	Gene	p-Value
1	TOPO2A	0.0002
2	E2F	0.0006
3	SURVIVIN	0.0016
4	MMP1	0.0035
5	TS	0.0052
6	CATB	0.0053
7	CATD	0.0124
8	TK	0.0176
9	UPA	0.0195
10	5alphar1	0.0307
11	UPAR	0.0341
12	TIMP4	0.0393
13	LKB1	0.0420
14	KIP	0.0455
15	PAI1	0.0466
16	VEGFR1	0.0483
17	VEGFR2	0.0505
18	ERBB2	0.0513
19	BAX	0.0541
20	MMP11	0.0586

Cox in ErbB2+ (MFS)

	Gene	p-Value
1	UPA	0.0097
2	MMP11	0.0176
3	UPAR	0.0190
4	MMP1	0.0247
5	MMP3	0.0311
6	ERBB3	0.0552
7	TIMP3	0.0564
8	TP	0.0778
9	PAI1	0.0892
10	MMP2	0.0918
11	CATB	0.1096
12	DPD	0.1485
13	SFRP	0.1568
14	TIMP1	0.1577
15	TOPO2A	0.1716
16	AREG	0.1760
17	VEGFC	0.2244
18	VEGFA	0.2527
19	TCF4	0.2756
20	TGFB	0.2780

Cox in ErbB2- (MFS)

	Gene	p-Value
1	SURVIVIN	0.0006
2	E2F	0.0010
3	TOPO2A	0.0025
4	TK	0.0036
5	TS	0.0069
6	VEGFA	0.0185
7	KIP	0.0295
8	BAX	0.0302
9	5alphar1	0.0304
10	CATD	0.0338
11	CATB	0.0431
12	BCATENIN	0.0467
13	ADR	0.0564
14	LKB1	0.0660
15	VEGFR1	0.0690
16	VEGFR2	0.0779
17	TIMP4	0.0802
18	MYC	0.0978
19	MMP1	0.1005
20	CK7	0.1052

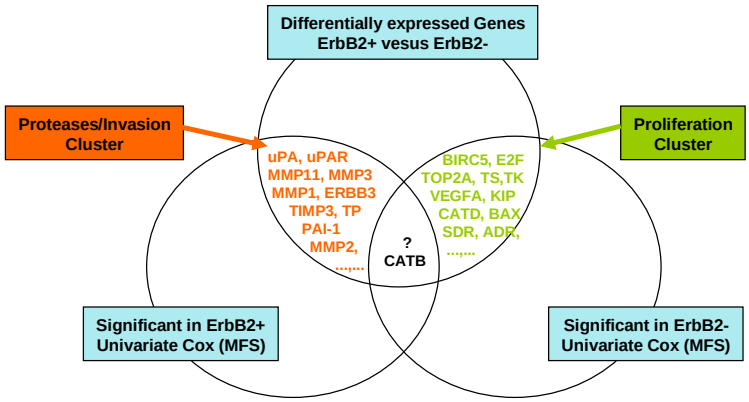


Figure 31. *Example of candidate genes prognostic in subgroups of ERBB2+ or ERBB2- tumors. Proliferation is the dominating cluster in ERBB2- tumors while proteases are the main prognostic indicators in ERBB2+ tumors.*

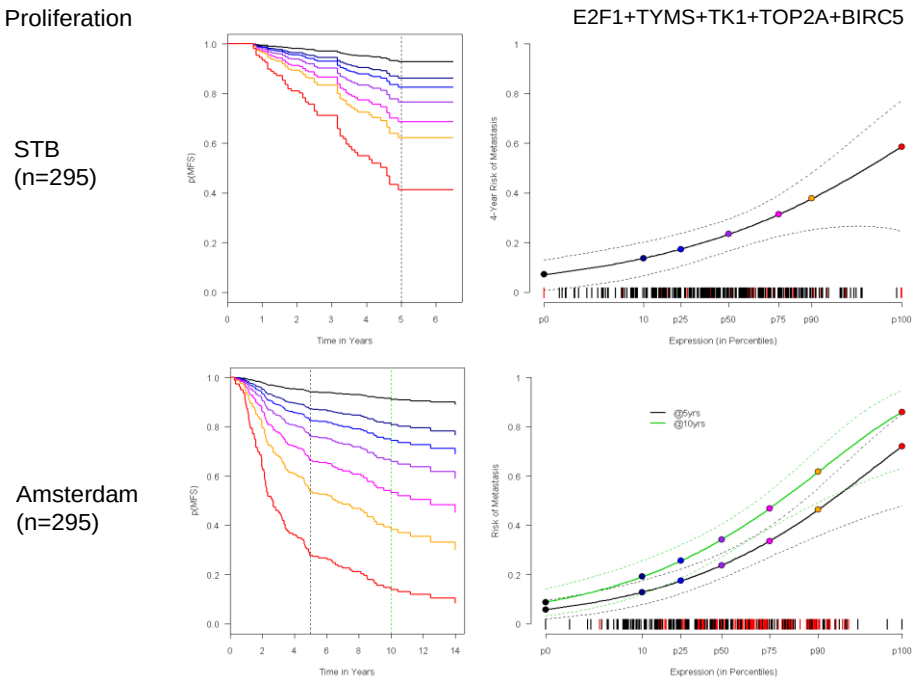


Figure 32. *Kaplan-Meier survival analysis. Based on a proliferation signature (average of 5 proliferation genes) survival is predicted based on a Cox regression model. Left: Survival curves for minimum value, 5th, 25th 50th, 75th, 90th percentile and maximum value. Right: Modeling of the risk of developing metastasis after 4-years (top), and 5 and 10 years (bottom) depending on the expression levels of these 5 proliferation genes ("risk score"). The same analysis was performed using two independent data sets (STB data set: 317 breast cancer samples; Amsterdam data set: 295 breast cancer samples)*

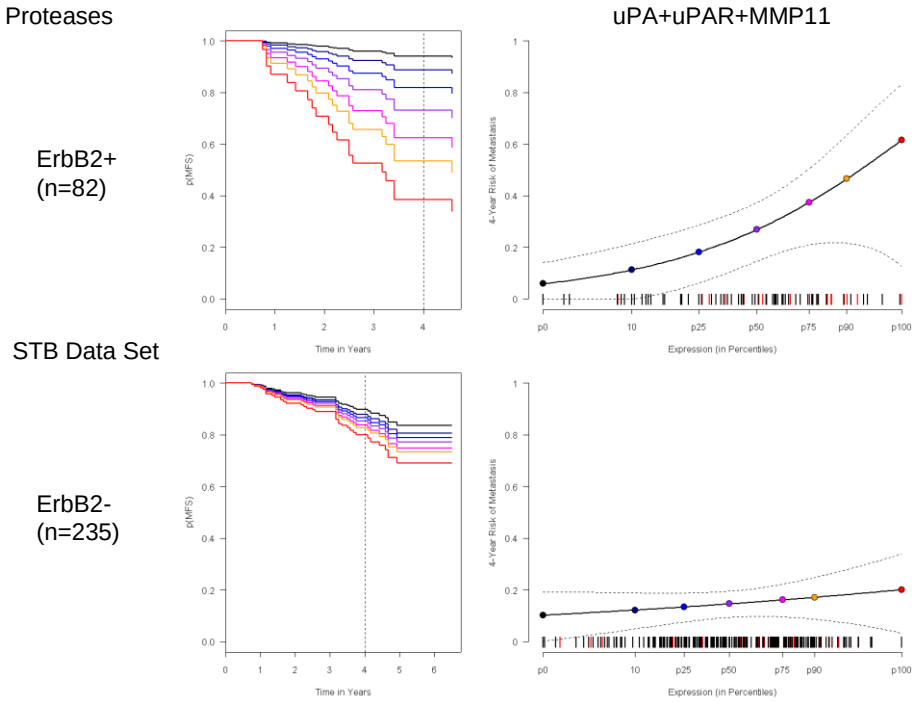


Figure 33. *Kaplan-Meier survival analysis. Based on a 3 proteases survival is predicted based on a Cox regression model. Left: Survival curves for minimum value, 5th, 25th 50th, 75th, 90th percentile and maximum value. Right: Modeling of the risk of developing metastasis after 4-years (top). Stratification according ERBB2 status (Top: ERBB2+ tumors; bottom: ERBB2- tumors). Interestingly, the prognostic value of proteases is much less prominent in ERBB2- tumors.*

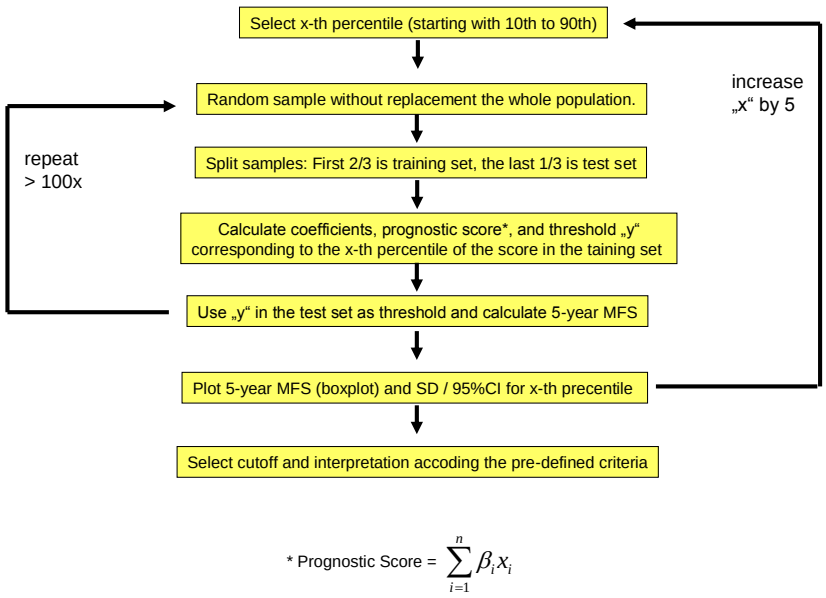


Figure 34. *Example of a prognostic score algorithm. Several genes are selected according to their association with survival in a training set. A prognostic score is calculated for each sample based on Cox coefficients and the actual expression value of the respective gene. The performance is evaluated in the training set for various cutoffs thereafter.*

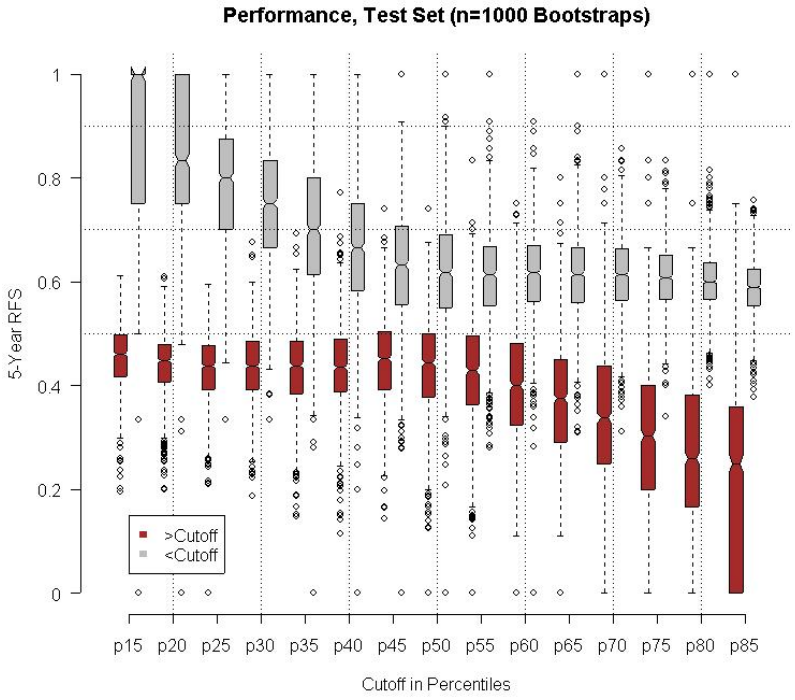


Figure 35. *Performance (5-year relapse-free survival) of the prognostic score in 317 primary breast cancer samples based on 5 genes. The score for a new sample could be calculated according to the formula and the coefficients derived from the training set and estimate the 5-year survival of that patient by comparing the score to the corresponding score in the test set.*

Outlook

Possible extensions and/or follow-up through collaboration partners were thought about: (i) laser capture micro dissection (LCM) to investigate role of stroma/epithelium, (ii) refine methods and protocols to develop diagnostic tools for the routine clinical setting, (iii) prospective clinical trial to assess predictive value and/or select patients for novel therapies (e.g. Herceptin ®), (iv) model system to explore/validate some of findings, in particular the relationship between ERBB2 overexpression and the mechanisms of invasion and metastasis through the regulation of proteases, (v) further explore potential of publicly available gene expression data sets.

Participants and Collaborations

(See also NIRA data base)

Patrick Urban, Vincent Vuaroqueaux, Martin Labuhn, Jonas Cicenas, Serenella Eppenberger-Castori, Urs Eppenberger, Willy Küng, Otto Hagenbuchle, Keith Harshman, Mauro Delorenzi, Darlene Goldstein, Therry Sengstag, Asa Wirapati

Collaborations

Main: NCCR Molecular Oncology, Lausanne
Stiftung Tumorbank Basel, University of Basel
Swiss Institute for Experimental Cancer Research (ISREC), Lausanne
DNA Array Facility Lausanne, CIG, University of Lausanne
Bioinformatics Core Facility, Lausanne
Swiss Institute for Bioinformatics, Lausanne
OncoScore AG, Basel
Molekulare Tumorbiologie, University Basel, Department for Gynecology and Oncology

Others: University Hospital, Department for Gynecology and Oncology, Basel
University Institute of Pathology, Basel
Buck Institute for Aging, Novato, CA, USA
ZeTuP, St. Gallen

References

1. American Cancer Society: Cancer Facts and Figures 2006. Atlanta, Ga: American Cancer Society, 2006

2. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 365:1687-717, 2005

3. Nystrom L, Andersson I, Bjurstam N, et al: Long-term effects of mammography screening: updated overview of the Swedish randomised trials. Lancet 359:909-19, 2002

4. Esteva FJ, Hortobagyi GN: Prognostic molecular markers in early breast cancer. Breast Cancer Res 6:109-18, 2004

5. Goldhirsch A, Glick JH, Gelber RD, et al: Meeting Highlights: International Expert Consensus on the Primary Therapy of Early Breast Cancer 2005. Ann Oncol, 2005

6. Sorlie T, Perou CM, Tibshirani R, et al: Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 98:10869-74, 2001

7. van 't Veer LJ, Dai H, van de Vijver MJ, et al: Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530-6, 2002

8. Wang Y, Klijn JG, Zhang Y, et al: Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 365:671-9, 2005

9. Paik S, Tang G, Shak S, et al: Gene Expression and Benefit of Chemotherapy in Women With Node-Negative, Estrogen Receptor-Positive Breast Cancer. J Clin Oncol, 2006

10. Paik S, Shak S, Tang G, et al: A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 351:2817-26, 2004

11. Schulze A, Downward J: Navigating gene expression using microarrays--a technology review. Nat Cell Biol 3:E190-5, 2001

12. Lockhart DJ, Winzeler EA: Genomics, gene expression and DNA arrays. Nature 405:827-36, 2000

13. Jain KK: Biochips for gene spotting. Science 294:621-3, 2001

14. Quackenbush J: Computational analysis of microarray data. Nat Rev Genet 2:418-27, 2001

15. Eisen MB, Spellman PT, Brown PO, et al: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 95:14863-8, 1998

16. Planet PJ, DeSalle R, Siddall M, et al: Systematic analysis of DNA microarray data: ordering and interpreting patterns of gene expression. Genome Res 11:1149-55, 2001

17. Khan J, Wei JS, Ringner M, et al: Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nat Med 7:673-9, 2001

18. Nadon R, Shoemaker J: Statistical issues with microarrays: processing and analysis. Trends Genet 18:265-71, 2002

19. Quackenbush J: Microarray analysis and tumor classification. N Engl J Med 354:2463-72, 2006

20. Eppenberger-Castori S, Kueng W, Benz C, et al: Prognostic and predictive significance of ErbB-2 breast tumor levels measured by enzyme immunoassay. J Clin Oncol 19:645-56, 2001

21. Thor AD, Liu S, Edgerton S, et al: Activation (tyrosine phosphorylation) of ErbB-2 (HER-2/neu): a study of incidence and correlation with outcome in breast cancer. J Clin Oncol 18:3230-9, 2000

22. Ross DT, Scherf U, Eisen MB, et al: Systematic variation in gene expression patterns in human cancer cell lines. Nat Genet 24:227-35, 2000

23. Perou CM, Jeffrey SS, van de Rijn M, et al: Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proc Natl Acad Sci U S A 96:9212-7, 1999

24. Perou CM, Sorlie T, Eisen MB, et al: Molecular portraits of human breast tumours. Nature 406:747-52, 2000

25. Sorlie T, Tibshirani R, Parker J, et al: Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A 100:8418-23, 2003

26. Hedenfalk I, Duggan D, Chen Y, et al: Gene-expression profiles in hereditary breast cancer. N Engl J Med 344:539-48, 2001

27. Zhao H, Langerod A, Ji Y, et al: Different gene expression patterns in invasive lobular and ductal carcinomas of the breast. Mol Biol Cell 15:2523-36, 2004

28. Bertucci F, Finetti P, Rougemont J, et al: Gene expression profiling for molecular characterization of inflammatory breast cancer and prediction of response to chemotherapy. Cancer Res 64:8558-65, 2004

29. Desai KV, Xiao N, Wang W, et al: Initiating oncogenic event determines gene-expression patterns of human breast cancer models. Proc Natl Acad Sci U S A 99:6967-72, 2002

30. Ma XJ, Salunga R, Tuggle JT, et al: Gene expression profiles of human breast cancer progression. Proc Natl Acad Sci U S A 100:5974-9, 2003

31. West M, Blanchette C, Dressman H, et al: Predicting the clinical status of human breast cancer by using gene expression profiles. Proc Natl Acad Sci U S A 98:11462-7, 2001

32. Weigelt B, Wessels LF, Bosma AJ, et al: No common denominator for breast cancer lymph node metastasis. Br J Cancer 93:924-32, 2005

33. Zhu G, Reynolds L, Crnogorac-Jurcevic T, et al: Combination of microdissection and microarray analysis to identify gene expression changes between differentially located tumour cells in breast cancer. Oncogene 22:3742-8, 2003

34. van de Vijver MJ, He YD, van't Veer LJ, et al: A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347:1999-2009, 2002

35. Huang E, Cheng SH, Dressman H, et al: Gene expression predictors of breast cancer outcomes. Lancet 361:1590-6, 2003

36. Rouzier R, Perou CM, Symmans WF, et al: Breast cancer molecular subtypes respond differently to preoperative chemotherapy. Clin Cancer Res 11:5678-85, 2005

37. Jansen MP, Foekens JA, van Staveren IL, et al: Molecular classification of tamoxifen-resistant breast carcinomas by gene expression profiling. J Clin Oncol 23:732-40, 2005

38. Chang JC, Wooten EC, Tsimelzon A, et al: Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. Lancet 362:362-9, 2003

39. Gianni L, Zambetti M, Clark K, et al: Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer. J Clin Oncol 23:7265-77, 2005

40. Weigelt B, Hu Z, He X, et al: Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer. Cancer Res 65:9155-8, 2005

41. Minn AJ, Gupta GP, Siegel PM, et al: Genes that mediate breast cancer metastasis to lung. Nature 436:518-24, 2005

42. Sotiriou C, Wirapati P, Loi S, et al: Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst 98:262-72, 2006

43. Hyman E, Kauraniemi P, Hautaniemi S, et al: Impact of DNA amplification on gene expression patterns in breast cancer. Cancer Res 62:6240-5, 2002

44. Chang HY, Nuyten DS, Sneddon JB, et al: Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. Proc Natl Acad Sci U S A 102:3738-43, 2005

45. West RB, Nuyten DS, Subramanian S, et al: Determination of stromal signatures in breast carcinoma. PLoS Biol 3:e187, 2005

46. Mackay A, Jones C, Dexter T, et al: cDNA microarray analysis of genes associated with ERBB2 (HER2/neu) overexpression in human mammary luminal epithelial cells. Oncogene 22:2680-8, 2003

47. Bertucci F, Borie N, Ginestier C, et al: Identification and validation of an ERBB2 gene expression signature in breast cancers. Oncogene 23:2564-75, 2004

48. Brenton JD, Carey LA, Ahmed AA, et al: Molecular classification and molecular forecasting of breast cancer: ready for clinical application? J Clin Oncol 23:7350-60, 2005

49. Simon R: Roadmap for developing and validating therapeutically relevant genomic classifiers. J Clin Oncol 23:7332-41, 2005

50. Weigelt B, Peterse JL, van 't Veer LJ: Breast cancer metastasis: markers and models. Nat Rev Cancer 5:591-602, 2005

51. Fan C, Oh DS, Wessels L, et al: Concordance among gene-expression-based predictors for breast cancer. N Engl J Med 355:560-9, 2006

52. Gaasterland T, Bekiranov S: Making the most of microarray data. Nat Genet 24:204-6, 2000

53. Eccles SA: The role of c-erbB-2/HER2/neu in breast cancer progression and metastasis. J Mammary Gland Biol Neoplasia 6:393-406, 2001

54. Olayioye MA, Neve RM, Lane HA, et al: The ErbB signaling network: receptor heterodimerization in development and cancer. Embo J 19:3159-67, 2000

55. Yarden Y, Sliwkowski MX: Untangling the ErbB signalling network. Nat Rev Mol Cell Biol 2:127-37, 2001

56. Yarden Y: Biology of HER2 and its importance in breast cancer. Oncology 61 Suppl 2:1-13, 2001

57. Hynes NE, Lane HA: ERBB receptors and cancer: the complexity of targeted inhibitors. Nat Rev Cancer 5:341-54, 2005

58. de Bono JS, Rowinsky EK: The ErbB receptor family: a therapeutic target for cancer. Trends Mol Med 8:S19-26, 2002

59. Nahta R, Esteva FJ: HER-2-targeted therapy: lessons learned and future directions. Clin Cancer Res 9:5078-84, 2003

60. Nahta R, Esteva FJ: HER2 therapy: Molecular mechanisms of trastuzumab resistance. Breast Cancer Res 8:215, 2006

61. Schulze WX, Deng L, Mann M: Phosphotyrosine interactome of the ErbB-receptor kinase family. Mol Syst Biol 1:2005 0008, 2005

62. Schlessinger J: Cell signaling by receptor tyrosine kinases. Cell 103:211-25, 2000

63. Jones RB, Gordus A, Krall JA, et al: A quantitative protein interaction network for the ErbB receptors using protein microarrays. Nature 439:168-74, 2006

64. Slamon DJ, Clark GM, Wong SG, et al: Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science 235:177-82, 1987

65. Troyer KL, Lee DC: Regulation of mouse mammary gland development and tumorigenesis by the ERBB signaling network. J Mammary Gland Biol Neoplasia 6:7-21, 2001

66. D'Souza B, Taylor-Papadimitriou J: Overexpression of ERBB2 in human mammary epithelial cells signals inhibition of transcription of the E-cadherin gene. Proc Natl Acad Sci U S A 91:7202-6, 1994

67. Marone R, Hess D, Dankort D, et al: Memo mediates ErbB2-driven cell motility. Nat Cell Biol 6:515-22, 2004

68. Bosc DG, Goueli BS, Janknecht R: HER2/Neu-mediated activation of the ETS transcription factor ER81 and its target gene MMP-1. Oncogene 20:6215-24, 2001

69. Mazumdar A, Adam L, Boyd D, et al: Heregulin regulation of urokinase plasminogen activator and its receptor: human breast epithelial cell invasion. Cancer Res 61:400-5, 2001

70. Gum R, Wang SW, Lengyel E, et al: Up-regulation of urokinase-type plasminogen activator expression by the HER2/neu proto-oncogene. Anticancer Res 15:1167-72, 1995

71. Chakraborti S, Mandal M, Das S, et al: Regulation of matrix metalloproteinases: an overview. Mol Cell Biochem 253:269-85, 2003

72. Tan M, Yao J, Yu D: Overexpression of the c-erbB-2 gene enhanced intrinsic metastasis potential in human breast cancer cells without increasing their transformation abilities. Cancer Res 57:1199-205, 1997

73. Adam L, Vadlamudi R, Kondapaka SB, et al: Heregulin regulates cytoskeletal reorganization and cell migration through the p21-activated kinase-1 via phosphatidylinositol-3 kinase. J Biol Chem 273:28238-46, 1998

74. Spencer KS, Graus-Porta D, Leng J, et al: ErbB2 is necessary for induction of carcinoma cell invasion by ErbB family receptor tyrosine kinases. J Cell Biol 148:385-97, 2000

75. Bergers G, Benjamin LE: Tumorigenesis and the angiogenic switch. Nat Rev Cancer 3:401-10, 2003

76. Kumar R, Yarmand-Bagheri R: The role of HER2 in angiogenesis. Semin Oncol 28:27-32, 2001

77. Xiong S, Grijalva R, Zhang L, et al: Up-regulation of vascular endothelial growth factor in breast cancer cells by the heregulin-beta1-activated p38 signaling pathway enhances endothelial cell migration. Cancer Res 61:1727-32, 2001

78. Laughner E, Taghavi P, Chiles K, et al: HER2 (neu) signaling increases the rate of hypoxia-inducible factor 1alpha (HIF-1alpha) synthesis: novel mechanism for HIF-1-mediated vascular endothelial growth factor expression. Mol Cell Biol 21:3995-4004, 2001

79. Klos KS, Wyszomierski SL, Sun M, et al: ErbB2 increases vascular endothelial growth factor protein synthesis via activation of mammalian target of rapamycin/p70S6K leading to increased angiogenesis and spontaneous metastasis of human breast cancer cells. Cancer Res 66:2028-37, 2006

80. Deryugina EI, Quigley JP: Matrix metalloproteinases and tumor metastasis. Cancer Metastasis Rev 25:9-34, 2006

81. Kao J, Pollack JR: RNA interference-based functional dissection of the 17q12 amplicon in breast cancer reveals contribution of coamplified genes. Genes Chromosomes Cancer 45:761-9, 2006

82. Kauraniemi P, Barlund M, Monni O, et al: New amplified and highly expressed genes discovered in the ERBB2 amplicon in breast cancer by cDNA microarrays. Cancer Res 61:8235-40, 2001

83. Jacobson KK, Morrison LE, Henderson BT, et al: Gene copy mapping of the ERBB2/TOP2A region in breast cancer. Genes Chromosomes Cancer 40:19-31, 2004

84. Leyland-Jones B: Trastuzumab: hopes and realities. Lancet Oncol 3:137-44, 2002

85. Vogel CL, Cobleigh MA, Tripathy D, et al: Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. J Clin Oncol 20:719-26, 2002

86. Slamon DJ, Leyland-Jones B, Shak S, et al: Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med 344:783-92, 2001

87. Nahta R, Hortobagyi GN, Esteva FJ: Growth factor receptors in breast cancer: potential for therapeutic intervention. Oncologist 8:5-17, 2003

88. Cobleigh MA, Vogel CL, Tripathy D, et al: Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. J Clin Oncol 17:2639-48, 1999

89. Herbst RS, Fukuoka M, Baselga J: Gefitinib--a novel targeted approach to treating cancer. Nat Rev Cancer 4:956-65, 2004

90. Nagata Y, Lan KH, Zhou X, et al: PTEN activation contributes to tumor inhibition by trastuzumab, and loss of PTEN predicts trastuzumab resistance in patients. Cancer Cell 6:117-27, 2004

91. Lee KF, Simon H, Chen H, et al: Requirement for neuregulin receptor erbB2 in neural and cardiac development. Nature 378:394-8, 1995

92. Crone SA, Zhao YY, Fan L, et al: ErbB2 is essential in the prevention of dilated cardiomyopathy. Nat Med 8:459-65, 2002

93. Cicenas J, Urban P, Kung W, et al: Phosphorylation of tyrosine 1248-ERBB2 measured by chemiluminescence-linked immunoassay is an independent predictor of poor prognosis in primary breast cancer patients. Eur J Cancer 42:636-45, 2006

94. Cicenas J, Urban P, Vuaroqueaux V, et al: Increased level of phosphorylated akt measured by chemiluminescence-linked immunosorbent assay is a predictor of poor prognosis in primary breast cancer overexpressing ErbB-2. Breast Cancer Res 7:R394-401, 2005

95. Labuhn M, Vuaroqueaux V, Fina F, et al: Simultaneous quantitative detection of relevant biomarkers in breast cancer by quantitative real-time PCR. Int J Biol Markers 21:30-9, 2006

96. Van Gelder RN, von Zastrow ME, Yool A, et al: Amplified RNA synthesized from limited quantities of heterogeneous cDNA. Proc Natl Acad Sci U S A 87:1663-7, 1990

97. Breimann L, Friedman JH, Olsen RA: Classification and Regression Trees. Wadsworth and Brooks Publishing, Monterey, CA., 1984

98. Zanetti-Dallenbach R, Vuaroqueaux V, Wight E, et al: Comparison of gene expression profiles in core biopsies and corresponding surgical breast cancer samples. Breast Cancer Res 8:R51, 2006

99. Vuaroqueaux V, Urban P, Labuhn M, et al: Low E2F1 transcript levels are a strong determinant of favorable breast cancer outcome. Breast Cancer Res 9:R33, 2007

100. Urban P, Vuaroqueaux V, Labuhn M, et al: Increased expression of urokinase-type plasminogen activator mRNA determines adverse prognosis in ErbB2-positive primary breast cancer. J Clin Oncol 24:4245-53, 2006

101. Hanahan D, Weinberg RA: The hallmarks of cancer. Cell 100:57-70, 2000

102. van Diest PJ, van der Wall E, Baak JP: Prognostic value of proliferation in invasive breast cancer: a review. J Clin Pathol 57:675-81, 2004

103. Pollack JR, Sorlie T, Perou CM, et al: Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. Proc Natl Acad Sci U S A 99:12963-8, 2002

104. Guerreiro Da Silva ID, Hu YF, Russo IH, et al: S100P calcium-binding protein overexpression is associated with immortalization of human breast epithelial cells in vitro and early stages of breast cancer development in vivo. Int J Oncol 16:231-40, 2000

105. Mazzucchelli L: Protein S100A4: too long overlooked by pathologists? Am J Pathol 160:7-13, 2002

106. White SL, Gharbi S, Bertani MF, et al: Cellular responses to ErbB-2 overexpression in human mammary luminal epithelial cells: comparison of mRNA and protein expression. Br J Cancer 90:173-81, 2004

107. Ahmed AA, Brenton JD: Microarrays and breast cancer clinical studies: forgetting what we have not yet learnt. Breast Cancer Res 7:96-9, 2005

108. Ein-Dor L, Kela I, Getz G, et al: Outcome signature genes in breast cancer: is there a unique set? Bioinformatics 21:171-8, 2005

109. Wang SC, Hung MC: HER2 overexpression and cancer targeting. Semin Oncol 28:115-24, 2001

110. Longley DB, Harkin DP, Johnston PG: 5-fluorouracil: mechanisms of action and clinical strategies. Nat Rev Cancer 3:330-8, 2003

111. Di Leo A, Cardoso F, Durbecq V, et al: Predictive molecular markers in the adjuvant therapy of breast cancer: state of the art in the year 2002. Int J Clin Oncol 7:245-53, 2002

112. Dai H, van't Veer L, Lamb J, et al: A cell proliferation signature is a marker of extremely poor outcome in a subpopulation of breast cancer patients. Cancer Res 65:4059-66, 2005

113. Rosty C, Sheffer M, Tsafrir D, et al: Identification of a proliferation gene cluster associated with HPV E6/E7 expression level and viral DNA load in invasive cervical carcinoma. Oncogene 24:7094-104, 2005

114. Allgayer H, Babic R, Gruetzner KU, et al: c-erbB-2 is of independent prognostic relevance in gastric cancer and is associated with the expression of tumor-associated protease systems. J Clin Oncol 18:2201-9, 2000

115. Berney CR, Yang J, Fisher RJ, et al: Correlates of urokinase-type plasminogen activator in colorectal cancer: positive relationship with nm23 and c-erbB-2 protein expression. Oncol Res 10:47-54, 1998

116. Umeda T, Eguchi Y, Okino K, et al: Cellular localization of urokinase-type plasminogen activator, its inhibitors, and their mRNAs in breast cancer tissues. J Pathol 183:388-97, 1997

117. Nielsen BS, Sehested M, Duun S, et al: Urokinase plasminogen activator is localized in stromal cells in ductal breast cancer. Lab Invest 81:1485-501, 2001

118. Baker EA, Leaper DJ: Proteinases, their inhibitors, and cytokine profiles in acute wound fluid. Wound Repair Regen 8:392-8, 2000

119. Stadelmann WK, Digenis AG, Tobin GR: Physiology and healing dynamics of chronic cutaneous wounds. Am J Surg 176:26S-38S, 1998

120. Coussens LM, Werb Z: Inflammation and cancer. Nature 420:860-7, 2002

121. Dvorak HF: Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing. N Engl J Med 315:1650-9, 1986

122. Balkwill F, Mantovani A: Inflammation and cancer: back to Virchow? Lancet 357:539-45, 2001

123. Harbeck N, Kates RE, Gauger K, et al: Urokinase-type plasminogen activator (uPA) and its inhibitor PAI-I: novel tumor-derived factors with a high prognostic and predictive impact in breast cancer. Thromb Haemost 91:450-6, 2004

124. Duffy MJ, Duggan C: The urokinase plasminogen activator system: a rich source of tumour markers for the individualised management of patients with cancer. Clin Biochem 37:541-8, 2004

125. Decock J, Paridaens R, Cufer T: Proteases and metastasis: clinical relevance nowadays? Curr Opin Oncol 17:545-50, 2005

126. Cianfrocca M, Goldstein LJ: Prognostic and predictive factors in early-stage breast cancer. Oncologist 9:606-16, 2004

127. Look M, van Putten W, Duffy M, et al: Pooled analysis of prognostic impact of uPA and PAI-1 in breast cancer patients. Thromb Haemost 90:538-48, 2003

128. Adam R, Drummond DR, Solic N, et al: Modulation of the receptor binding affinity of amphiregulin by modification of its carboxyl terminal tail. Biochim Biophys Acta 1266:83-90, 1995

129. Neelam B, Richter A, Chamberlin SG, et al: Structure-function studies of ligand-induced epidermal growth factor receptor dimerization. Biochemistry 37:4884-91, 1998

130. Wong L, Deb TB, Thompson SA, et al: A differential requirement for the COOH-terminal region of the epidermal growth factor (EGF) receptor in amphiregulin and EGF mitogenic signaling. J Biol Chem 274:8900-9, 1999

131. Beerli RR, Hynes NE: Epidermal growth factor-related peptides activate distinct subsets of ErbB receptors and differ in their biological activities. J Biol Chem 271:6071-6, 1996

132. Salomon DS, Normanno N, Ciardiello F, et al: The role of amphiregulin in breast cancer. Breast Cancer Res Treat 33:103-14, 1995

133. Normanno N, Ciardiello F: EGF-related peptides in the pathophysiology of the mammary gland. J Mammary Gland Biol Neoplasia 2:143-51, 1997

134. Luetteke NC, Qiu TH, Fenton SE, et al: Targeted inactivation of the EGF and amphiregulin genes reveals distinct roles for EGF receptor ligands in mouse mammary gland development. Development 126:2739-50, 1999

Publications

Vuaroqueaux V*, Urban P* et al. *Low E2F1 transcript levels are a strong determinant of favorable breast cancer outcome*. Breast Cancer Res. 2007;9(3):R33.

Urban P et al. *Increased expression of urokinase-type plasminogen activator mRNA determines adverse prognosis in ErbB2-positive primary breast cancer*. J Clin Oncol. 2006 Sep 10;24(26):4245-53.

Zanetti-Dällenbach R, Vuaroqueaux V, Wight E, Labuhn M, Singer G, Urban P et al. *Comparison of gene expression profiles in core biopsies and corresponding surgical breast cancer samples*. Breast Cancer Res. 2006;8(4):R51.

Cicenas J*, Urban P* et al. *Phosphorylation of tyrosine 1248-ERBB2 measured by chemiluminescence-linked immunoassay is an independent predictor of poor prognosis in primary breast cancer patients.* Eur J Cancer. 2006 Mar;42(5):636-45.

Cicenas J, Urban P et al. *Increased level of phosphorylated akt measured by chemiluminescence-linked immunosorbent assay is a predictor of poor prognosis in primary breast cancer overexpressing ErbB-2*. Breast Cancer Res. 2005;7(4):R394-401.

* equal contribution

Abstracts and Meetings

Prediction of Distant Recurrence in Primary Breast Cancer Patients Stratified by ER and ERBB2 status. Urban P., et al., 2006 ASCO Annual Meeting, 2006.

Identification of Breast Cancer Patients at minimal Risk with quantitative rt-PCR. Vuaroqueaux V, Labuhn M, Urban P et al. AACR, 96th Annual Meeting, Anaheim, CA, 2005.

Molecular profiles of prostate cancer and its surrounding non-malignant tissue. V. Vuaroqueaux, M. Labuhn, R. Morant, P. A. Diener, C. Horica, T. Németh, M. Sulmoni, P. Urban et al, 2005 ASCO Annual Meeting, 2005.

Early Detection of Breast Cancer in Core Biopsies by Means of Quantitative RT-PCR. Vuaroqueaux V, Zanetti R, Labuhn M, Urban P et al. 9th International Conference on Primary Therapy of Early Breast Cancer, St. Gallen, CH, 2005.

Appendix

Probe Preparation (Labeling), Hybridization and Washing Protocols

What you need to prepare the probe:

- Oligo dT (Gibco=Life Tech. 25µg no 18418012)
- Reverse transcriptase + 5x Superscript buffer (Gibco no 18064014)
- 0.1 M DTT (comes with RT)
- Cy 3-dCTP (Amersham no PA53021) and Cy 5-dCTP (PA55021)
- DNTPs (BIL-Boehringer no 1969064)
- RNAse inhibitor (Gibco 1000 U no 15518012)
- Cot 1 human DNA (Gibco no 18440016)
- Poly A (Sigma 25mg no P9403)
- 25mM EDTA
- 1 M NaOH
- 1 M HCl
- 1 M Tris pH 6,8
- 3 M Na Ac pH 4.9
- 20x SSC and 3x SSC
- yeast tRNA (Sigma 500 U no R8759)
- 0.45 millipore filter (Millipore PVDF membrane o, 45 µm red no UFC3OHVNB)
- 2% SDS

What you need for hybridization:

- 3x SSC
- Probe
- 64°C waterbath

What you need for washes:

- 2x SSC 0.1% SDS
- 0.2x SSC
- 0.1x SSC

Probe preparation.

Prepare in two 200 µl PCR tubes:

- 2 µg oligo-dT (21 mer)
- 2 µg mRNA (control (CTL) or target (TR) RNA) or 25-50 µg totalRNA
- in 13.4 µl H2O

Heat 5 min, 70°C (PCR machine)
Leave 5 min at RT.

To each tube add:

- 6 µl 5X SuperScript II buffer
- 3 µl 0.1 M DTT
- 3 µl Cy3-dCTP (1 mM) or Cy5-dCTP (1 mM) (Amersham Cat. # PA53021, PA55021)
- 0.6 µl dNTPs (25 mM dATP, dTTP, dGTP; 10 mM dCTP)
- 2 µl SuperScript II Reverse Transcriptase (Gibco BRL Cat. # 18064-14)
- 2 µl RNAse Inhibitor (15 U/µl) (Gibco BRL Cat. # 15518-012)
- Incubate 1.30 hr, 42°C (PCR machine).

Pool the two tubes (CTL + TR)

Add:

- 2.65 µl 25 mM EDTA
- 3.3 µl 1 M NaOH

Incubate 10 min, 65°C (PCR machine).

Add:

- 3.3 µl of 1 M HCl
- 5 µl of 1 M Tris pH 6.8

- 20 µg Cot 1 DNA
- 10 µg Poly A RNA

Place the sample in a Microcon 30 microcencentrator (Amicon)

Add:
- 400 µl TE [10 mM Tris (pH 7.5)
- 1 mM EDTA]

Spin[1] 12 min, 14'000 g, to a volume of 10 to 20 µl. Discard flowthrough.

Repeat twice[2]
Concentrate down to 9 µ□l
Collect the probe by inverting the filter and centrifuging for 1 min at 1000 g

Add:
- 1.9 µl 20X SSC
- 1 µl E.Coli tRNA (10 µg/µl)
- 0.5 µl 10% SDS

Hybridization.

- Put cDNA array slide in hybridization chamber (TeleChem International Cat. # AHC-1*).
- Add 10 µl of 3X SSC in the two grooves at both ends of the slide to humidify chamber.
- Heat probe 1 min, 95°C.
- Spin 1 min, max speed.
- Add probe to slide[3], lay cover slip slowly on top of solution[4].
- Close chamber, immerge into 64°C bath. WORK QUICKLY.
- Hybridize O/N without agitation.

Washes.

Dismount chamber, put slide in glass slide holder with coverslip[5].

Place slide holder in glass dish containing washing solutions (ca 500 ml), move slide holder up and down during washing time for good agitation then transfer to next dish. Alternatively use Microarray Wash Station (TeleChem International Cat. # AW-1*).

Washes (RT):
- 5 min 2X SSC, 0.1% SDS
- 5 min 2X SSC, 0.1% SDS
- 1 min 0.2X SSC
- 1 min 0.2X SSC
- 1 min 0.1X SSC
- 1 min 0.1X SSC

Dry slide in centrifuge 2 min, 2650 rpm.
Store in light-tight box until scanning[6].

1. All washing/centrifugation steps are done at room temperature. Chilling results in precipitation of free label. Avoid putting the probe on ice at any time.
2. After the third centrifugation, probe retained by the Microcon should be brighter than flowthrough. This is an indication of good labeling.
3. Apply probe where genes are located (use provided mask).
4. Cover slips must be dust- and particle-free to allow even seating on the array. Air bubbles trapped under the cover slip exit after several minutes at 64°C.
5. Cover slip comes off during washes.
6. Cy3 or Cy5 are scanned dry. Storage of up to 2 weeks (Dark, RT) is OK.

*(http://www.wenet.net/~telechem/)

TRIZOL-RNAeasy (RNA Isolation and Purification) Protocol

Wear gloves, eye protection and work in a fume hood when using Trizol !

Isolation of RNA from small quantities of tissue (1-10 ug) or cells (10^2 to 10^4)

Avoid washing cells as this might increase the possibility of mRNA degradation.

1 -Add 800 ul of Trizol to tissue or cells
2 -Shear genomic DNA with 2 passes through a 26 gauge need (brown ones)

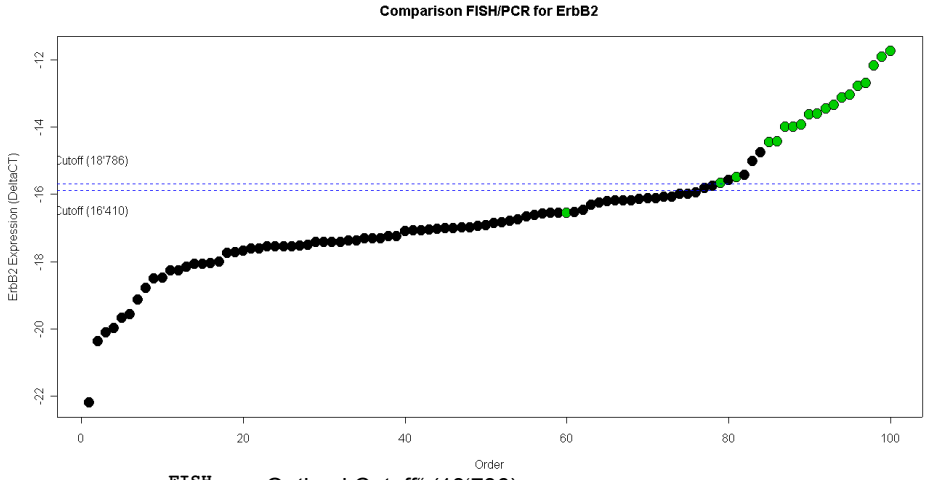
At this step, samples can be stored at -70^{o}C for at least a month.

3 -Incubate homogenized samples for 10' at RT (15 to 30^{o}C)
4 -Add 160 ul of PheChCl3isoamylOH (200ul per 1 ml of Trizol)
5 -Shake tubes vigorously by hand for 15 sec - Incubate at RT 2-3'
6 -Spin 15' at 4^{o}C and no more than 12 000 x g
7 -Aq phase + 500 ml PheChCl3isoamylOH
8 -Spin 2'
9 -Repeat 7 and 8
10 -Aqueous phase (500 ul) + an equal vol of 70% EtOH – mix well
11 -Load on RNAeasy column (up to 700 ul can be loaded)
12 -Spin 15 sec at 10 000 rpm
13 -Wash with 700 ul RW1 – Spin 15 sec at 10 000 rpm
14 -On a new collection tube, add 500 ul RPE – Spin as in 11
15 -Add 500 ul RPE – Spin 2' at 10 000 rpm
16 -In a fresh tube, elute samples with 30-50 ul RNAse free H_2O by spinning 1' at 10 000 rpm

Selected Cancer-related Genes assessed by qrt-PCR

Gene/Group	Gene Symbol	RefSequence
Proteases and Inhibitors		
Cathepsin B	CTSB	NM_001908
Cathepsin D	CTSD	NM_001909
Urokinase-type plasminogen activator	PLAU	NM_002658
Urokinase-type plasminogen activator receptor	PLAUR	NM_002659
Plasminogen Activator Inhibitor	SERPINE1	NM_000602
Metalloprotease 1	MMP1	NM_002421
Metalloprotease 2	MMP2	NM_004530
Metalloprotease 3	MMP3	NM_002422
Metalloprotease 7	MMP7	NM_002423
Metalloprotease 9	MMP9	NM_004994
Metalloprotease 11	MMP11	NM_005940
Tissue inhibitor of Metalloproteases 1	TIMP1	NM_003254
Tissue inhibitor of Metalloproteases 2	TIMP2	NM_003255
Tissue inhibitor of Metalloproteases 3	TIMP3	NM_000362
Tissue inhibitor of Metalloproteases 4	TIMP4	NM_003256
Proliferation		
Thymidilate synthase	TYMS	NM_001071
Thymidine kinase 1	TK1	NM_003258
Topoisomerase II-2alpha	TOP2A	NM_001067
Transcription factor E2F1	E2F1	NM_005225
Survivin	BIRC5	NM_001168
Dihydropyrimidine deshydrogenase	DPYD	NM_000110
Growth Factors and Receptors		
Transforming growth factor alpha	TGFA	NM_003236
Transforming growth factor beta 1	TGFB1	NM_000660
Epidermal growth factor	EGF	NM_001963
Amphiregulin	AREG	NM_001657
Epidermal growth factor receptor 1	EGFR	NM_005228
Epidermal growth factor receptor 2	ERBB2	NM_004448
Epidermal growth factor receptor 3	ERBB3	NM_001982
Epidermal growth factor receptor 4	ERBB4	NM_005235
Insulin-like growth factor 1	IGF1	NM_000618
Insulin-like growth factor 2	IGF2	NM_000612
Insulin-like growth factor receptor 1	IGF1R	NM_000875
Insulin-like growth factor receptor 2	IGF2R	NM_000876
Angiogenesis		
Vascular endothelial growth factor A	VEGF	NM_003376
Vascular endothelial growth factor B	VEGFB	NM_003377
Vascular endothelial growth factor C	VEGFC	NM_005429
Vascular endothelial growth factor D	FIGF	NM_004469
Vascular endothelial growth factor receptor 1	FLT1	NM_002019
Vascular endothelial growth factor receptor 2	KDR	NM_002253
Vascular endothelial growth factor receptor 3	FLT4	NM_002020
Thymidine Phosphorylase	ECGF1	NM_001953
Endocrine System		
Estrogen receptor 1	ESR1	NM_000125
Progesterone receptor	PGR	NM_000926
17 beta-hydroxy steroid deshydrogenase 1	HSD17B1	NM_000413
Cyclooxygenase-2	PTGS2	NM_000963
Peptidylglycine alpha-amidating monooxygenase	PAM	NM_000919
Epithelial Markers		
Cytokeratin 19	KRT19	NM_002276
CytoKeratin 7	KRT7	NM_005556
Glutathione S-transferase pi	GSTP1	NM_000852
Cell cycle and Apoptosis		
CyclinD1	CCND1	NM_053056
p21/Cip1	CDKN1A	NM_078467
p27/Kip1	CDKN1B	NM_004064
Retinoblastoma	RB1	NM_000321
Adrenomedullin	ADM	NM_001124
Bcl2-associated X protein	BAX	NM_138761
B-cell lymphoma 2	BCL2	NM_000633
Transcription Factors		
Hypoxia inducible factor 1-alpha	HIF1A	NM_001530
Transcription Factor 4	TCF4	NM_003199
beta-Catenin	CTNNB1	NM_001904
Amplified in breast cancer1 (AIB1)	NCOA3	NM_006534

ERBB2 Status Determination based on mRNA Expression Levels



	FISH	
	0	1
PCR		
FALSE	77	1
TRUE	4	18

„Optimal Cutoff“ (18‘786) X-squared = 67.1819, df = 1, p-value = 2.476e-16

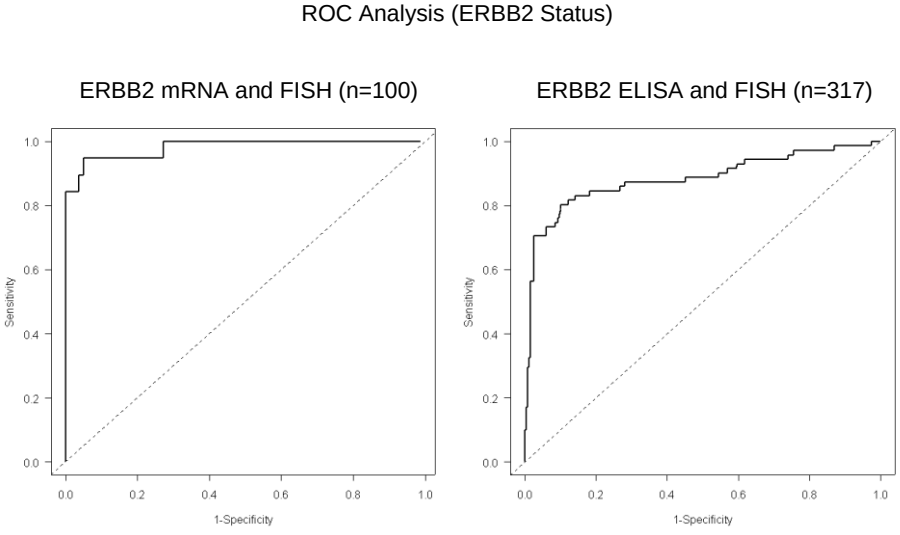
Sensitivity: 94.7%, Specificity: 95.1%, Accuracy: 95% (MissClassification: 5%)

	FISH	
	0	1
FALSE	75	1
TRUE	6	18

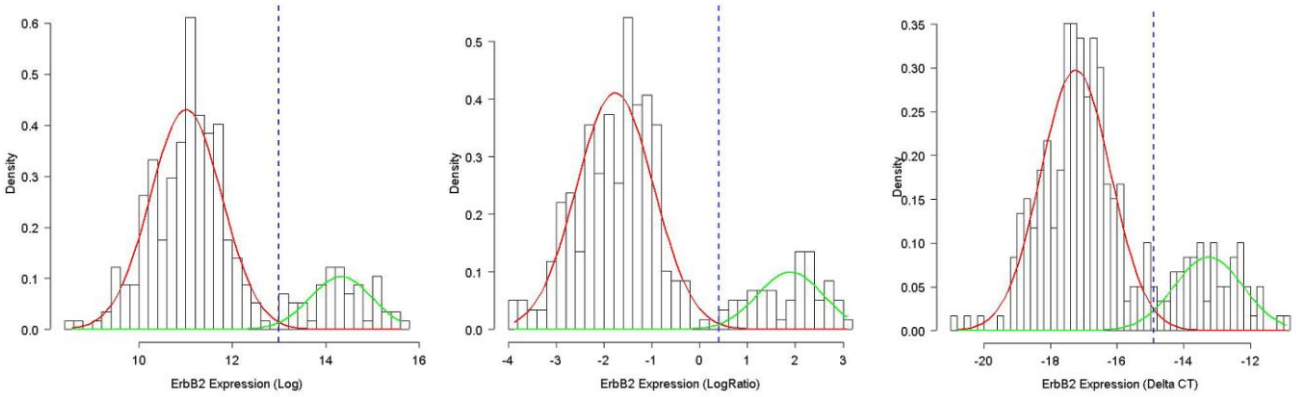
„OS Cutoff“ (16‘410) X-squared = 59.6493, df = 1, p-value = 1.134e-14

Sensitivity: 94.7%, Specificity: 92.6%, Accuracy: 93% (MissClassification: 7%)

Comparison of ERBB2 status measured by FISH (fluorescence in situ hybridization) and mRNA expression level detected by qrt-PCR. The graph shows the ranked expression values of ERBB2 mRNA levels in 100 samples and the corresponding ERBB2 status by FISH coded in color (black=ERBB2-, green=ERBB2+). Depending on the selection of the cutoff the misclassification rate was between 5 and 7%.

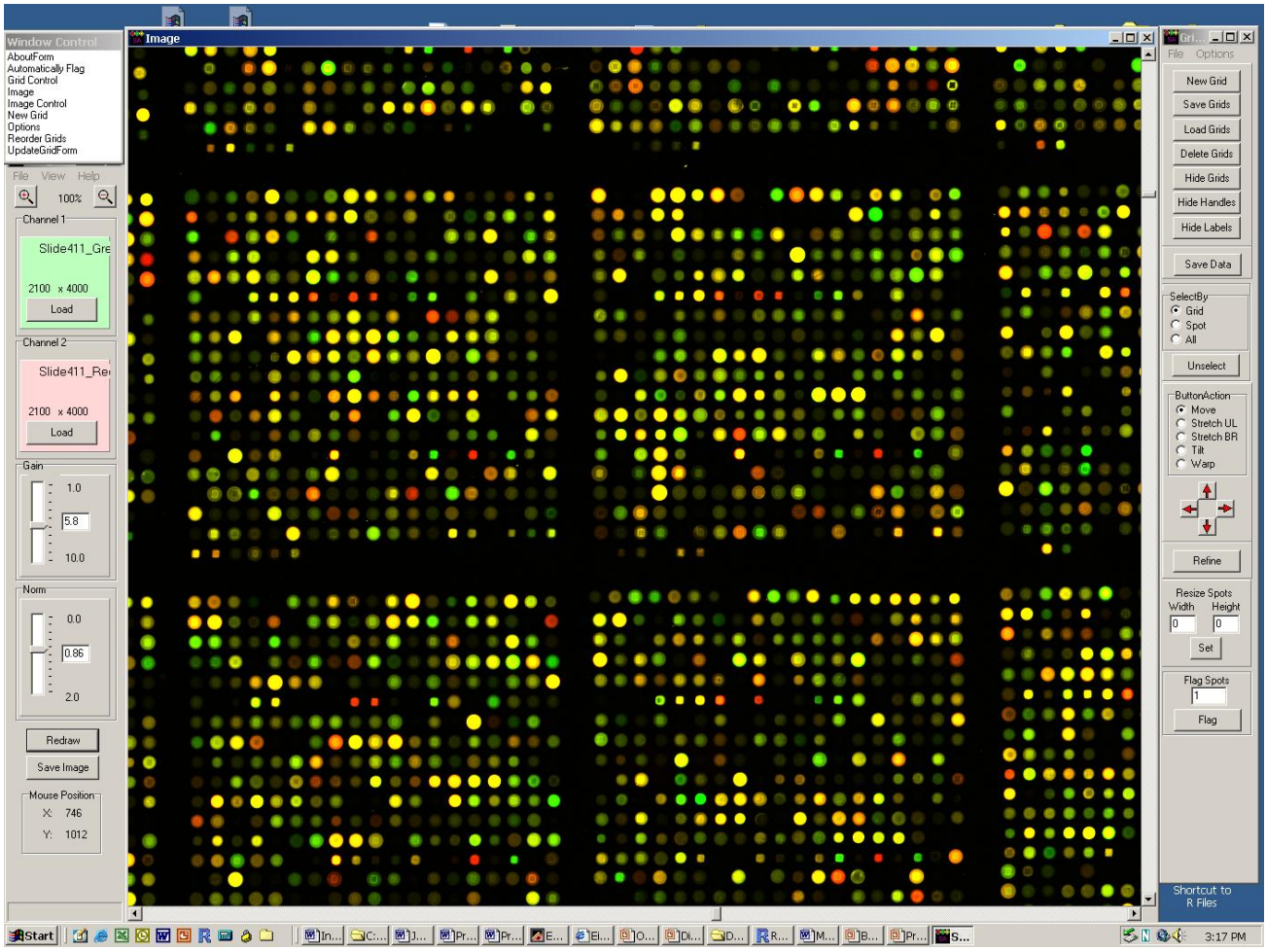


ROC (receiver operator characteristics): Performance of ERBB2 mRNA (qrt-PCR) and ERBB2 protein (ELISA) to predict ERBB2 status. ERBB2 status determined by FISH was used as gold standard (comparator).



Examples of the distribution of ERBB2 mRNA expression levels in three independent breast cancer data sets detected by qrt-PCR (left) and two different microarray platforms (middle and left). A mixture model was used to estimate the cutoff for ERBB2 status since all three displayed a clear bi-model distribution of the ERBB2 mRNA expression levels. The cutoff determined by the mixture model matched very well with the ERBB2 status detected by FISH.

Example Human 10K Microarray image



Screen shot of human 10K array (ScanAlyze software) after hybridization of a human breast cancer biopsy vs. reference RNA and overlay of the two pictures. This image also illustrates significant differences in spot size and irregularities of spot/sub grids alignment as a consequence of array printing.

Early Development: Summary Results and Conclusions

Early results obtained with RNA from human breast cancer biopsies as well as various cell lines showed that:

- RNeasy (Qiagen) was the preferred method for RNA extraction for the main project. Bioanalyzer proved to and important tool to assure and assess RNA quality, degradation and quantity after extraction, and was part of standard quality controls in future experiments.
- Amplification of RNA proved to be a robust and highly reproducible procedure yielding typically 2000 to 3000 fold amplification. It was to be preformed for both, biopsy and reference RNA in the future project.
- Refined probe preparation with additional Microcon filtering step protocols were yielded homogenous hybridizations with low background. Hybridization introduced bigger variation than amplification.
- RNA from test biopsies banked for many years had acceptable to good quality RNA for both, cryo-sections and powederized tissue. Review of histology (cryo-cuts) demonstrated almost complete infiltration with tumor with little or no necrosis. RNA derived from tumor tissue was successfully amplified and hybridized to microarrays.
- Preliminary analysis of the test biopsy data showed that tumors clustered into groups according to distinct biological features (e.g. sub-types according the hormone receptor status as previously described by Sorlie et al. [6]) with clear trends with respect to survival. Findings were in accordance with previously reported results in the literature suggesting that the data obtained with the cDNA microarrays, protocols and STB tissue were meaningful and of high quality.
- Clustering of various cell lines and test experiments revealed characteristic gene expression pattern reflecting the tissue of origin or the specific features accompanied with nature of the experimental design (Figures A and B). For example, repeat hybridizations always clustered together as well as did dye-swap. The main cluster separated biopsies (mixed tissues samples) from cells lines and "self-self" hybridization. The genes driving the individual cluster were consistent with the composition of the samples (eg. biopsy composed of various cell types vs. cell line of epithelial origin) as well as the organ specific characteristics.
- Many discussions and experiments were performed together with the SIB/BCF allowing to optimize and establish standard procedures for scanning, image analysis, normalization, background handling, and quality controls. Many of the steps were automated and implemented in a web-based application including a repository for raw (images) and numerical data.
- There was high correlation between gene expression levels measured by microarray and protein expression levels measured by ELISA for selected breast cancer related genes (e.g. estrogen receptor)
- RNA was extracted from 100 breast cancer biopsies for the main project was of high quality; adequate quantities for subsequent microarray analysis in 93 cases. Out of these, 92 RNA samples were successfully amplified
- After careful evaluation, the use of the Agilent scanner and the GenePix software for feature extraction and image analysis was recommended
- An extended pilot experiment demonstrated that hybridizations protocols work fine and reproducibility was good. However, randomization should be preformed in order to control for inter-batch and between-days variability. Moreover, hybridizations should be performed in duplicates without dye-swap (see below)
- Significant improvement were achieved with the "second" generation of slides/protocols, such as introduction of new slide surface, spotting solution and Arabidopsis spike controls
- New bioinformatics procedures improved quality assessment of slides and annotation files
- 89 biopsies were successfully analyzed by qrt-PCR and profiled for 60 well-defined cancer-related genes. Generally, there was high correlation between mRNA expression levels measured by microarray and qrt-PCR
- p-ERBB2 was assessed in 70 samples and correlated positively with survival, ERBB1/2 RNA and protein levels, and correlated negatively with ERBB4 RNA levels.
- p-AKT was measured in 65 samples and correlated with survival as well as proliferation genes

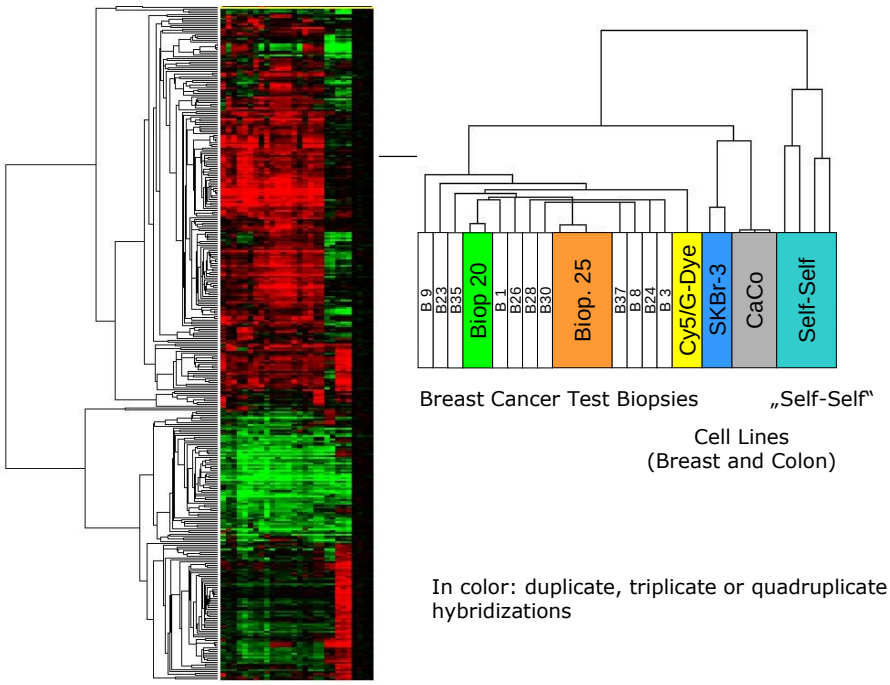


Figure A. *Example of hierarchical clustering using breast cancer biopsies, breast and colon cancer cell lines, and "self-self" hybridizations (same RNA used in target and control sample). Several samples were hybridized in replicates (underlined in color), and clustered always together. All breast cancer biopsies cluster together as well as the two cell lines, although much bigger differences can be observed. A filter was applied based on standard deviation (SD) with approx. 800 genes passing the criteria (shown). Red and green color indicates high and low expression as compared to reference (Human Reference RNA, Stratagene). Black: no difference between sample and reference (e.g. "self-self" hybridization)*

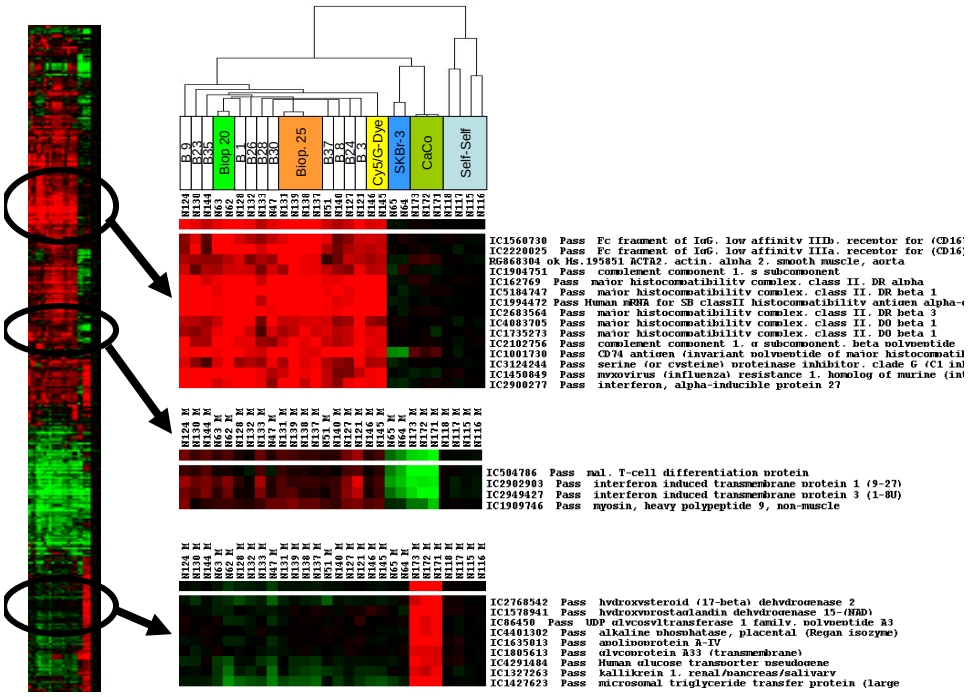


Figure B. *Hierarchical Clustering using breast cancer biopsies, breast and colon cancer cell lines: Genes related to immune response, inflammation and stroma were expressed in tissues only. In contrast, apolipoprotein and glucose transporters or trigryceride transfer protein were found in colon cancer cell line only (CaCo cell line), representing some of the tissue and/or cell line specific characteristics of the samples.*

Microarray Pilot Experiment

In preparation for the large experiment and as a guide to its design, a pilot experiment of 28 hybridizations was set up in close collaboration with the Bioinformatics Core Facility (BCF) to assess relative importance of different sources of variability because the whole experiment/hybridization would take several weeks and array printing is done in batches of 100 slides. The following parameters (sources of variance) were investigated:

- experimenter (are there differences between and within experimenters)
- batch differences (are there differences between batches of slides that have been printed at different times)
- effect of position within batch (are the differences between the slides printed at the beginning versus in the middle versus at the end of each batch)
- time difference between hybridizations (how does the time between replicate hybridizations affect the result)
- dye bias/normalization (is there a dye-bias which requires day-swap and how does normalization interfere)

The findings from the pilot experiment showed that overall slide quality is good. There were some differences between experimenters, and that there was variability within experimenters. Differences within/between batch did not seem to be very large whereas there were fairly large differences between days (time point of hybridization). These findings had the following impact on the design of the main experiment:

- duplicate hybridizations to be performed
- replicate experiments to be carried out on different days (randomization for time)
- position within batch not necessary to be controlled
- randomization necessary to control for inter-batch variability
- dye-swap not necessary; samples to be labeled with Cy3, reference (pooled RNA) with Cy5

Amphiregulin Microarray Project

Patrick Urban, Laura Ciarloni, Cathrin Brisken

Introduction

Epidermal growth factor (EGF), TGFa, HB-EGF and Amphiregulin (AR) are ligands that bind to the epithelial growth factor receptor (EGFR). AR has homology with EGF but is truncated at the C-terminus and exhibits a lower affinity to EGFR than EGF [128]. In a structure-function study of ligand-induced EGFR dimerization, AR could not dimerize EGFR (mechanism of receptor-activation) whereas EGF could [129]. Cell line experiments with truncated and wild-type EGFR demonstrated that the cytoplamic tail of EGFR plays a critical role in AR signaling but is dispensable for EGF [130] suggesting, that the two EGF-related peptides might differ in their biological activities. Furthermore, they can activate different subsets of ErbB receptors [131]. Besides, AR and EGF can function as autocrine growth factors (e.g. AR enhanced by EGF) [132,133].
Comparison of knockout mice models lacking functional EGF, TGFa, HB-EGF or AR demonstrated a specific and unique requirement for AR during morphogenesis in mammary gland development whereas EGF, TGFa, HB-EGF were clearly dispensable for this process [65,134]. Together, this suggests, that AR may have unique functions. We want to use gene expression profiling by microarrays to discover genes that are specifically regulated by AR and not other family members, as exemplified by EGF.

Hypothesis and Aims

AR is the only EGF family member that is essential to mammary gland development. It was hypothesized that this is due to the unique capacity of AR to induce a particular subset of target genes. To identify these candidate target genes differentially expressed genes were analyzed in cells treated with either AR or EGF by microarray profiling.

Experimental Procedures

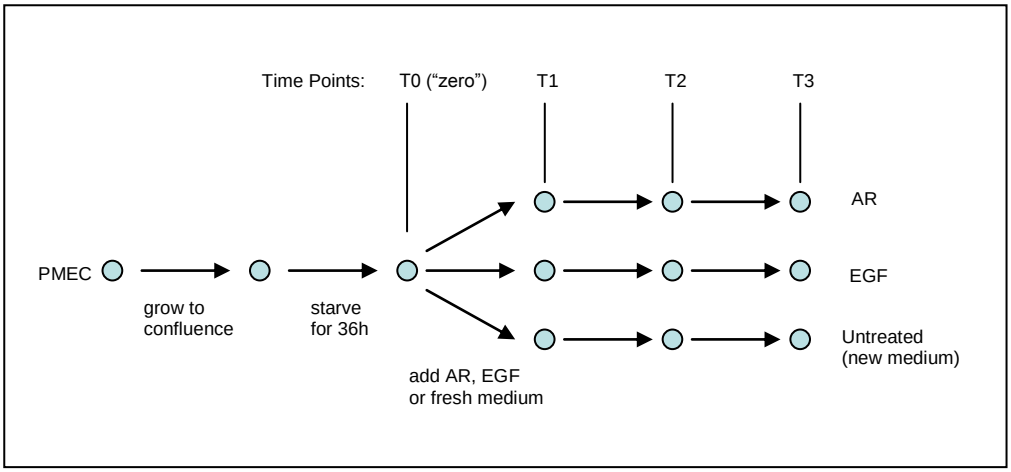
As model system primary mammary epithelial cells were used which were established in culture from mammary reduction mammoplasty (obtained from surgery at CHUV). After starvation for 36 hrs and cells were subsequently stimulated with AR or EGF. Prior to that, WT3 cells (mouse mammary cell line) were cultured and grown to confluence, serum starved for 36h and afterwards stimulated with different concentrations of AR or EGF. 8 hours later, cyclin D1 protein levels were used as a surrogate marker for response allowing optimizing individual concentrations for each ligand. Untreated cells were also sampled. Once the doses for each ligand were determined, the following experimental design was applied (using human primary mammary epithelial cells PMEC):

1. Part 1: One time point, 3 concentrations:

    - 3 different concentrations of each ligand (AR_{C1-3}, EGF_{C1-3}) and untreated (4h after stimulation), plus "zero" = 8 samples
    - each sample was hybridized against a pool of RNA consisting of all 8 samples
    - each sample was hybridized once, AR_{C3}, EGF_{C3} and untreated in duplicate with dye-swap = 11 slides

2. Part 2: One concentration, 3 time points:

    - "optimal" concentration (according to exp.1) was used in a time-course experiment
    - the exact design was discussed upon results from exp.1. (maybe direct comparisons could be done)
    - proposed time points were: T1= 0.5h (1h), T2= 4h and T3= 12h (8h)
    - slides required for the 2nd part: appr. 20



A big batch of pooled aRNA was prepared. Non-amplified and amplified RNA were checked for quality and quantity. DAF/BCF printed about 50 slides for this project.

Amphiregulin Project

Chip-ID	Chip-Nr	Colour	RNA	for 5ug	H2O	Tube	Comment
54	1	Cy3	Pool	7.0 ul	4.5 ul	1	
		Cy5	EGF 0.5	8.0 ul	3.5 ul	2	repeat
55	2	Cy3	Pool	7.0 ul	4.5 ul	3	
		Cy5	AR	7.3 ul	4.2 ul	4	
56	3	Cy3	Pool	7.0 ul	4.5 ul	5	
		Cy5	Control	5.7 ul	5.8 ul	6	
57	4	Cy3	EGF 0.5	8.0 ul	3.5 ul	7	
		Cy5	Pool	7.0 ul	4.5 ul	8	
58	5	Cy3	AR	7.3 ul	4.2 ul	9	
		Cy5	Pool	7.0 ul	4.5 ul	10	
59	6	Cy3	Control	5.7 ul	5.8 ul	11	
		Cy5	Pool	7.0 ul	4.5 ul	12	repeat
60	7	Cy3	Pool	7.0 ul	4.5 ul	13	
		Cy5	EGF 5			14	
61	8	Cy3	Pool	7.0 ul	4.5 ul	15	
		Cy5	EGF 0.1			16	

Pool:	AR 1/3	Conc AR: 680ng/ul	30ng	44ul
	EGF (0.5) 1/3	Conc EGF: 620ng/ul	30ng	48ul
	Control 1/3	Conc Contr: 876ng/ul	30ng	34ul
Total			90ng	126ul
Concentration Pool: **714ng/ul**		Lot Nr	Cy3: 212947	
Spikes: DilOct3			Cy5: 213053	
			SS: 1163838	

Results of differentially expressed genes are shown on Table A on next page.

Table A. *Array Results (Top 30 Differentially Expressed Genes; red over-, blue under-expressed)*

A 55	M 55	A57	M57	AR vs EGF	GeneSymbl	Unigene	Description	AccessionNr
12.24	-1.21	12.56	-0.22	-1.43	EST	Hs.380959	ESTs, Moderately similar to B Chain B, 14-3-3 ZetaPHOSPHOPEPTIDE COMPLEX	BM475869
11.93	-0.58	11.48	-0.73	-1.31	THBS1	Hs.87409	thrombospondin 1	NM_003246
11.05	-0.67	10.95	-0.61	-1.27	CPE	Hs.75360	carboxypeptidase E	NM_001873
11.25	-0.69	11.48	-0.53	-1.22	CAST	Hs.359682	calpastatin	NM_001750
10.71	-0.33	10.55	-0.84	-1.18	SDNSF	Hs.84775	neural stem cell derived neuronal survival protein	NM_139279
11.63	-0.77	12.25	-0.36	-1.13	RBM3	Hs.301404	RNA binding motif protein 3	NM_006743
12.37	-0.51	11.00	-0.62	-1.13	ANXA3	Hs.1378	annexin A3	NM_005139
11.56	-0.80	11.33	-0.32	-1.12	USP14	Hs.75981	ubiquitin specific protease 14 (tRNA-guanine transglycosylase)	NM_005151
11.78	-0.79	11.63	-0.30	-1.09	TM9SF2	Hs.28757	transmembrane 9 superfamily member 2	NM_004800
12.38	-0.74	11.51	-0.34	-1.08	UMPS	Hs.2057	uridine monophosphate synthetase (orotate phosphoribosyl transferase and orotidine-5'-decarboxylase)	NM_000373
10.64	-0.57	10.67	-0.49	-1.06	GALNT3	Hs.278611	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 3 (GalNAc-T3)	NM_004482
10.39	-0.69	9.63	-0.34	-1.04	EST	Hs.433347	Homo sapiens cDNA FLJ30687 fis, clone FCBBF2000379.	BQ717423
11.83	-0.60	12.00	-0.43	-1.03	SFPQ	Hs.180610	splicing factor proline/glutamine rich (polypyrimidine tract binding protein associated)	NM_005066
11.65	-0.39	10.83	-0.64	-1.03	WHIP	Hs.236828	Werner helicase interacting protein	NM_020135
11.57	-0.45	11.04	-0.56	-1.01	SEPW1	Hs.14231	selenoprotein W, 1	NM_003009
12.71	-0.02	12.28	-0.96	-0.98	HNRPH2	Hs.278857	heterogeneous nuclear ribonucleoprotein H2 (H')	NM_019597
10.07	-0.55	9.69	-0.42	-0.97	EPS15	Hs.79095	epidermal growth factor receptor pathway substrate 15	NM_001981
9.97	-0.65	9.66	-0.32	-0.97	MTRR	Hs.153792	5-methyltetrahydrofolate-homocysteine methyltransferase reductase	NM_002454
11.37	-0.67	11.90	-0.29	-0.96	SNRPD3	Hs.1575	small nuclear ribonucleoprotein D3 polypeptide 18kDa	NM_004175
10.22	-0.25	9.91	-0.71	-0.96	CTNNB1	Hs.171271	catenin (cadherin-associated protein), beta 1 (88kD)	
11.11	-0.29		-0.67	-0.96	AMD1	Hs.262476	S-adenosylmethionine decarboxylase 1	NM_001634
11.03	-0.56	11.08	-0.40	-0.95	HCCS	Hs.211571	holocytochrome c synthase (cytochrome c heme-lyase)	NM_005333
10.58	-0.39	10.28	-0.53	-0.92	KIAA0101	Hs.81892	KIAA0101 gene product	NM_014736
11.18	-0.43	10.93	-0.49	-0.91	ADAM10	Hs.172028	a disintegrin and metalloproteinase domain 10	NM_001110
10.84	-0.31	10.97	-0.59	-0.90	MARCKS	Hs.75607	myristoylated alanine-rich protein kinase C substrate	NM_002356
11.43	-0.17	11.43	-0.74	-0.90	ROBO1	Hs.301198	roundabout, axon guidance receptor, homolog 1 (Drosophila)	NM_002941
11.77	-0.02	11.30	-0.88	-0.90	ADAM9	Hs.2442	a disintegrin and metalloproteinase domain 9 (meltrin gamma)	NM_003816
11.39	-0.33	11.37	-0.57	-0.90	EST		No Siginificant similarities	
11.47	-0.04	11.07	-0.86	-0.90	RAP1A	Hs.865	RAP1A, member of RAS oncogene family	NM_002884
11.65	0.37	11.92	0.50	0.87	ENO3	Hs.118804	enolase 3, (beta, muscle)	NM_001976
10.98	-0.06	10.75	0.93	0.87	APG3	Hs.26367	autophagy Apg3p/Aut1p-like	NM_022488
9.81	0.91	8.48	-0.03	0.88	E2F1	Hs.96055	E2F transcription factor 1	NM_005225
13.69	0.82	13.99	0.06	0.88	ACTB	Hs.426930	actin, beta	NM_001101
11.00	0.51	10.96	0.39	0.90	ENO2	Hs.146580	enolase 2, (gamma, neuronal)	NM_001975
13.16	0.28	12.80	0.64	0.92	HNK-1ST	Hs.155553	HNK-1 sulfotransferase	NM_004854
12.77	0.43	12.71	0.53	0.96	ACTA2	Hs.195851	actin, alpha 2, smooth muscle, aorta	NM_001613
12.80	0.32	12.99	0.67	0.99	EIF3S6	Hs.106673	eukaryotic translation initiation factor 3, subunit 6 48kDa	NM_001568
8.20	-0.01	12.32	1.00	0.99	MGC16028	Hs.8984	MGC16028 similar to RIKEN cDNA 1700019E19 gene	NM_052873
10.09	0.07	10.64	0.93	0.99	SLA/LP	Hs.161436	soluble liver antigen/liver pancreas antigen	AI635383
10.60	0.66	9.85	0.38	1.04	HNRPK	Hs.129548	heterogeneous nuclear ribonucleoprotein K	
12.34	0.35	13.09	0.73	1.07	CLK2	Hs.73986	CDC-like kinase 2	NM_001291
10.36	0.13	10.43	0.96	1.10	ASXL1	Hs.3686	additional sex combs like 1 (Drosophila)	AB023195
10.87	0.30	12.49	0.79	1.10	UGP2	Hs.77837	UDP-glucose pyrophosphorylase 2	NM_006759
9.59	0.75	9.33	0.35	1.10	CRMP1	Hs.155392	collapsin response mediator protein 1	NM_001313
10.56	0.13	10.76	1.00	1.13	DDAH2	Hs.247362	dimethylarginine dimethylaminohydrolase 2	NM_013974
9.58	-0.02	12.08	1.18	1.16	FMO2	Hs.132821	flavin containing monooxygenase 2	NM_001460
10.99	1.17	8.93	0.01	1.18	PYY	Hs.169249	peptide YY	NM_004160
12.45	0.38	12.84	0.91	1.29	KRT16	Hs.432448	keratin 16 (focal non-epidermolytic palmoplantar keratoderma)	NM_005557
13.04	0.46	13.19	0.86	1.32	ANXA2	Hs.217493	annexin A2	NM_004039
10.68	1.64	8.42	-0.02	1.62	EST	Hs.164005	Homo sapiens cDNA FLJ13432 fis, clone PLACE1002537.	AK096685
9.91	1.62	8.20	0.02	1.63	GK	Hs.1466	glycerol kinase	NM_000167
10.80	1.65	8.36	0.00	1.65	TH	Hs.178237	tyrosine hydroxylase	
11.61	1.87	9.05	-0.18	1.68	DZIP1	Hs.60177	zinc-finger protein DZIP1	NM_014934
8.89	-0.05	9.82	1.84	1.79	EST	Hs.287687	Homo sapiens cDNA: FLJ21960 fis, clone HEP05517.	AW968569
8.67	-0.04	10.25	1.92	1.88	EVER1	Hs.16165	epidermodysplasia verruciformis 1	NM_007267
12.53	2.27	8.36	0.00	2.28	LTK	Hs.210	leukocyte tyrosine kinase	NM_002344
15.01	2.14	10.65	0.22	2.36	PA2G4	Hs.374491	proliferation-associated 2G4, 38kDa	NM_006191
11.70	2.55	8.80	0.01	2.56	PNLIPRP2	Hs.143113	pancreatic lipase-related protein 2	NM_005396

Curriculum Vitae

Personal

Name: Patrick Urban
Address: Stallenrain 13, 4104 Oberwil, Switzerland
Phone: +41 76 585 80 66
E-mail: patrick.urban@bluewin.com

Work Experience

2006-	Clinical Research Physician, Oncology Global Development, Novartis Pharma AG, Basel, CH
2002-06	Medical and Scientific Advisor, OncoScore AG, Riehen, CH
2002-06	MD/PhD program, NCCR Molecular Oncology, ISREC, Lausanne; University of Basel and Stiftung Tumorbank Basel, CH
2000-01	Assistenzarzt, Kantonsspital Liestal, CH
1997/98	Externships in Saudia Arabia (Thoracic Surgery, Kind Fahd Hospital, Jeddah) and Australia (Dermatology, Skin an Cancer Foundation, Sydney)
1996	Exchange Semester (Erasmus Program), University XI, Paris

Education

2007	6th Masterclass in Clinical Oncology, European School of Oncology, Malta
2002-06	PhD in Molecular Biology, University of Basel, CH
2002	Statistical Analysis of Microarray Data, Swiss Bioinformatics Institute, Lausanne, CH
2001	Dissertation (MD), University of Basel
1993-99	Medical Studies and Staatsexamen, University of Basel
1995	Biometriekurs für Mediziner
1988-92	Gymnasium and Matura Type C, Münchenstein, CH

Publications

Vuaroqueaux V, Urban P et al. *Low E2F1 transcript levels are a strong determinant of favorable breast cancer outcome*. Breast Cancer Res. 2007;9(3):R33.

Urban P et al. *Increased expression of urokinase-type plasminogen activator mRNA determines adverse prognosis in ErbB2-positive primary breast cancer*. J Clin Oncol. 2006 Sep 10;24(26):4245-53.

Zanetti-Dällenbach R, Vuaroqueaux V, Wight E, Labuhn M, Singer G, Urban P et al. *Comparison of gene expression profiles in core biopsies and corresponding surgical breast cancer samples*. Breast Cancer Res. 2006;8(4):R51.

Cicenas J, Urban P et al. *Phosphorylation of tyrosine 1248-ERBB2 measured by chemilumines-cence-linked immunoassay is an independent predictor of poor prognosis in primary breast cancer patients.* Eur J Cancer. 2006 Mar;42(5):636-45.

Cicenas J, Urban P et al. Increased level of phosphorylated akt measured by chemilumines-cence-linked immunosorbent assay is a predictor of poor prognosis in primary breast cancer overexpressing ErbB-2. Breast Cancer Res. 2005;7(4):R394-401.

VOLUME 24 · NUMBER 26 · SEPTEMBER 10 2006

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Increased Expression of Urokinase-Type Plasminogen Activator mRNA Determines Adverse Prognosis in ErbB2-Positive Primary Breast Cancer

Patrick Urban, Vincent Vuaroqueaux, Martin Labuhn, Mauro Delorenzi, Pratyaksha Wirapati, Edward Wight, Hans-Jörg Senn, Christopher Benz, Urs Eppenberger, and Serenella Eppenberger-Castori

From the Swiss Institute for Experimental Cancer Research, National Center of Competence in Research, Molecular Oncology; Swiss Institute of Bioinformatics; Epalinges s/Lausanne; OncoScore AG, Riehen; Department of Obstetrics and Gynaecology, University of Basel; Stiftung Tumorbank Basel, Riehen and University of Basel; Zentrum für Tumordiagnostik und Prävention, St Gallen, Switzerland; and the Buck Institute for Age Research, Novato, CA.

Submitted December 5, 2005; accepted May 19, 2006.

Supported by the Stiftung Tumorbank Basel, the Swiss National Science Foundation (Grant No. 3100-059819.99/1), and the National Center of Competence in Research Molecular Oncology—a research instrument of the Swiss National Science Foundation. Quantitative real-time polymerase chain reaction analysis was supported by OncoScore AG.

P.U. and V.V. contributed equally to this article.

Terms in blue are defined in the glossary, found at the end of this article and online at www.jco.org.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Address reprint requests to Patrick Urban, MD, Stiftung Tumorbank Basel, Lörracherstrasse 50, CH-4125 Riehen, Switzerland; e-mail: patrick.urban@unibas.ch.

© 2006 by American Society of Clinical Oncology

0732-183X/06/2426-4245/$20.00

DOI: 10.1200/JCO.2005.05.1912

ABSTRACT

Purpose
To evaluate and validate mRNA expression markers capable of identifying patients with ErbB2-positive breast cancer associated with distant metastasis and reduced survival.

Patients and Methods
Expression of 60 genes involved in breast cancer biology was assessed by quantitative real-time PCR (qrt-PCR) in 317 primary breast cancer patients and correlated with clinical outcome data. Results were validated subsequently using two previously published and publicly available microarray data sets with different patient populations comprising 295 and 286 breast cancer samples, respectively.

Results
Of the 60 genes measured by qrt-PCR, urokinase-type plasminogen activator (*uPA* or *PLAU*) mRNA expression was the most significant marker associated with distant metastasis-free survival (MFS) by univariate Cox analysis in patients with ErbB2-positive tumors and an independent factor in multivariate analysis. Subsequent validation in two microarray data sets confirmed the prognostic value of *uPA* in ErbB2-positive tumors by both univariate and multivariate analysis. *uPA* mRNA expression was not significantly associated with MFS in ErbB2-negative tumors. Kaplan-Meier analysis showed in all three study populations that patients with ErbB2-positive/uPA–positive tumors exhibited significantly reduced MFS (hazard ratios [HR], 4.3; 95% CI, 1.6 to 11.8; HR, 2.7; 95% CI, 1.2 to 6.2; and, HR, 2.8; 95% CI, 1.1 to 7.1; all $P < .02$) as compared with the group with ErbB2-positive/uPA–negative tumors who exhibited similar outcome to those with ErbB2-negative tumors, irrespective of uPA status.

Conclusion
After evaluation of 898 breast cancer patients, *uPA* mRNA expression emerged as a powerful prognostic indicator in ErbB2-positive tumors. These results were consistent among three independent study populations assayed by different techniques, including qrt-PCR and two microarray platforms.

J Clin Oncol 24:4245-4253. © 2006 by American Society of Clinical Oncology

INTRODUCTION

The oncogenic receptor tyrosine kinase, *ErbB2* (*HER2/neu*), has received major attention as a breast cancer biomarker and is mechanistically linked with a variety of malignant processes including dysregulated cell growth and proliferation, tumor angiogenesis, tissue invasion, and metastases (reviewed previously[1-4]). Overexpression of *ErbB2* mRNA and protein resulting from the amplified oncogene is observed in 15% to 30% of all newly diagnosed breast cancer patients, in which it is associated with poor prognosis and serves as a predictor of clinical responsiveness to the anti-ErbB2 therapeutic trastuzumab (Herceptin; Genentech, South San Francisco, CA). Based on commonly used breast cancer risk assessment criteria, including the most recent St Gallen international expert consensus criteria, most patients with ErbB2-positive breast cancer are assigned to the highest risk category for metastatic relapse.[5] However, not all patients with ErbB2-positive breast cancer develop nodal or distant metastases; moreover, the molecular mechanisms by which *ErbB2* overexpression results in clinically more aggressive disease remain poorly understood. Therefore, identification of additional risk markers is needed to elucidate pathways responsible for metastatic relapse as well as to

Downloaded from www.jco.org on September 19, 2006 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

improve the risk classification of patients diagnosed with ErbB2-positive breast cancer.

Recently, expression profiling has shown promise in being able to use breast cancer gene expression signatures to predict metastatic relapse.[6-8] However, to date these studies have not focused on the identification of outcome predictors for patients newly diagnosed with ErbB2-positive breast cancer. To address this need, 60 candidate genes were selected for expression profiling based on their reported links to malignant cell behavior and breast cancer biology and their representation in all major cancer pathways, including cell proliferation and survival, invasiveness, angiogenic potential, and endocrine dependence (Table 1). All 60 candidates were assessed by quantitative real-time polymerase chain reaction (qrt-PCR) in a set of 317 primary breast cancer biopsies and their expression levels were correlated with clinical outcome data. Five candidate genes were identified whose expression levels showed significant univariate association with distant metastasis-free survival (MFS) in ErbB2-positive tumors. The most significant of these markers, urokinase-type plasminogen activator (*uPA* or *PLAU*), was further validated in two independent and previously reported breast cancer study populations for which gene expression (microarray) and patient outcome data are publicly accessible.[6,7]

PATIENTS AND METHODS

Study Patients

Fresh frozen tissue from primary breast cancer patients was obtained from the Stiftung Tumorbank Basel (STB, Basel, Switzerland). Exclusion criteria included neoadjuvant therapy, insufficient tumor content (< 50%), and poor RNA quality resulting in 317 eligible samples (STB study). All patients underwent surgery between 1992 and 1996. Quantitative protein levels of hormone receptors, ErbB2 and uPA, were prospectively determined at the time of surgery.[9,10] Median patient age at diagnosis was 60 years (range, 27 to 88 years). Fifty-seven patients (18%) developed distant metastasis within a median MFS time of 44 months (range, 8 to 90 months). Forty-six percent of tumors were node-positive, 73% were estrogen-receptor (ER) –positive (> 20 fmol/mg protein by Enzyme Immuno Assay[11]), and 22% were ErbB2-positive (see Statistical Analysis). Systemic adjuvant hormone therapy was administrated to 43% of patients, chemotherapy to 22% of patients, and combination adjuvant therapy to 16% of patients; virtually all hormonal therapy was in the form of tamoxifen, with chemotherapy primarily consisting of cyclophosphamide, methotrexate, and fluorouracil and 9% of patients receiving anthracyclines. STB is an approved nonprofit laboratory that performs routine tumor biomarker measurements and clinical research under Swiss health regulatory office compliance with specific regard to ethical standards and patient confidentiality.

For validation purposes, two additional breast cancer study populations were also analyzed using the publicly accessible microarray and outcome data sets reported by Van de Vijver et al (Amsterdam study; Agilent microarray platform; Agilent Technologies, Palo Alto, CA)[6] and Wang et al (Rotterdam study; Affymetrix microarray platform; Affymetrix, Santa Clara, CA).[7] Major differences in the three study populations included patient age, nodal status, and adjuvant treatments. MFS was the only outcome parameter commonly available for all three study populations. Comparative patient and tumor characteristics from the STB, Amsterdam, and Rotterdam study populations are summarized in Table 2.

RNA Extraction and Quantitative Real-Time PCR

Gene expression measurements by quantitative real-time PCR were performed as reported.[12] In brief, RNA was extracted using the RNAeasy Mini Kit (Qiagen, Hilden, Germany), quantified and quality-checked on a Bioanalyzer 2100 (RNA 6000 Nano LabChip-Kit; Agilent Technologies). High quality RNA samples (28S:18S > 1) were reverse transcribed (1μg total RNA,

Table 1. List of the Candidate Genes (N = 60)

Gene	Symbol	Ref Sequence
Adrenomedullin	*ADM*	NM_001124
Amphiregulin	*AREG*	NM_001657
Bcl2-associated X protein	*BAX*	NM_138761
B-cell lymphoma 2	*BCL2*	NM_000633
Survivin	*BIRC5*	NM_001168
Cyclin D1	*CCND1*	NM_053056
p21/Cip1	*CDKN1A*	NM_078467
p27/Kip1	*CDKN1B*	NM_004064
Beta-Catenin	*CTNNB1*	NM_001904
Cathepsin B	*CTSB*	NM_001908
Cathepsin D	*CTSD*	NM_001909
Dihydropyrimidine deshydrogenase	*DPYD*	NM_000110
Transcription factor E2F1	*E2F1*	NM_005225
Thymidine phosphorylase	*ECGF1*	NM_001953
Epidermal growth factor	*EGF*	NM_001963
Epidermal growth factor receptor 1	*EGFR*	NM_005228
Epidermal growth factor receptor 2	*ERBB2*	NM_004448
Epidermal growth factor receptor 3	*ERBB3*	NM_001982
Epidermal growth factor receptor 4	*ERBB4*	NM_005235
Estrogen receptor 1	*ESR1*	NM_000125
Vascular endothelial growth factor receptor 1	*FLT1*	NM_002019
Vascular endothelial growth factor receptor 3	*FLT4*	NM_002020
Glutathione *S*-transferase pi	*GSTP1*	NM_000852
Hypoxia inducible factor 1-alpha	*HIF1A*	NM_001530
17 beta-hydroxy steroid deshydrogenase 1	*HSD17B1*	NM_000413
Insulin-like growth factor 1	*IGF1*	NM_000618
Insulin-like growth factor receptor 1	*IGF1R*	NM_000875
Insulin-like growth factor 2	*IGF2*	NM_000612
Insulin-like growth factor receptor 2	*IGF2R*	NM_000876
Vascular endothelial growth factor receptor 2	*KDR*	NM_002253
Cytokeratin 19	*KRT19*	NM_002276
Cytokeratin 7	*KRT7*	NM_005556
Metalloprotease 1	*MMP1*	NM_002421
Metalloprotease 11	*MMP11*	NM_005940
Metalloprotease 2	*MMP2*	NM_004530
Metalloprotease 3	*MMP3*	NM_002422
Metalloprotease 7	*MMP7*	NM_002423
Metalloprotease 9	*MMP9*	NM_004994
Amplified in breast cancer 1 (AIB1)	*NCOA3*	NM_006534
Peptidylglycine alpha-amidating monooxygenase	*PAM*	NM_000919
Progesterone receptor	*PGR*	NM_000926
Urokinase-type plasminogen activator (uPA)	*PLAU*	NM_002658
Urokinase-type plasminogen activator receptor (uPAR)	*PLAUR*	NM_002659
Cyclooxygenase-2	*PTGS2*	NM_000963
Retinoblastoma	*RB1*	NM_000321
Plasminogen activator inhibitor	*SERPINE1*	NM_000602
Transcription factor 4	*TCF4*	NM_003199
Transforming growth factor alpha	*TGFA*	NM_003236
Transforming growth factor beta 1	*TGFB1*	NM_000660
Tissue inhibitor of metalloproteases 1	*TIMP1*	NM_003254
Tissue inhibitor of metalloproteases 2	*TIMP2*	NM_003255
Tissue inhibitor of metalloproteases 3	*TIMP3*	NM_000362
Tissue inhibitor of metalloproteases 4	*TIMP4*	NM_003256
Thymidine kinase 1	*TK1*	NM_003258
Topoisomerase II-2 alpha	*TOP2A*	NM_001067
Thymidilate synthase	*TYMS*	NM_001071
Vascular endothelial growth factor A	*VEGF*	NM_003376
Vascular endothelial growth factor B	*VEGFB*	NM_003377
Vascular endothelial growth factor C	*VEGFC*	NM_005429
Vascular endothelial growth factor D	*FIGF*	NM_004469

NOTE. Official symbol provided by the HUGO Nomenclature Committee (http://www.gene.ucl.ac.uk/nomenclature).

Downloaded from www.jco.org on September 19, 2006 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

Table 2. Patient and Tumor Characteristics

Characteristic	Study: STB, Quantitative RT-PCR: All (N = 317) No.	%	ErbB2+ (n = 70) No.	%	ErbB2− (n = 247) No.	%	Amsterdam, Microarray (Agilent): All (N = 295) No.	%	ErbB2+ (n = 52) No.	%	ErbB2− (n = 243) No.	%	Rotterdam, Microarray (Affymetrix): All (N = 286) No.	%	ErbB2+ (n = 51) No.	%	ErbB2− (n = 235) No.	%
Age, years																		
Mean	60		58		60		44		43		44		54		NA		NA	
Median	59		57		60		44		43		45		NA		NA		NA	
≤ 40	20	6	7	10	13	5	75	25	16	31	59	24	36	13	NA		NA	
41–55	110	35	25	36	85	35	220	75	36	69	184	76	129	45	NA		NA	
≥ 56	187	59	38	54	149	60	0	0	0		0		121	42	NA		NA	
pT stage																		
pT1	100	32	21	30	79	32	155	53	27	52	128	53	146	51	NA		NA	
pT2	183	58	41	60	142	58	140	47	25	48	115	47	132	46	NA		NA	
pT3/4	33	10	7	10	26	10	0		0		0		8	3	NA		NA	
pN status																		
Negative	161	51	32	46	129	52	151	51	26	50	125	51	286	100	51	100	235	100
Positive	136	43	36	51	100	41	144	49	26	50	118	49	0		0		0	
Unknown	20	6	2	3	18	7	0		0		0		0		0		0	
ER status*																		
Positive	231	73	38	54	193	78	226	77	32	62	194	80	209	73	32	61	177	75
Negative	86	27	32	46	54	22	69	23	20	38	49	20	77	27	19	39	58	25
ErbB2 status																		
Positive	70	22	70	100	0		52	18	52	100	0		51	18	51	100	0	
Negative	247	78	0		247	100	243	82	0		243	100	235	82	0		235	100
Adjuvant therapy																		
Total	245	77	50	71	195	79	130	44	20	38	110	45	0		0		0	
Hormone	135	43	21	30	114	46	20	7	2	4	18	7	0		0		0	
Chemotherapy	72	22	16	23	56	23	90	30	17	33	73	30	0		0		0	
Combination	38	12	13	18	25	10	20	7	1	2	19	8	0		0		0	
None	60	19	15	21	45	18	165	56	32	62	133	65	286	100	51	100	235	100
Unknown	12	3	5	7	7	3	0		0		0		0		0		0	
Follow-up																		
Metastasis	57	18	17	24	40	16	101	34	23	44	78	32	107	37	19	37	88	37
Mean MFS, years	3.7		3.5		3.8		7.3		6.4		7.5		6.5		6.3		6.5	
Median MFS, years	3.6		3.3		3.7		6.8		6.0		7.0		7.2		6.8		7.3	

Abbreviations: STB, Stiftung Tumorbank Basel; RT-PCR, real-time polymerase chain reaction; ER, estrogen receptor; NA, not available; MFS, metastasis-free survival.
*ER positive ≥ 20 fmol/mg protein (EIA) for STB; for Amsterdam and Rotterdam studies. [6,7] ER status was the only parameter that differed significantly between patients with ErbB2 ± tumors: $P < .001$ (STB); $P = .007$ (Amsterdam); $P = .081$ (Rotterdam).

10 mmol/L DTT, 1μg of hexamer primers, 2 U of MMLV Reverse Transcriptase [Invitrogen, Carlsbad, CA], 40 U of RNasin [Promega, Madison, WI], 0.5 mmol/L each dNTP [Promega], 1× reaction buffer). PCR primers were designed to be cDNA specific and ordered at GeneScan Europe (Freiburg, Germany). PCR was performed in 40 cycles on a ABI Prism 7000 using 2× SYBR Green I Master Mix (Applied Biosystems, Foster City, CA) in a final volume of 25 μL. Relative quantities (Δ cycle threshold values) were obtained by normalization against ribosomal 18S RNA. The 60 candidate genes measured by quantitative real-time PCR are listed in Table 1.

Statistical Analysis

ErbB2 status was determined based on mRNA expression levels using a mixture model in all three study populations, because ErbB2 status was not expressly available for the Amsterdam and Rotterdam studies and *ErbB2* mRNA expression revealed a clear bimodal distribution (Online Only Appendix; Tables A1, -A3 and Figs A1-A4). Subsequent comparison of *ErbB2* mRNA expression levels and *ErbB2* amplification measured by fluorescence in situ hybridization (FISH) in a subset of 100 STB tumor samples demonstrated close agreement between mRNA and FISH, as well as between ErbB2 status determined by the mixture model and FISH, suggesting that the statistical model provides accurate assessment of ErbB2 status (95% agreement, kappa statistic 0.83;[13] Online Only Appendix). The prognostic value of biomarkers was assessed by univariate and multivariate Cox regression against distant MFS, with and without stratification by ER status, treatment group, and surrogate markers of tumor proliferation (average mRNA expression levels of *BIRC5, TOP2A, TYMS, TK1*, and *E2F1*). Results were summarized in Tables, including HR, level of significance (P value), and 95% CI. P values in multivariate analysis were based on Wald tests. Cutoff values for *uPA* were evaluated in ErbB2-positive and ErbB2-negative groups separately by calculating the 5-year MFS as a function of putative *uPA* cutoff (Online Only Appendix). Survival probabilities for MFS were calculated according to the Kaplan-Meier method and group differences assessed by the log-rank test. Spearman rank correlation was used to calculate correlations among biomarkers, the t test was used to investigate differences in mean expression values between groups, and the Fisher's exact test was used to assess differences between categoric data. Log-expression ratios from the Amsterdam study were retransformed to log basis 2. P values of .05 or less were considered to be significant. All statistical

Downloaded from www.jco.org on September 19, 2006 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

analyses were preformed using "R" statistical software version 2.0.1 (http://www.r-project.org).

RESULTS

Candidate Genes Associated With MFS in the ErbB2-Positive STB Breast Cancer Population

The prognostic value of all 60 biomarkers detected by quantitative real-time PCR was individually assessed in the 70 STB ErbB2-positive tumors (22%) by univariate Cox regression revealing five candidates (*uPA, MMP 3, MMP11,* uPAR [PLAUR], and *MMP1*) significantly associated with MFS (Table 3). All five genes encode proteases and their levels of mRNA expression correlated strongly with one another (Online Only Appendix). Compared to *uPA*, the effect of *MMP11, uPAR, MMP1*, and *MMP3* was weaker and nonsignificant when taken together with *uPA* in a multivariate model (data not shown). In contrast, none of the five candidates showed significant prognostic value with respect to MFS by univariate analysis in the 247 STB patients with ErbB2-negative tumors (Table 3). However, *uPA, uPAR*, and *MMP1* were significantly associated with MFS in the overall population. To investigate possible confounding by treatment group, ER status, or proliferation (as defined by the average expression value of *BIRC5, TOP2A, TYMS, TK1*, and *E2F1*), stratified Cox analysis was performed; in all of these models uPA retained significant prognostic value in ErbB2-positive but not ErbB2-negative tumors (data not shown).

Prognostic Validation of uPA in Amsterdam and Rotterdam Breast Cancer Populations

Amsterdam and Rotterdam breast cancer study sets were dichotomized into ErbB2-positive and ErbB2-negative patients as described (Patients and Methods and Online Only Appendix). Using the established 0.3 and 12.9 mRNA cutoff values to dichotomize the Amsterdam and Rotterdam patients, respectively, 52 of the Amsterdam patients (18%) and 51 of the Rotterdam patients (18%) were classified as ErbB2-positive (Table 2). Within each of these study sets, the prognostic value of *uPA* was independently assessed against both the ErbB2-positive and ErbB2-negative tumors, with and without stratification by ER status, treatment group (Amsterdam study only), and tumor proliferation. As presented in Table 4, *uPA* demonstrated significant prognostic association with MFS in ErbB2-positive but not in ErbB2-negative breast cancers. *uPA* was also significantly associated with MFS in the overall population, although displaying a weaker prognostic effect (Table 4).

Table 4. Univariate Cox Analysis (MFS) for *uPA* in ErbB2-Positive Tumors, ErbB2-Negative Tumors, and the Overall Population According to the Three Data Sets

Group	Univariate MFS		
	P	HR*	95% CI
STB			
ErbB2+ (n = 70)	.005	3.14	1.36 to 7.27
ErbB2− (n = 247)	.364	1.21	0.80 to 1.82
All (N = 317)	.021	1.53	1.07 to 2.19
Amsterdam			
ErbB2+ (n = 52)	.005	2.14	1.25 to 3.66
ErbB2− (n = 243)	.148	1.24	0.80 to 1.66
All (N = 295)	.005	1.44	1.11 to 1.85
Rotterdam			
ErbB2+ (n = 51)	.003	2.88	1.44 to 5.74
ErbB2− (n = 235)	.305	1.15	0.88 to 1.49
All (N = 286)	.039	1.29	1.01 to 1.63

Abbreviations: MFS, metastasis-free survival; uPA, urokinase-type plasminogen activator; HR, hazard ratio; STB, Stiftung Tumorbank Basel.
*Estimated HR using standardization by interquartilerange (change in uPA expression from the first to the third quartile). Studies used different scales (STB: delta [Ct]; Amsterdam: log [ratio]; Rotterdam: log [intensity]).

Multivariate Cox Analysis

The prognostic value of *uPA* mRNA expression was further assessed by multivariate analysis across all three study populations against nodal status, grade, tumor size, age, hormone receptor status, tumor proliferation, and treatment. Of note, the Rotterdam study patients were all untreated and node-negative, with only information on ER status available. Across all three study populations, *uPA* retained independent prognostic value and was significantly associated with MFS in ErbB2-positive tumors (Table 5). In addition to *uPA*, nodal status and chemotherapy retained independent prognostic value. In multivariate analysis, *uPA* also retained significance as an independent prognostic marker across all breast cancer patients, but not for patients with ErbB2-negative tumors (data not shown). Overall survival analysis demonstrated consistent findings with MFS

Table 3. Univariate Cox Analysis (MFS) in ErbB2-Positive (n = 70), ErbB2-Negative (n = 247), and Overall STB Breast Cancer Patients

Gene	Univariate MFS								
	ErbB2+ (n = 70)			ErbB2− (n = 247)			All (N = 317)		
	P	HR*	95% CI	*P*	HR*	95% CI	*P*	HR*	95% CI
uPA	.005	3.14	1.36 to 7.27	.364	1.21	0.80 to 1.82	.021	1.53	1.07 to 2.19
MMP3	.009	2.88	1.30 to 6.40	.998	1.00	0.69 to 1.44	.237	1.23	0.88 to 1.72
MMP11	.012	3.02	1.26 to 7.24	.652	1.11	0.71 to 1.72	.064	1.45	0.98 to 2.14
uPAR	.014	2.88	1.22 to 6.78	.372	1.21	0.83 to 1.83	.035	1.48	1.03 to 2.12
MMP1	.032	2.23	1.05 to 4.74	.101	1.37	0.93 to 2.32	.003	1.74	1.20 to 2.51

NOTE. Top five genes significant in the ErbB2-positive group are shown.
Abbreviations: HR, hazard ratio; MFS, metastasis-free survival; STB, Stiftung Tumorbank Basel; uPA, urokinase-type plasminogen activator; MMP, matrix metalloproteinases; uPAR, urokinase-type plasminogen activator receptor.
*Estimated HR using standardization by interquartile range (change in expression from the first to the third quartile).

Downloaded from www.jco.org on September 19, 2006 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

Table 5. Multivariate Cox Analysis for Patients With ErbB2-Positive Tumors

Factor	ErbB2+ MFS								
	Univariate*			Multivariate*			Multivariate†		
	P	HR	95% CI	*P*	HR	95% CI	*P*	HR	95% CI
uPA status									
> 75 *v* ≤ 75th percentile	< .001	3.12	1.86 to 5.23	.001	3.24	1.60 to 6.57	< .001	3.08	1.76 to 5.38
ER status									
Positive *v* negative‡	.100	0.65	0.39 to 1.09	.535	0.79	0.38 to 1.65	.325	0.76	0.43 to 1.32
Nodal status									
> 3 *v* 0-3 positive nodes	< .001	8.63	3.92 to 18.9	< .001	12.2	4.72 to 31.5	< .001	11.5	4.62 to 28.9
Tumor size, cm									
> 2 *v* ≤ 2	.027	2.19	1.09 to 4.38	.108	1.98	0.86 to 4.55	NI	NI	NI
Age, years									
> 40 *v* ≤ 40	.421	0.71	0.30 to 1.65	.222	0.46	0.13 to 1.60	NI	NI	NI
Grade									
3 *v* 1 or 2	.099	1.72	0.90 to 3.29	.780	1.12	0.49 to 2.56	NI	NI	NI
Proliferation									
> median *v* ≤ median	.140	1.58	0.86 to 2.89	.197	1.85	0.73 to 4.73	.107	1.73	0.89 to 3.35
Chemotherapy	.881	0.95	0.49 to 1.84	.012	0.30	0.12 to 0.77	.045	0.45	0.21 to 0.98

Abbreviations: MFS, metastasis-free survival; HR, hazard ratio; uPA, urokinase-type plasminogen activator; ER, estrogen receptor.
*Univariate and multivariate analysis based on Cox regression stratified by data set. All variables were coded 0 or 1 according to the criteria below. Since information on tumor size, age, and grade were not available for the Rotterdam data set this analysis corresponds to the STB and Amsterdam data sets (n = 110; 12 excluded due to missing values).
†Cox regression including all three study populations stratified by data set. Tumor size, age, and grade were not included in this model since not available for the Rotterdam data set (n = 165; 8 excluded due to missing values). For multivariate analysis of each individual data set see supplementary information.
‡≥ 20 *v* < 20 fmol/mg (STB); for Amsterdam and Rottendam studies, see references.[6,7]

among the Amsterdam study population, for which this outcome parameter was also available. Multivariate analysis for each of the three data sets can be found in the Online Only Appendix.

Stratification of ErbB2-Positive Tumors by uPA

The cutoff value for uPA status was set at the 75th percentile after evaluation of 5-year MFS as a function of *uPA* mRNA expression in the STB study population (Online Only Appendix). The *uPA* cutoff was chosen to identify ErbB2-positive tumors with poor prognosis (5-year MFS of less than 50%) and to optimize separation between the groups. The 75th percentile agreed with previously published uPA cutoff values based on protein determination.[14,15] This same *uPA* cutoff value was used to dichotomize ErbB2-positive and ErbB2-negative tumors in all three study populations and for outcome determination by Kaplan-Meier analyses (Fig 1). Within each study set, uPA status proved to be a strong prognostic factor for the development of distant metastasis, but only among ErbB2-positive breast cancer patients (HRs: STB study: 4.3; 95% CI, 1.6 to 11.8; Rotterdam study: 2.8; 95% CI, 1.1 to 7.1]; Amsterdam study: 2.7; 95% CI, 1.2 to 6.2; all $P < .02$). The prognostic value of *uPA* overexpression was even more pronounced for overall survival among ErbB2-positive Amsterdam patients (HR, 3.5; 95% CI, 1.5 to 8.6; $P < .001$, Fig 1). A search for alternative cutoff values for *uPA* failed to identify any level of *uPA* mRNA expression significantly associated with MFS among patients with ErbB2-negative breast cancer (Online Only Appendix). These stratification findings were independent of tumor treatment or stage, since multivariate analysis (see Multivariate Cox Analysis section) confirmed the independent prognostic value of *uPA* in ErbB2-positive but not in ErbB2-negative tumors (Table 5 and the Online Only Appendix).

Clinical and Prognostic Parameters Associated With uPA *Stratified ErbB2-Positive Tumors*

In all three study sets, ER expression was significantly lower in ErbB2-positive as compared with ErbB2-negative tumors. However, no significant difference was found with respect to ER status, nodal status, tumor size, age, or treatment group between the dichotomized ErbB2-positive/uPA–negative and ErbB2-positive/uPA–positive tumors in any of the three study populations (Table 6). In addition, the 70-gene prognostic signature[6] was investigated in the Amsterdam study population and found to classify 85% of ErbB2-positive tumors into a poor prognosis group; however, this signature showed no difference with regard to uPA status among patients with ErbB2-positive tumors (Table 6 and the online-only Appendix). Except for the Rotterdam study population, *uPA* expression was significantly higher in ErbB2-positive compared with ErbB2-negative tumors.

DISCUSSION

Initial evaluation of mRNA profiles from 60 cancer-related genes identified five protease-related prognostic candidates whose increased expression appeared significantly associated with MFS in ErbB2-positive breast cancer. The most significant of these, *uPA*, was further validated in two independent breast cancer study populations.[6,7] Despite analysis of three distinct study populations by different assay methods to measure *uPA* mRNA (quantitative real-time PCR *v* Agilent and Affymetrix microarrays, Table 2), expression levels of this biomarker showed its consistent and independent prognostic significance that was restricted to patients with ErbB2-positive breast cancer.

Matrix metalloproteinases (MMP) and the uPA system are known to play important roles in cancer cell invasion and metastasis;[16-18] and the initially identified five protease-related prognostic candidates (Table 3) all showed a strong pattern of correlation among the STB study population (Online Only Appendix). uPA and other protein members of the uPA system are well established breast cancer prognostics.[10] However, this study showing that *uPA* mRNA

Downloaded from www.jco.org on September 19, 2006 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

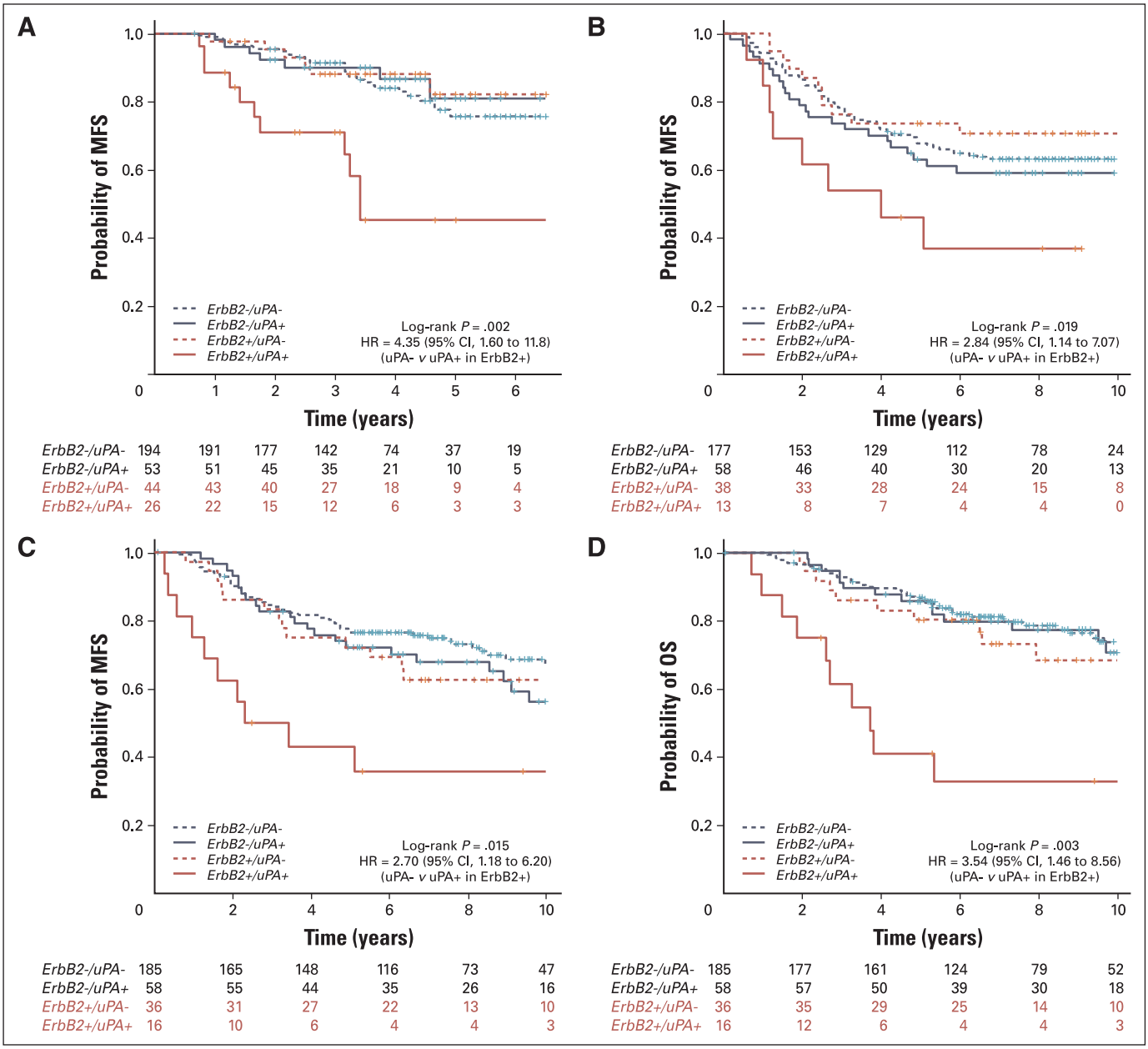


Fig 1. Combined Kaplan-Meier analysis (ErbB2/uPA). ErbB2-status is indicated by color (red: ErbB-positive; black: ErbB2-negative); urokinase-type plasminogen activator (uPA) status by line type (solid: uPA-positive; dashed: uPA-negative). (A) Stiftung Tumorbank Basel study (metastasis-free survival [MFS]); (B) Rottendam study (MFS); (C) Amsterdam study (MFS); (D) Amsterdam study (overall survival).

expression is of significant prognostic value only in ErbB2-positive patients is a novel observation. In conjunction with ErbB2 and uPA status, Zemzoum et al[19] used another member of the uPA system, plasminogen activator inhibitor-1 (PAI-1), to define a high-risk subgroup of breast cancer patients. Among the STB patients, a strong correlation was observed between *uPA* and *PAI-1* mRNA levels (r_s = 0.64). In this same population, *PAI-1* was significantly associated with MFS by univariate Cox regression, but showed only borderline significance for ErbB2-positive patients (P = .08), hence it was not included among the five protease-related prognostic candidates (Table 3). Like *uPA*, *PAI-1* was not a significantly associated with MFS in ErbB2-negative tumors. Previously published reports have also demonstrated the predictive value of uPA with respect to adjuvant chemotherapy.[20,21] However, the uniform association of *uPA* mRNA expression with survival observed across three distinct study populations was not likely due to predictive value with regard to chemotherapy because one of the validating study populations (Rotterdam) received no chemotherapy; as well, by multivariate analysis including chemotherapy, *uPA* remained an independent and significant factor in determining patient outcome in the other two study populations (Table 5).

Konecny et al[14] and Zemzoum et al[19] previously investigated uPA and ErbB2 in breast cancer and reported their independent prognostic value, also illustrating that observed prognostic relationships may be dependent on the method of biomarker assessment (for example,

Downloaded from www.jco.org on September 19, 2006 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

Table 6. Associations Between Clinical Parameters and uPA Status in ErbB2-Positive Tumors

Factor	STB Study					Amsterdam Study					Rotterdam Study				
	uPA−		uPA+			uPA−		uPA+			uPA−		uPA+		
	No.	%	No.	%	*P*	No.	%	No.	%	*P*	No.	%	No.	%	*P*
Nodal status					.211					.369					.99
Negative	23	53	9	36		20	56	6	38		38	100	13	100	
Positive	20	47	16	64		16	44	10	62		0		0		
Tumor size, cm															
Mean	3.0		2.7		.620	2.2		2.3		.450	NA		NA		—
Differentiation					.799					.99	NA		NA		—
Good/moderate	20	53	15	58		15	42	7	44		NA		NA		
Poor	18	47	11	42		21	58	9	56		NA		NA		
Age, years															—
Mean	57.5		58.5		.753	43.3		41.6		.458	NA		NA		
ER status					.99					.759					.192
Negative	20	45	12	46		13	36	7	44		12	32	7	54	
Positive	24	55	14	54		23	64	9	56		26	68	6	46	
Treatment					.748					.393					—
Chemotherapy	10	24	6	26		13	36	4	25		0		0		
Hormone	14	33	7	30		2	6	0	0		0		0		
Combined	7	17	6	26		0	0	1	6		0		0		
None	11	26	4	17		21	58	11	69		0		0		
70-gene signature[6]					—					.99					—
Good prognosis	NA		NA			6	17	2	13		NA		NA		
Poor prognosis	NA		NA			30	83	14	87		NA		NA		

Abbreviations: uPA, urokinase-type plasminogen activator; STB, Stiftung Tumorbank Basel; ER, estrogen receptor; NA, not available.

immunohistochemistry *v* FISH for *ErbB2*) as well as the choice of the clinical outcome parameter (MFS *v* disease-free or overall survival). Unlike these earlier reports which evaluated uPA by immunoassays (enzyme-linked immunosorbent assay [ELISA]) on protein level and ErbB2 by protein immunohistochemistry and DNA FISH, this study quantitated expression of *uPA* and *ErbB2* at the mRNA level. *ErbB2* mRNA expression has been shown by several groups including our own to correlate well with more standard protein and DNA measures of ErbB2 status.[12, 22-24] Although the prognostic value of microarray and quantitative real-time PCR determined *uPA* mRNA levels were found in this study to be highly comparable, prior analysis of STB patients demonstrated only a moderate correlation between uPA protein content (ELISA) and mRNA levels ($r_s = 0.46$), suggesting some biologic discordance between these two measures of uPA bioactivity. Using an optimized *uPA* mRNA cutoff value capable of prognostically subsetting ErbB2-positive tumors, 25% of all breast cancer patients in this study scored as uPA–positive, which is consistent with the reported proportion of poor-risk breast cancers overexpressing *uPA* when measured by protein content.[14,15] To avoid overfitting, cutoff optimization was performed only on the STB study population and this optimized value was subsequently validated in the Amsterdam and Rotterdam study populations. Of note, we were unable to find any uPA cutoff showing prognostic significance among the ErbB2-negative breast cancer tumors (Online Only Appendix).

Recently, gene expression profiling using microarrays or quantitative real-time PCR have proved useful in predicting breast cancer outcomes.[6-8] Interestingly, neither the prognostic gene signatures published by the Amsterdam[6] nor the Rotterdam[7] groups contained *uPA*. It was speculated that *uPA* as a single gene may have limited prognostic value compared with gene signatures and that mRNA expression may not be as informative as protein biomarker expression.[25] This study revealed, however, that *uPA* mRNA level was highly informative and a prognostically valuable risk indicator when used for ErbB2-positive breast cancers. Given the finding that *uPA* mRNA expression had no apparent prognostic value among ErbB2-negative breast cancers, it is possible that the overall prognostic importance of *uPA* (and other genes) has been underestimated in past breast cancer studies not stratified by ErbB2 status. Interestingly, the 70-gene prognostic signature published by the Amsterdam group[6] classified almost all ErbB2-positive patients into the poor prognosis group. In terms of survival there was no difference between the poor prognosis/ErbB2-positive and poor prognosis/ErbB2-negative tumors (Online Only Appendix). In contrast to *uPA* and other proteases, the 70-gene prognostic signature appears to largely discriminate a low-risk group from among patients with ErbB2-negative tumors.

The existence of a molecular mechanism linking uPA upregulation with ErbB2 overexpression is supported by numerous reports that ErbB2 activation increases the transcriptional expression of *uPA* and various members of the *MMP* family.[4,26-28] Consistent with these reports, we observed significantly higher *uPA* expression levels in ErbB2-positive relative to ErbB2-negative tumors; and similar findings have been reported for other cancers.[29-31] However, the mechanistic interactions linking *ErbB2* and *uPA* overexpression with increased breast cancer metastatic potential are likely more complex, also because ErbB2 overexpression occurs only within tumor epithelium while uPA and other proteases are primarily overexpressed in tumor stroma.[32,33] Stromal production of *uPA* (and other proteases) must be considered when microdissection techniques are used to obtain breast cancer RNA for expression profiling. The functional interaction between overexpressed *ErbB2* and *uPA* also deserves

Downloaded from www.jco.org on September 19, 2006 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

greater study at both cellular and clinical levels since it is unclear why ErbB2-negative/uPA–positive and ErbB2-positive/uPA–negative breast cancers are no more clinically aggressive than ErbB2-negative/uPA–negative breast cancers. Apparently the proteolytic activity of uPA alone is insufficient to determine the metastatic outcome, and perhaps requires the cell proliferation and survival advantages provided by activation of ErbB2. Similarly, ErbB2 activation may only predispose to a metastatic outcome in conjunction with cellular expression of the uPA and other proteases.

In conclusion, this study indicates that overexpression of *uPA* mRNA levels in ErbB2-positive breast cancers determines the aggressive and highly metastatic clinical behavior previously attributed only to activation of the ErbB2 receptor tyrosine kinase. ErbB2 status is determined on all newly diagnosed breast cancers. While the prognostic and predictive potential of uPA ELISA-based measurements is supported by the highest level of clinical evidence, such uPA assays have proven impractical for widespread clinical application. Future studies showing that the *uPA* quantitative real-time PCR assay as employed in this study may also be adapted to small paraffin-archived breast cancer samples will undoubtedly lead to more rapid and widespread clinical measurement of *uPA*. With expression profiling now integrated into several ongoing multinational breast cancer adjuvant trials, attention to breast tumor *uPA* mRNA levels, measured alone or as part of gene expression signature, will further enhance the prognostic and predictive value of this important protease biomarker.

REFERENCES

1. Yarden Y: Biology of HER2 and its importance in breast cancer. Oncology 61:1-13, 2001 (suppl 2)

2. Olayioye MA, Neve RM, Lane HA, et al: The ErbB signaling network: Receptor heterodimerization in development and cancer. Embo J 19:3159-3167, 2000

3. Neve RM, Lane HA, Hynes NE: The role of overexpressed HER2 in transformation. Ann Oncol 12:S9-13, 2001 (Suppl 1)

4. Eccles SA: The role of c-erbB-2/HER2/neu in breast cancer progression and metastasis. J Mammary Gland Biol Neoplasia 6:393-406, 2001

5. Goldhirsch A, Glick JH, Gelber RD, et al: Meeting highlights: International Expert Consensus on the primary therapy of early breast cancer 2005. Ann Oncol 16:1569, 2005

6. van de Vijver MJ, He YD, van't Veer LJ, et al: A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 347:1999-2009, 2002

7. Wang Y, Klijn JG, Zhang Y, et al: Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet 365:671-679, 2005

8. Paik S, Shak S, Tang G, et al: A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med 351: 2817-2826, 2004

9. Eppenberger-Castori S, Kueng W, Benz C, et al: Prognostic and predictive significance of ErbB-2 breast tumor levels measured by enzyme immunoassay. J Clin Oncol 19:645-656, 2001

10. Look MP, van Putten WL, Duffy MJ, 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 94:116-128, 2002

11. Blamey RW: Guidelines on endocrine therapy of breast cancer EUSOMA. Eur J Cancer 38:615-634, 2002

12. Labuhn M, Vuaroqueaux V, Fina F, et al: Simultaneous quantitative detection of relevant biomarkers in breast cancer by quantitative real-time PCR. Int J Biol Markers 21:30-39, 2006

13. Landis JR, Koch GG: The measurement of observer agreement for categorical data. Biometrics 33:159-174, 1977

14. Konecny G, Untch M, Arboleda J, et al: Her-2/neu and urokinase-type plasminogen activator and its inhibitor in breast cancer. Clin Cancer Res 7:2448-2457, 2001

15. Bouchet C, Ferrero-Pous M, Hacene K, et al: Limited prognostic value of c-erbB-2 compared to uPA and PAI-1 in primary breast carcinoma. Int J Biol Markers 18:207-217, 2003

16. Chapman HA: Plasminogen activators, integrins, and the coordinated regulation of cell adhesion and migration. Curr Opin Cell Biol 9:714-724, 1997

17. Andreasen PA, Kjoller L, Christensen L, et al: The urokinase-type plasminogen activator system in cancer metastasis: A review. Int J Cancer 72:1-22, 1997

18. Duffy MJ, Maguire TM, Hill A, et al: Metalloproteinases: Role in breast carcinogenesis, invasion and metastasis. Breast Cancer Res 2:252-257, 2000

19. Zemzoum I, Kates RE, Ross JS, et al: Invasion factors uPA/PAI-1 and HER2 status provide independent and complementary information on patient outcome in node-negative breast cancer. J Clin Oncol 21:1022-1028, 2003

20. Janicke F, Prechtl A, Thomssen C, et al: Randomized adjuvant chemotherapy trial in high-risk, lymph node-negative breast cancer patients identified by urokinase-type plasminogen activator and plasminogen activator inhibitor type 1. J Natl Cancer Inst 93:913-920, 2001

21. Harbeck N, Kates RE, Look MP, et al: Enhanced benefit from adjuvant chemotherapy in breast cancer patients classified high-risk according to urokinase-type plasminogen activator (uPA) and plasminogen activator inhibitor type 1 (n = 3424). Cancer Res 62:4617-4622, 2002

22. Vinatzer U, Dampier B, Streubel B, et al: Expression of HER2 and the coamplified genes GRB7 and MLN64 in human breast cancer: Quantitative real-time reverse transcription-PCR as a diagnostic alternative to immunohistochemistry and fluorescence in situ hybridization. Clin Cancer Res 11:8348-8357, 2005

23. Ginestier C, Charafe-Jauffret E, Penault-Llorca F, et al: Comparative multi-methodological measurement of ERBB2 status in breast cancer. J Pathol 202:286-298, 2004

24. Vanden Bempt I, Vanhentenrijk V, Drijkoningen M, et al: Real-time reverse transcription-PCR and fluorescence in-situ hybridization are complementary to understand the mechanisms involved in HER-2/neu overexpression in human breast carcinomas. Histopathology 46:431-441, 2005

25. van 't Veer LJ, Dai H, van de Vijver MJ, et al: Gene expression profiling predicts clinical outcome of breast cancer. Nature 415:530-536, 2002

26. Bosc DG, Goueli BS, Janknecht R: HER2/Neu-mediated activation of the ETS transcription factor ER81 and its target gene MMP-1. Oncogene 20: 6215-6224, 2001

27. Mazumdar A, Adam L, Boyd D, et al: Heregulin regulation of urokinase plasminogen activator and its receptor: Human breast epithelial cell invasion. Cancer Res 61:400-405, 2001

28. Gum R, Wang SW, Lengyel E, et al: Upregulation of urokinase-type plasminogen activator expression by the HER2/neu proto-oncogene. Anticancer Res 15:1167-1172, 1995

29. Allgayer H, Babic R, Gruetzner KU, et al: C-erbB-2 is of independent prognostic relevance in gastric cancer and is associated with the expression of tumor-associated protease systems. J Clin Oncol 18:2201-2209, 2000

30. O-charoenrat P, Rhys-Evans PH, Archer DJ, et al: C-erbB receptors in squamous cell carcinomas of the head and neck: Clinical significance and correlation with matrix metalloproteinases and vascular endothelial growth factors. Oral Oncol 38:73-80, 2002

31. Berney CR, Yang J, Fisher RJ, et al: Correlates of urokinase-type plasminogen activator in colorectal cancer: Positive relationship with nm23 and c-erbB-2 protein expression. Oncol Res 10:47-54, 1998

32. Umeda T, Eguchi Y, Okino K, et al: Cellular localization of urokinase-type plasminogen activator, its inhibitors, and their mRNAs in breast cancer tissues. J Pathol 183:388-397, 1997

33. Nielsen BS, Sehested M, Duun S, et al: Urokinase plasminogen activator is localized in stromal cells in ductal breast cancer. Lab Invest 81:1485-1501, 2001

Acknowledgment

We thank pathologists for providing tumor tissues in particular R. Caduff, MD, H. Moch, MD, M. Mihatsch, MD, and G. Sauter, MD, for the FISH assessment of ErbB-2. We are indebted to all clinicians for their collaboration in collecting clinical data in particular H. Dieterich, MD, D. Fink, MD, and K. Lüscher, MD. We are grateful to Thierry Sengstag, PhD, for implementing mixture model statistics, to Sabine Ehret for technical assistance, and to Ulrike Weissenstein, PhD, and Chrisitine Wullschleger for data management.

Downloaded from www.jco.org on September 19, 2006 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

Appendix

The Appendix is included in the full-text version of this article, available online at www.jco.org. It is not included in the PDF version (via Adobe® Reader®).

Authors' Disclosures of Potential Conflicts of Interest

Although all authors completed the disclosure declaration, the following authors or their immediate family members indicated a financial interest. No conflict exists for drugs or devices used in a study if they are not being evaluated as part of the investigation. For a detailed description of the disclosure categories, or for more information about ASCO's conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Authors	Employment	Leadership	Consultant	Stock	Honoraria	Research Funds	Testimony	Other
Vincent Vuaroqueaux	OncoScore AG (N/R)							
Martin Labuhn	OncoScore AG (N/R)							
Urs Eppenberger				OncoScore AG (A)				
Serenella Eppenberger-Castori				OncoScore AG (A)				
Dollar Amount Codes (A) < $10,000 (B) $10,000-99,999 (C) ≥ $100,000 (N/R) Not Required								

Author Contributions

Conception and design: Patrick Urban, Vincent Vuaroqueaux, Martin Labuhn, Urs Eppenberger, Serenella Eppenberger-Castori
Administrative support: Edward Wight, Urs Eppenberger, Serenella Eppenberger-Castori
Collection and assembly of data: Patrick Urban, Vincent Vuaroqueaux, Martin Labuhn, Edward Wight, Serenella Eppenberger-Castori
Data analysis and interpretation: Patrick Urban, Vincent Vuaroqueaux, Mauro Delorenzi, Pratyaksha Wirapati
Manuscript writing: Patrick Urban
Final approval of manuscript: Patrick Urban, Vincent Vuaroqueaux, Martin Labuhn, Mauro Delorenzi, Pratyaksha Wirapati, Edward Wight, Hans-Jörg Senn, Christopher Benz, Urs Eppenberger, Serenella Eppenberger-Castori
Other: Hans-Jörg Senn, Christopher Benz

GLOSSARY

Gene expression profiling: Identifying the expression of a set of genes in a biologic sample (eg, blood, tissue) using microarray technology.

Microarray: A miniature array of regularly spaced DNA or oligonucleotide sequences printed on a solid support at high density that is used in a hybridization assay. The sequences may be cDNAs or oligonucleotide sequences that are synthesized in situ to make a DNA chip.

MMP (matrix metalloprotease [metalloproteinases]): MMPs belong to a family of enzymes (zinc-dependent endoproteinases) that are involved in the degradation of the extracellular matrix. MMPs are involved in both normal and pathologic tissue remodeling, where their selective proteolysis is now appreciated to help regulate cell growth, angiogenesis, and invasiveness.

PAI-1 (plasminogen activator inhibitor): PAI-1 is serine protease inhibitor, and an inhibitor of uPA and tPA. Together with other members of the uPA system it is involved in extracellular matrix degradation, stimulation of cell migration and control of cell adhesion which are important for invasion and metastasis in cancer. Official gene symbol: SERPINE1[r].

Prognostic (prognostic marker): A marker that predicts the prognosis of a patient (eg, the likelihood of relapse, progression, and/or death) independent of future treatment effects. A factor can be both prognostic and predictive.

Predictive (predictive marker): Markers, biologic or molecular, that determine which treatment will increase the efficacy and improve outcome.

qrt-PCR: Quantitative polymerase chain reaction (qPCR), also known as real-time PCR, consists of detecting PCR products as they accumulate. It can be applied to gene expression quantification by reverse transcription of RNA into cDNA, thus receiving the name of quantitative reverse transcriptase polymerase chain reaction (qRT-PCR). In spite of its name—quantitative—results are usually normalized to an endogenous reference. Current devices allow the simultaneous assessment of many RNA sequences.

uPA (Urokinase-type plasminogen activator): A molecule with chemotactic activity when bound to its receptor, uPAR. Soluble uPA or uPA bound to uPAR also generates plasmin, which degrades extracellular matrix components leading to invasion and metastasis. The chemotactic activity is responsible for cell recruitment, which occurs in inflammation, neo-angiogenesis and cancer invasiveness.

uPAR: Also called CD87, uPAR is a GPI (glycophosphatidylinositol)-anchored protein that is expressed by various cells, including neutrophils, T lymphocytes, monocytes, macrophages and fibroblasts. In the absence of an intracytoplasmic region, the GPI acts as a tether, with the transmembrane adaptor(s) mediating the activation of intracellular signal transduction molecules.

Downloaded from www.jco.org on September 19, 2006 . For personal use only. No other uses without permission.
Copyright © 2006 by the American Society of Clinical Oncology. All rights reserved.

Research article

Open Access

Low E2F1 transcript levels are a strong determinant of favorable breast cancer outcome

Vincent Vuaroqueaux[1], Patrick Urban[1], Martin Labuhn[2], Mauro Delorenzi[3], Pratyaksha Wirapati[4,5], Christopher C Benz[6], Renata Flury[7], Holger Dieterich[8], Frédérique Spyratos[9], Urs Eppenberger[1] and Serenella Eppenberger-Castori[1]

[1]Stiftung Tumorbank Basel, Lörracherstrasse 50, 4125 Riehen, Switzerland
[2]OncoScore AG, Lörracherstrasse 50, 4125 Riehen, Switzerland
[3]Swiss Institute for Experimental Cancer Research and Swiss Institute of Bioinformatics, Boveresses 155, Office D208, CH-1006 Epalinges, Switzerland
[4]Chemin des Boveresses 155, CH-1066 Epalinges, Switzerland
[5]Room 2021, Genopode, UNIL Sorge, CH-1015 Lausanne, Switzerland
[6]Therapeutics Buck Institute for Age Research 8001 Redwood Blvd., Novato, CA 94945, USA
[7]Fachärztin FMH für Pathologie, Chefärztin Pathologie, Brauerstrasse 15,8401 Winterthur, Switzerland
[8]Women's Hospital and Breast Cancer Center, Therese-Herzog-Weg 2, 79618 Rheinfelden, Germany
[9]Laboratoire d'Oncogénétique/INSERM U735, Centre René Huguenin, 35 rue Dailly, 92210 St-Cloud, France

Corresponding author: Vincent Vuaroqueaux, v.vuaroqueaux@tumorbank.org

Received: 1 Mar 2007 Revisions requested: 9 Apr 2007 Revisions received: 3 May 2007 Accepted: 29 May 2007 Published: 29 May 2007

Breast Cancer Research 2007, **9**:R33 (doi:10.1186/bcr1681)
This article is online at: http://breast-cancer-research.com/content/9/3/R33
© 2007 Vuaroqueaux *et al.*: licensee BioMed Central Ltd.
This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Introduction We investigated whether mRNA levels of E2F1, a key transcription factor involved in proliferation, differentiation and apoptosis, could be used as a surrogate marker for the determination of breast cancer outcome.

Methods E2F1 and other proliferation markers were measured by quantitative RT-PCR in 317 primary breast cancer patients from the Stiftung Tumorbank Basel. Correlations to one another as well as to the estrogen receptor and ERBB2 status and clinical outcome were investigated. Results were validated and further compared with expression-based prognostic profiles using The Netherlands Cancer Institute microarray data set reported by Fan and colleagues.

Results E2F1 mRNA expression levels correlated strongly with the expression of other proliferation markers, and low values were mainly found in estrogen receptor-positive and ERBB2-negative phenotypes. Patients with low E2F1-expressing tumors were associated with favorable outcome (hazard ratio = 4.3 (95% confidence interval = 1.8–9.9), $P = 0.001$). These results were consistent in univariate and multivariate Cox analyses, and were successfully validated in The Netherlands Cancer Institute data set. Furthermore, E2F1 expression levels correlated well with the 70-gene signature displaying the ability of selecting a common subset of patients at good prognosis. Breast cancer patients' outcome was comparably predictable by E2F1 levels, by the 70-gene signature, by the intrinsic subtype gene classification, by the wound response signature and by the recurrence score.

Conclusion Assessment of E2F1 at the mRNA level in primary breast cancer is a strong determinant of breast cancer patient outcome. E2F1 expression identified patients at low risk of metastasis irrespective of the estrogen receptor and ERBB2 status, and demonstrated similar prognostic performance to different gene expression-based predictors.

Introduction

A variety of genes involved in breast cancer biology have been studied and proposed as prognostic or predictive biomarkers, but only a few of them, such as hormone receptors and ERBB2, are used today to classify breast cancer patients and to make treatment decisions in the clinical routine [1,2]. The introduction of microarray analysis recently lead to a better characterization of breast cancer on a molecular level, under-

ER = estrogen receptor; MFS = metastasis-free survival; NKI = The Netherlands Cancer Institute; PCR = polymerase chain reaction; RT = reverse transcriptase; STB = Stiftung Tumorbank Basel.

lining its biological heterogeneity and revealing that breast tumors can be grouped into different subtypes with distinct gene expression profiles and prognosis [3]. Some of these subtypes confirmed the relevance of established differences between phenotypes such as the estrogen receptor (ER) and ERBB2 status, but also identified novel breast cancer subtypes or prognostic signatures of potential clinical value [3-7]. Although little overlap was observed between these gene signatures at the level of individual genes, recent data indicate that the underlying biological processes and pathways might be common [8-10].

In terms of tumor biology, proliferation has been recognized as a distinct hallmark of cancer and as an important determinant of cancer outcome [11-13]. Increased tumor cell proliferation is accompanied by cell matrix remodeling and neo-angiogenesis, which together form the basis for an aggressive tumor phenotype [14,15]. This observation was further underlined by recent reports showing that several genes involved in gene signatures discriminating clinically relevant breast cancer subtypes were related to proliferation [3,4,9,16,17].

In the context of breast cancer molecular screening, we recently investigated by quantitative RT-PCR the expression of 60 tumor-related genes in various subsets of breast cancers from the Stiftung Tumorbank Basel (STB) [18,19]. This gene set also comprised several genes involved in proliferation such as thymidilate synthase (TYMS), thymidine kinase 1 (TK1), topoisomerase 2-alpha (TOP2A), survivin (BIRC5) and the transcription factor E2F1. Since these genes strongly correlated to one another and since the assessment of a single gene able to accurately predict breast cancer patients' outcome would represent major advantages for standard clinical use, we focused our efforts on the evaluation of E2F1 transcript levels as surrogate marker for proliferation. This transcription factor is well known for being involved in the cyclin/cyclin-dependent kinase/retinoblastoma pathway and for controlling the expression of more than 1,000 genes involved in cell proliferation, differentiation and apoptosis [20-23]. In a set of 317 primary breast cancers patients with known clinical outcome (STB data set), we evaluated E2F1 mRNA expression levels with respect to other proliferation markers, ER and ERBB2 status and clinical outcome. All results obtained in our collective were subsequently validated in The Netherlands Cancer Institute (NKI) microarray data set comprising 295 breast cancer patients. Moreover, the prognostic value of E2F1 was compared with the 70-gene prognostic signature, and with other gene expression-based predictors such as the intrinsic subtypes, the wound response signature and the recurrence score available as reported by Fan and colleagues using the same NKI data set [8].

Methods

Study populations

Patients and methods have been described previously [18]. The 317 primary breast cancer tissue samples were obtained from the STB, Switzerland and were analyzed by quantitative RT-PCR (STB data set). The previously published microarray breast cancer data set reported by Van de Vijver and colleagues (NKI data set) [5] was used for validation and comparative analysis as reported by Fan and colleagues [8]. Major differences between the two study populations included the patient age, nodal status, adjuvant therapy and methodology (quantitative RT-PCR versus Agilent microarray). Detailed patient and tumor characteristics are summarized in Table 1.

Quantitative real-time PCR analysis

Gene expression measurements by quantitative RT-PCR were performed as reported previously [24]. Total RNA was extracted using the RNAeasy Mini Kit (Qiagen, Hilden, Germany) and was quality-checked on a Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA). High-quality RNA samples were reverse-transcribed and PCR was carried out in 40 cycles on a ABI Prism 7000 using 2× SYBR Green I Master Mix (Applied Biosystems, Forster City, CA, USA). Relative gene expression quantities (Δ[Ct] values) were obtained by normalization against ribosomal 18S RNA.

Statistical analysis

For the STB study the ER status was defined based on the mRNA level as reported previously [24], and for the NKI data set the status was defined as provided by the authors [5,8]. The ERBB2 status was determined in both the STB and NKI data sets using mRNA expression levels for all study populations as previously described by Urban and colleagues [18].

The prognostic value of biomarkers was assessed by univariate and multivariate Cox analysis against metastasis-free survival (MFS), and in different patient subgroups according to the ER and ERBB2 status. The association of E2F1 with MFS in particular was assessed by univariate Cox analysis for various cutoff values (data not shown). For all subsequent analysis, the 30th percentile was used as the cutoff point for E2F1. Survival probabilities for MFS were calculated according to the Kaplan–Meier method, and group differences were assessed by the logrank test. Multivariate *P* values were based on Wald statistics. Statistical analysis was performed with 'R' statistical software version 2.0.1 using the 'survival' package [25].

Results

E2F1 correlated with other proliferation markers and clinical outcome

A strong and significant correlation was found between the five proliferation markers analyzed in the STB data set (Table 2). Univariate Cox regression analysis demonstrated a significant association of E2F1 as well as TYMS, TK1, TOP2A and

Table 1

Patient and tumor characteristics

Characteristic	Stiftung Tumorbank Basel data set	The Netherlands Cancer Institute data set
Method	Quantitative RT-PCR	Agilent Microarray
n	317	295
Age		
Mean/median (years)	60/59	44/44
≤ 40 years	20 (6%)	75 (25%)
41–55 years	110 (35%)	220 (75%)
≥ 56 years	187 (59%)	0 (0%)
pT stage		
pT1	100 (32%)	155 (53%)
pT2	183 (58%)	140 (47%)
pT3/4	33 (10%)	0
pN status		
Negative	161 (54%)	151 (51%)
Positive	136 (46%)	144 (49%)
Unknown	20	0
Histological grade		
1 (good)	28 (9%)	75 (25%)
2 (intermediate)	137 (46%)	101 (34%)
3 (poor)	133 (45%)	119 (41%)
Unknown	19	0
Estrogen receptor status[a]		
Positive	231 (73%)	226 (77%)
Negative	86 (27%)	69 (23%)
ErbB2 status		
Positive	70 (22%)	52 (18%)
Negative	247 (78%)	243 (82%)
Adjuvant therapy		
None	60 (20%)	165 (56%)
Hormone	135 (44%)	20 (7%)
Chemotherapy	72 (24%)	90 (30%)
Combination	38 (12%)	20 (7%)
Total	245 (80%)	130 (44%)
Unknown	12	0
Follow-up		
Events (metastases)	57 (18%)	101 (34%)
Mean/median metastasis-free survival (years)	3.7/3.6	7.3/6.8

Data presented as *n* (%) unless stated otherwise. [a]Estrogen receptor positive, ≥ 20 fmol/mg protein (enzyme immunoassay) for the Stiftung Tumorbank Basel data set; for The Netherlands Cancer Institute study, see [5,6]

Table 2

Correlation among different proliferation markers in the Stiftung Tumorbank Basel data set and association with survival

	Correlation[a]				Univariate Cox regression[b] (*P* value)
	E2F1	BIRC5	TOP2A	TK1	
E2F1	-				<0.001
BIRC5	0.84				0.001
TOP2A	0.78	0.76			<0.001
TK1	0.79	0.88	0.67		0.018
TYMS	0.81	0.80	0.71	0.77	0.005

[a]Pearson correlation coefficient, all $P < 0.05$. [b]Metastasis-free survival.

BIRC5 expression levels with distant MFS (Table 2). Similar results were observed in the NKI data set (data not shown). In the NKI data set we also investigated Ki67. The RNA expression levels of this proliferation marker were positively correlated with E2F1 (correlation coefficient = 0.46) and were borderline significant ($P = 0.02$) in univariate Cox regression analysis.

Distinct E2F1 expression patterns according to ER and ERBB2 status determined the clinical outcome

Scatter plots of E2F1 versus ER and ERBB2 expression levels in the STB data set (Figure 1a,b) revealed that ER-negative and ERBB2-positive breast tumors typically expressed high levels of E2F1, whereas in contrast low E2F1 levels (below the 30th percentile of its distribution in this collective) were detected almost exclusively in ER-positive and ERBB2-negative breast tumors. The same pattern was observed in the NKI data set (Figure 1c,d). Similar scatter plots were obtained analyzing the other proliferation markers (data not shown).

Cox univariate survival analysis performed in subsets of patients according to their ER and ERBB2 status showed that E2F1 correlated with MFS in ER-positive and ERBB2-negative tumors, but not in ER-negative and ERBB2-positive tumors (data not shown). Combined Kaplan–Meier analysis using E2F1 and the ER or ERBB2 status revealed that patients whose tumors expressed low E2F1 levels, a situation found mainly in ER-positive and ERBB2-negative phenotypes, were associated with favorable outcome, whereas patients with tumors expressing high E2F1 levels revealed a poor outcome independent of the ER and ERBB2 status (Figure 1e-h).

E2F1 correlated well with the 70-gene signature

The majority of the patients in the NKI data set assigned to the good-prognosis group by the 70-gene signature expressed low E2F1 levels and were found to be ER-positive or ERBB2-negative (Figure 2a,b). In addition, there was a strong correlation ($r = 0.67$) between E2F1 and the 70-gene signature (Figure 2c). In particular, 77% (69 out of 90) of patients with low E2F1-expressing tumors overlapped with patients assigned to the good-prognosis group by the 70-gene signature and were indeed found to be at the lowest risk of metastatic events. Patients with low E2F1 and a poor-prognosis signature or patients with high E2F1 and a good-prognosis signature had a comparable incidence of metastases (Table 3).

E2F1 stratification showed similar prognostic value as the 70-gene and other gene-based predictors

Kaplan–Meier analysis displayed the similar prognostic value of E2F1 and the 70-gene signature (hazard ratio = 5.1 (95% confidence interval = 2.7–9.8) and hazard ratio = 4.6 (95% confidence interval = 2.7–7.8), respectively; Figure 3a). We obtained similar results (Figure 3b–d) when E2F1 levels were compared with the breast cancer intrinsic subtypes [3], with the recurrence score [17] and with the wound response signature [7], all of these gene expression-based predictors being reported by Fan and colleagues in the NKI data set [8].

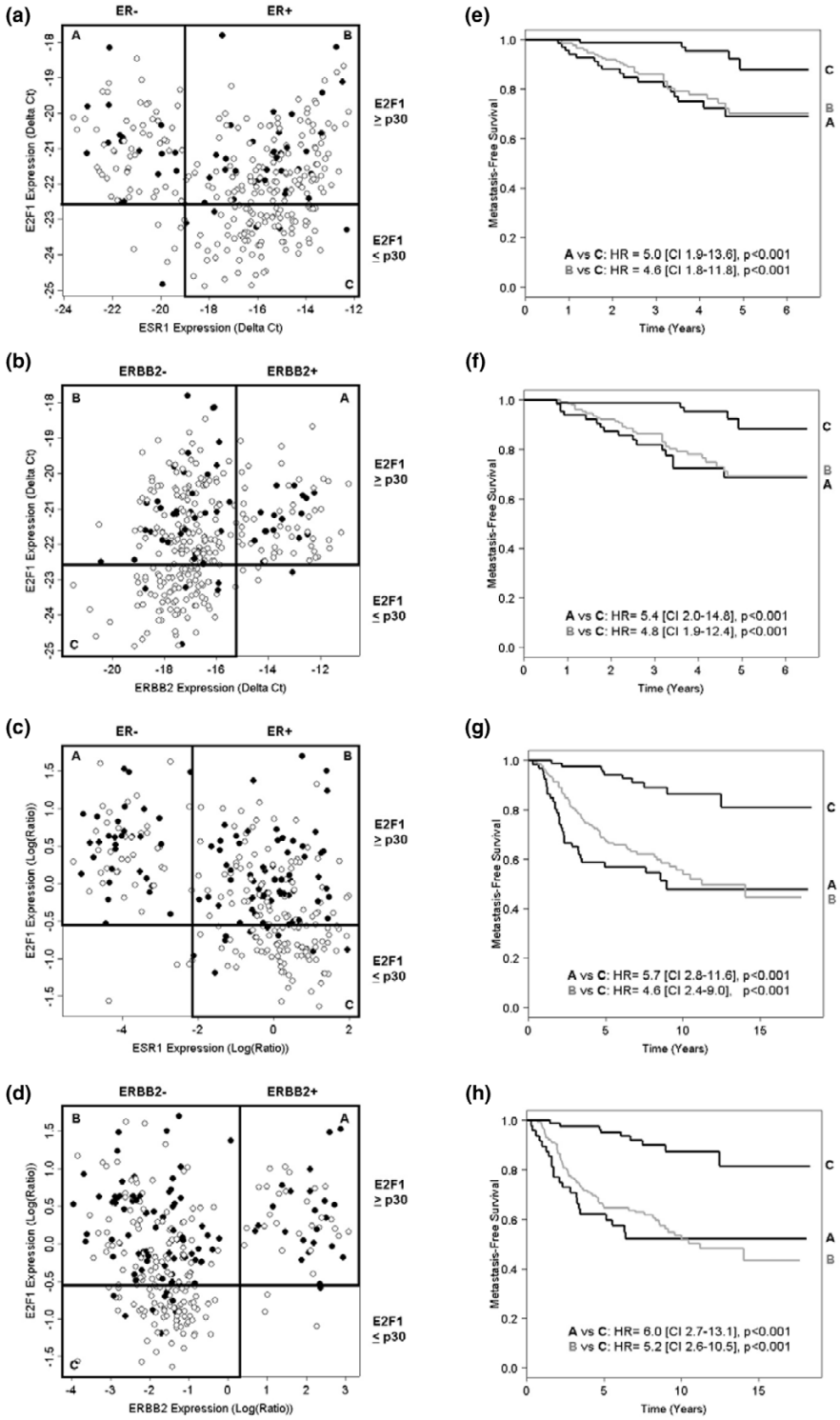
E2F1 was a strong and independent survival factor in multivariate analysis

Multivariate survival analysis including E2F1, nodal status, grade, tumor size, age, ER and ERBB2 status, and treatments revealed that only E2F1 and nodal status retained independent prognostic value in the STB data set (Table 4), and that E2F1, nodal status, tumor size, age and chemotherapy were significant in the NKI data set (Table 5). We performed a second multivariate Cox model including additionally the 70-gene signature in the NKI data set (Table 5), reconfirming that E2F1 and the 70-gene signature were significant and additive predictive survival factors together with the nodal status, tumor size and chemotherapy.

Discussion

In the present study we demonstrated that the assessment of E2F1 mRNA as a surrogate proliferation marker is a strong determinant of breast cancer outcome, particularly suitable for identifying patients at very low risk of metastasis, comparable with gene expression-based signatures such as the 70-gene signature. The prognostic component of the ER and ERBB2 status as well as different gene signatures were found to be

Figure 1



Estrogen receptor and ERBB2 versus E2F1 expression levels. Scatter plots of estrogen receptor (ER) ESR1 and ERBB2 versus E2F1 expression levels in **(a), (b)** the Stiftung Tumorbank Basel data (STB) set and **(c), (d)** The Netherlands Cancer Institute (NKI) data set. Open circles, no metastasis; filled circles, metastasis. Vertical lines, cutoff values for the estrogen receptor (ER) and ERBB2 status, respectively; horizontal lines, 30th percentile for E2F1. Combined Kaplan–Meier analysis (metastasis-free survival) using the ER or ERBB2 status and E2F1 (30th percentile) in **(e), (f)** the STB data set and **(g), (h)** the NKI data set. Labels of the survival curves correspond to the groups as indicated on the respective scatter plot. CI, 95% confidence interval; HR, hazard ratio.

Table 3

Concordance of E2F1 with the 70-gene signature in The Netherlands Cancer Institute data set

Proliferation status (E2F1)	NKI 70-gene signature prognosis		
	Good	Poor	Total
E2F1 ≤ p30	4/69 (5.7%)	6/21 (28.6%)	10/90 (11.1%)
E2F1 > p30	12/46 (26.1%)	79/159 (49.6%)	91/205 (44.4%)
Total	16/115 (13.9%)	85/180 (47.2%)	101/295 (34.2%)

The percentage indicates the number of metastatic events over the number of cases in each group.

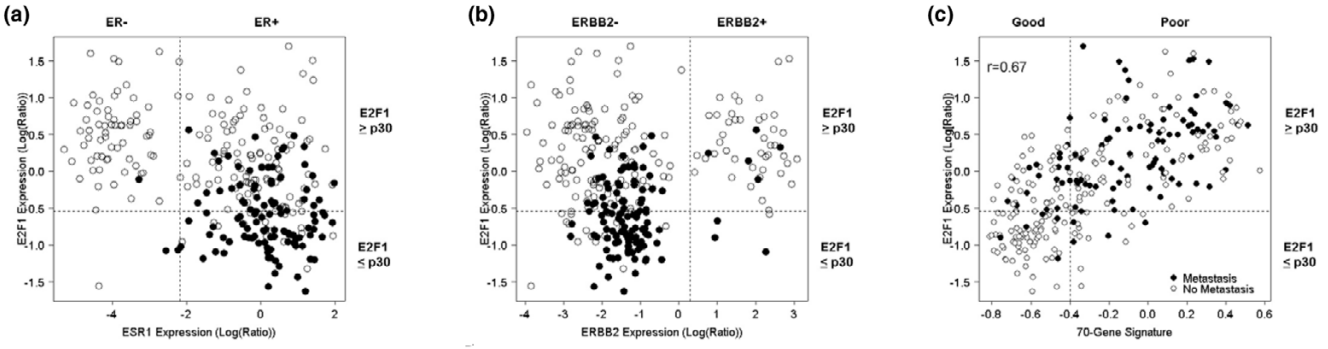
strongly related to tumor proliferation. In fact, a large subset of patients classified with very favorable outcome shared a common molecular tumor phenotype characterized by ER-positive and/or ERBB2-negative status and low proliferation (low levels of E2F1 as well as *BIRC5*,*TYMS*,*TOP2A* and *TK1*). Moreover, the results obtained in our data set analyzed by quantitative RT-PCR were successfully validated in an independent breast cancer data set using microarray technology.

Sotiriou and colleagues developed a gene expression grade index able to reclassify breast cancer patients with tumor histological grade 2 into groups with high risk of recurrence versus low risk [9]. The gene expression grade index was developed on the basis of the analysis of five breast cancer microarray data sets including more than 600 tumors, from which the authors extracted a list of 242 genes associated with tumor grade and predicting patient outcome. Most of these genes were related to proliferation and cell survival, such as E2F1 and MKI67, BIRC5, TOP2A and STK6, all being highly correlated and providing similar prognostic information. In our study, we demonstrated that the detection of a single gene is sufficient to select tumors at low proliferation. A single gene assessment requires high RNA quality from fresh (frozen) tissue, however, and might be insufficient in cases of more heterogeneous RNA quality (for example, RNA from paraffin-embedded tissues).

Breast cancer has been successfully classified using microarrays into clinically relevant subgroups based on variations in gene expression patterns. Sorlie and colleagues showed that ER-negative tumors grouped into basal-like and ERBB2 subtypes, both with poor prognosis [3]. In contrast, ER-positive breast cancers could be classified into luminal A and luminal B subtypes with significantly distinct prognosis: luminal A tumors displayed favorable outcome, whereas survival of patients with luminal B tumors was poor and comparable with those of the ER-negative ERBB2 and basal subtypes [3]. Our classification in the NKI data set revealed that 81% of the tumors expressing low E2F1 levels (below this study's cutoff point) corresponded with luminal A subtype as defined by Fan and colleagues [8], and subsequently had similar prognostic value (Figure 3b).

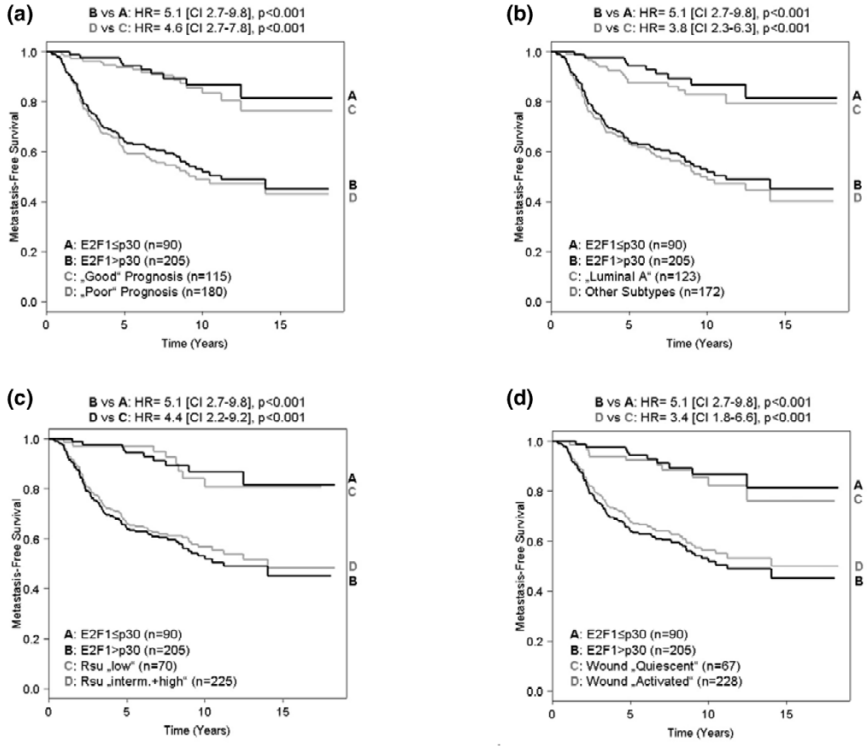
Van de Vijver and colleagues used a 70-gene prognostic signature to discriminate patients with good prognosis and poor prognosis [5], which according to our analysis strongly correlated with E2F1 expression levels. As shown in Figure 2, patients defined as of good prognosis by the 70-gene signa-

Figure 2



Comparison of E2F1 and the 70-gene signature in The Netherlands Cancer Institute data set. **(a), (b)** Scatter plots of estrogen receptor (ER = ESR1) and ERBB2 versus E2F1 expression levels. Open circles, poor-prognosis group as defined by [5]; filled circles, good-prognosis group [5]. **(c)** Correlation between the 70-gene prognostic signature and E2F1. Open circles, no metastasis; filled circles, metastasis.

Figure 3



Kaplan–Meier analysis of metastasis-free survival. Kaplan–Meier analysis (metastasis-free survival) using **(a)** E2F1 expression (30th percentile) and the 70-gene signature, **(B)** intrinsic subtypes, **(c)** the recurrence score (Rsu), and **(b)** the wound response signature. CI, 95% confidence interval; HR, hazard ratio.

ture had tumors expressing low E2F1 levels and were mainly ER-positive. Despite all observed correlations, multivariate Cox analysis of the NKI data set showed that E2F1 levels and the 70-gene prognostic signature retained additive significance when both covariates were included (Table 5). This is probably due to the fact that both markers classified, in addition to the overlapping patients at very low risk, patients at similar but higher risk who would not have been selected by either classifier alone (Table 3). Furthermore, we found that almost all ERBB2-positive and ER-negative tumors expressed high levels of E2F1 and were classified as of poor prognosis according to the 70-gene signature – suggesting an explanation of why Espinosa and colleagues were unsuccessful in improving the accuracy of the 70-gene signature by incorporating additional genes such as ERBB2 [26].

Fan and colleagues [8] recently demonstrated that the different gene-expression-based predictors including the 70 gene-signature, the intrinsic subtypes, the wound signature and the recurrence score were highly concordant to evaluate breast cancer outcome. Our analysis revealed that low proliferation as quantified by low levels of E2F1 represented a common determinant of patients with good prognosis (Figures 2 and 3). It has to be noted that the prognostic value of E2F1 was independent of the nodal status. Indeed, 40% of the STB tumors and 50% of the NKI tumors with low E2F1 expression levels belonged to nodal-positive patients at very low risk of metastases, reconfirming the impact of proliferation recently reported in a study evaluating breast cancer patients with 10 and more positive lymph nodes [27,28].

The STB and NKI data sets differed in adjuvant treatment modalities; in general, patients of the STB collective were older and consequently received more hormone therapy but less chemotherapy as compared with patients of the NKI collective. In this context, it has to be emphasized that treatment regiments were chosen independent of the E2F1 status (Additional file 1) and that E2F1 levels retained predictive survival significance in patients with and without different adjuvant treatments (Additional file 2). Multivariate analyses, however, revealed different treatment impacts in the two data sets (Tables 4 and 5). In the STB collective, chemotherapy was particularly significant in univariate Cox analysis but was nonsignificant in multivariate Cox models, suggesting that information about the higher risk cases receiving chemotherapy is already included in the combination of the other covariates. Since E2F1 is co-expressed or regulates genes such as TYMS, TK1 and TOP2A, which were mechanistically linked with response to 5-fluorouracil and anthracycline-based therapy [16,29-32], however, our results with respect to specific chemotherapy response should be further investigated.

Table 4

Univariate and multivariate Cox analyses in the Stiftung Tumorbank Basel data set ($n = 317$)

Factor	Univariate metastasis-free survival		Multivariate metastasis-free survival	
	Hazard ratio (95% confidence interval)	*P* value	Hazard ratio (95% confidence interval)	*P* value
E2F1 (>p30 versus ≤ p30)	4.27 (1.83–9.96)	0.001	2.95 (1.10–7.93)	0.032
Grade (3 versus 1 + 2)	2.03 (1.17–3.51)	0.011	1.56 (0.77–3.14)	0.213
Estrogen receptor status (positive versus negative)	0.64 (0.37–1.12)	0.120	1.21 (0.56–2.61)	0.625
ERBB2 status (positive versus negative)	1.69 (0.98–2.92)	0.058	1.41 (0.70–2.81)	0.335
pN (>3 nodes versus ≤ 3 nodes)	3.14 (1.76–5.61)	<0.001	2.36 (1.14–4.88)	0.021
Size (>2 cm versus ≤ 2 cm)	1.95 (1.05–3.62)	0.036	1.19 (0.58–2.43)	0.641
Age (≤ 40 years versus >40 years)	0.30 (0.15–0.59)	0.001	0.59 (0.24–1.43)	0.246
Chemotherapy	2.65 (1.54–4.55)	<0.001	1.27 (0.45–3.59)	0.654
Hormone therapy	0.50 (0.28–0.88)	0.017	0.76 (0.27–2.17)	0.605
70-gene signature (poor versus good prognosis)	Not available	Not available	Not available	Not available

Table 5

Univariate and multivariate Cox analyses in The Netherlands Cancer Institute data set ($n = 295$)

Factor	Univariate metastasis-free survival		Multivariate metastasis-free survival[a]			
	Hazard ratio (95% confidence interval)	*P* value	With 70-gene signature		Without 70-gene signature	
			Hazard ratio (95% confidence interval)	*P* value	Hazard ratio (95% confidence interval)	*P* value
E2F1 (>p30 versus ≤ p30)	5.09 (2.65–9.78)	<0.001	3.76 (1.90–7.45)	<0.001	2.47 (1.20–5.10)	0.014
Grade (3 versus 1 + 2)	2.38 (1.60–3.52)	<0.001	1.28 (0.82–2.00)	0.277	1.02 (0.65–1.60)	0.931
Estrogen receptor status (positive versus negative)	0.54 (0.36–0.83)	0.005	0.99 (0.61–1.59)	0.956	1.11 (0.70–1.78)	0.658
ERBB2 status (positive versus negative)	1.61 (1.01–2.57)	0.045	1.35 (0.82–2.21)	0.234	1.28 (0.78–2.09)	0.330
pN (>3 nodes versus ≤ 3 nodes)	2.20 (1.37–3.53)	0.001	2.35(1.32–4.21)	0.004	2.69 (1.47–4.91)	0.001
Size (>2 cm versus ≤ 2 cm)	2.08 (1.39–3.10)	<0.001	1.70 (1.12–2.58)	0.013	1.73 (1.14–2.62)	0.010
Age (≤ 40 years versus >40 years)	0.50 (0.33–0.75)	0.001	0.59 (0.38–0.89)	0.013	0.67 (0.43–1.02)	0.063
Chemotherapy	0.79 (0.52–1.19)	0.254	0.61 (0.38–1.00)	0.051	0.55 (0.33–0.91)	0.020
Hormone therapy	0.58 (0.28–1.19)	0.139	0.60 (0.28–1.27)	0.181	0.58 (0.27–1.23)	0.157
70-gene signature (poor versus good prognosis)	4.55 (2.67–7.77)	<0.001	Not included	Not included	2.78 (1.49–5.21)	0.001

[a]Multivariate metastasis-free survival calculated once with and once without the 70-gene signature.

Conclusion

Since accurate monitoring of proliferation assessing the mRNA E2F1 levels together with the determination of the ER and ERBB2 status can be performed easily by quantitative RT-PCR even in small amounts of tissue such as core biopsies [19], we encourage the inclusion of such analyses in protocols of ongoing clinical and translational research investigations, including predictive studies with respect to specific chemotherapies.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

VV, ML and SE-C designed the study. VV and ML contributed to the selection of the genes, selected primers, and supervised the RNA extraction and quantitative RT-PCR. VV and PU performed statistical analysis under the supervision of MD and PW. VV, PU, MD, PW and SE-C contributed to data interpretation. RF performed the pathological analysis of several samples and asserved the surgical samples for molecular analysis. CCB, HD, RF, FS and UE participated in designing the study and writing the manuscript. VV, PU and SE drafted the manuscript. All authors read and approved the final manuscript.

Additional files

The following Additional files are available online:

Additional File 1

A Word file containing a table presenting the treatment distribution according to the E2F1 status in both data sets.
See http://www.biomedcentral.com/content/supplementary/bcr1681-S1.doc

Additional File 2

A pdf file containing a figure showing Kaplan–Meier analysis (MFS) using E2F1 (30th percentile) performed in data subsets with defined adjuvant treatments: (a) none, (b) hormone, (c) chemotherapy and (d) combined.
See http://www.biomedcentral.com/content/supplementary/bcr1681-S2.pdf

Acknowledgements

VV and PU contributed equally to the work. This work was supported by the Stiftung Tumor Bank Basel, the Swiss National Foundation (Grant 3100-059819.99/1) and the NCCR Molecular Oncology of the Swiss National Science Foundation. OncoScore AG paid the costs for quantitative RT-PCR analysis. The authors are very grateful to Sabine Ehret for technical support and data management. They thank pathologists for providing tumor tissues, in particular Prof. H. Moch, Prof. M. Mihatsch and Prof. W. Wegmann. The authors are indebted to all clinicians for their collaboration in collecting clinical data, especially to Prof. D. Fink, PD Dr E. Wight and Dr K. Lüscher.

References

1. Esteva FJ, Hortobagyi GN: **Prognostic molecular markers in early breast cancer.** *Breast Cancer Res* 2004, **6:**109-118.
2. Goldhirsch A, Glick JH, Gelber RD, Coates AS, Thurlimann B, Senn HJ: **Meeting highlights: International Expert Consensus on the Primary Therapy of Early Breast Cancer 2005.** *Ann Oncol* 2005, **16:**1569-83.
3. 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 USA* 2001, **98:**10869-10874.
4. 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:**530-536.
5. 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:**1999-2009.
6. 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:**671-679.
7. Chang HY, Sneddon JB, Alizadeh AA, Sood R, West RB, Montgomery K, Chi JT, van de Rijn M, Botstein D, Brown PO: **Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds.** *PLoS Biol* 2004, **2:**E7.
8. Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DS, Nobel AB, van't Veer LJ, Perou CM: **Concordance among gene-expression-based predictors for breast cancer.** *N Engl J Med* 2006, **355:**560-569.
9. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, Nordgren H, Farmer P, Praz V, Haibe-Kains B, *et al.*: **Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis.** *J Natl Cancer Inst* 2006, **98:**262-272.
10. Brenton JD, Carey LA, Ahmed AA, Caldas C: **Molecular classification and molecular forecasting of breast cancer: ready for clinical application?** *J Clin Oncol* 2005, **23:**7350-7360.
11. Hanahan D, Weinberg RA: **The hallmarks of cancer.** *Cell* 2000, **100:**57-70.
12. van Diest PJ, van der Wall E, Baak JP: **Prognostic value of proliferation in invasive breast cancer: a review.** *J Clin Pathol* 2004, **57:**675-681.
13. Desmedt C, Sotiriou C: **Proliferation: the most prominent predictor of clinical outcome in breast cancer.** *Cell Cycle* 2006, **5:**2198-2202.
14. Ellis LM, Fidler IJ: **Angiogenesis and metastasis.** *Eur J Cancer* 1996, **32A:**2451-2460.
15. Eppenberger U, Kueng W, Schlaeppi 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:**3129-3136.
16. Dai H, van't Veer L, Lamb J, He YD, Mao M, Fine BM, Bernards R, van de Vijver M, Deutsch P, Sachs A, *et al.*: **A cell proliferation signature is a marker of extremely poor outcome in a subpopulation of breast cancer patients.** *Cancer Res* 2005, **65:**4059-4066.
17. 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:**2817-2826.
18. Urban P, Vuaroqueaux V, Delorenzi M, Wirapati P, Labuhn M, Benz C, Wight E, Senn HJ, Eppenberger U, Eppenberger-Castori S: **Increased expression of urokinase-type plasminogen activator (uPA) mRNA determines adverse prognosis in ErbB2-positive breast cencer patients.** *J Clin Oncol* 2006, **24:**4245-53.
19. Zanetti-Dallenbach R, Vuaroqueaux V, Wight E, Labuhn M, Singer G, Urban P, Eppenberger U, Holzgreve W, Eppenberger-Castori S: **Comparison of gene expression profiles in core biopsies and corresponding surgical breast cancer samples.** *Breast Cancer Res* 2006, **8:**R51.
20. Muller H, Bracken AP, Vernell R, Moroni MC, Christians F, Grassilli E, Prosperini E, Vigo E, Oliner JD, Helin K: **E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis.** *Genes Dev* 2001, **15:**267-285.
21. Bell LA, Ryan KM: **Life and death decisions by E2F-1.** *Cell Death Differ* 2004, **11:**137-142.
22. Zhang SY, Liu SC, Al-Saleem LF, Holloran D, Babb J, Guo X, Klein-Szanto AJ: **E2F-1: a proliferative marker of breast neoplasia.** *Cancer Epidemiol Biomarkers Prev* 2000, **9:**395-401.
23. Baldini E, Camerini A, Sgambato A, Prochilo T, Capodanno A, Pasqualetti F, Orlandini C, Resta L, Bevilacqua G, Collecchi P: **Cyclin A and E2F1 overexpression correlate with reduced disease-free survival in node-negative breast cancer patients.** *Anticancer Res* 2006, **26:**4415-4421.

24. Labuhn M, Vuaroqueaux V, Fina F, Schaller A, Nanni-Metellus I, Kung W, Eppenberger-Castori S, Martin PM, Eppenberger U: **Simultaneous quantitative detection of relevant biomarkers in breast cancer by quantitative real-time PCR.** *Int J Biol Markers* 2006, **21:**30-39.
25. **'R' Statistical Software** [http://www.r-project.org]
26. Espinosa E, Vara JA, Redondo A, Sanchez JJ, Hardisson D, Zamora P, Pastrana FG, Cejas P, Martinez B, Suarez A, *et al.*: **Breast cancer prognosis determined by gene expression profiling: a quantitative reverse transcriptase polymerase chain reaction study.** *J Clin Oncol* 2005, **23:**7278-7285.
27. Cobleigh MA, Tabesh B, Bitterman P, Baker J, Cronin M, Liu ML, Borchik R, Mosquera JM, Walker MG, Shak S: **Tumor gene expression and prognosis in breast cancer patients with 10 or more positive lymph nodes.** *Clin Cancer Res* 2005, **11:**8623-8631.
28. Miyoshi Y, Ando A, Egawa C, Taguchi T, Tamaki Y, Tamaki H, Sugiyama H, Noguchi S: **High expression of Wilms' tumor suppressor gene predicts poor prognosis in breast cancer patients.** *Clin Cancer Res* 2002, **8:**1167-1171.
29. Longley DB, Harkin DP, Johnston PG: **5-Fluorouracil: mechanisms of action and clinical strategies.** *Nat Rev Cancer* 2003, **3:**330-338.
30. Di Leo A, Cardoso F, Durbecq V, Giuliani R, Mano M, Atalay G, Larsimont D, Sotiriou C, Biganzoli L, Piccart MJ: **Predictive molecular markers in the adjuvant therapy of breast cancer: state of the art in the year 2002.** *Int J Clin Oncol* 2002, **7:**245-253.
31. Gianni L, Zambetti M, Clark K, Baker J, Cronin M, Wu J, Mariani G, Rodriguez J, Carcangiu M, Watson D, *et al.*: **Gene expression profiles in paraffin-embedded core biopsy tissue predict response to chemotherapy in women with locally advanced breast cancer.** *J Clin Oncol* 2005, **23:**7265-7277.
32. Rosty C, Sheffer M, Tsafrir D, Stransky N, Tsafrir I, Peter M, de Cremoux P, de La Rochefordiere A, Salmon R, Dorval T, *et al.*: **Identification of a proliferation gene cluster associated with HPV E6/E7 expression level and viral DNA load in invasive cervical carcinoma.** *Oncogene* 2005, **24:**7094-7104.

Research article

Open Access

Comparison of gene expression profiles in core biopsies and corresponding surgical breast cancer samples

Rosanna Zanetti-Dällenbach[1], Vincent Vuaroqueaux[2,3], Edward Wight[1], Martin Labuhn[3], Gad Singer[4], Patrick Urban[2], Urs Eppenberger[2], Wolfgang Holzgreve[1] and Serenella Eppenberger-Castori[2]

[1]Women's University Hospital Basel, Switzerland
[2]Stiftung Tumorbank Basel, Switzerland
[3]OncoScore AG, Riehen, Switzerland
[4]Department of Pathology, University Hospital Basel, Switzerland

Corresponding author: Serenella Eppenberger-Castori, s.eppenberger@tumorbank.org

Received: 14 Apr 2006 Revisions requested: 5 Jun 2006 Revisions received: 25 Jun 2006 Accepted: 18 Aug 2006 Published: 18 Aug 2006

Breast Cancer Research 2006, **8**:R51 (doi:10.1186/bcr1542)
This article is online at: http://breast-cancer-research.com/content/8/4/R51
© 2006 Zanetti-Dällenbach *et al.*; licensee BioMed Central Ltd.
This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Introduction Gene expression profiling has been successfully used to classify breast cancer into clinically distinct subtypes, and to predict the risk of recurrence and treatment response. The aim of this study was to investigate whether the gene expression profile (GEP) detected in a core biopsy (CB) is representative for the entire tumor, since CB is an important tool in breast cancer diagnosis. Moreover, we investigated whether performing CBs prior to the surgical excision could influence the GEP of the respective tumor.

Methods We quantified the RNA expression of 60 relevant genes by quantitative real-time PCR in paired CBs and surgical specimens from 22 untreated primary breast cancer patients. Subsequently, expression data were compared with independent GEPs obtained from tumors of 317 patients without preceding CB.

Results In 82% of the cases the GEP detected in the CB correlated very well with the corresponding profile in the surgical sample ($r_s \geq 0.95$, $p < 0.001$). Gene-by-gene analysis revealed four genes significantly elevated in the surgical sample compared to the CB; these comprised genes mainly involved in inflammation and the wound repair process as well as in tumor invasion and metastasis.

Conclusion A GEP detected in a CB are representative for the entire tumor and is, therefore, of clinical relevance. The observed alterations of individual genes after performance of CB deserve attention since they might impact the clinical interpretation with respect to prognosis and therapy prediction of the GEP as detected in the surgical specimen following CB performance.

Introduction

Gene expression profiling by parallel detection of thousands of genes permits the molecular signature (phenotype) of a tissue sample to be read and can, therefore, individually characterize a patient's tumor at the molecular level. Based on the gene expression profile (GEP) of a tumor, a molecular classification for breast cancer was proposed [1] and several molecular signatures were reported to predict the risk of recurrence and treatment response [2-5]. Such molecular analyses require only small amounts of material, such as tissue samples obtained by minimal invasive methods, for example, core biopsy (CB), which are used to assess the nature of palpable and non-palpable breast lesions to confirm or exclude the diagnosis of breast cancer [6-9].

Although CB investigations have become more and more important in the early workup of breast lesions, there are only a few investigations regarding the reliability of GEPs as detected in CBs [10]. However, wound healing subsequent to CB can potentially induce gene expression alterations in the injured tissue. Many of these normally occurring reparative

CB = core biopsy; COX = cyclooxygenase; Ct = cycle threshold; GEP = gene expression profile; MMP = matrix metalloproteinase; qrt-PCR = quantitative real-time PCR; PAI = plasminogen activator inhibitor; PAI-1 = Plasminogen Activator Inhibitor-1; ST = surgical tumor tissue specimen; uPAR = urokinase plasminogen activator receptor; VEGF = Vascular Endothelial Growth Factor A.

Table 1

Clinicopathological characteristics and gene expression variation of each patient

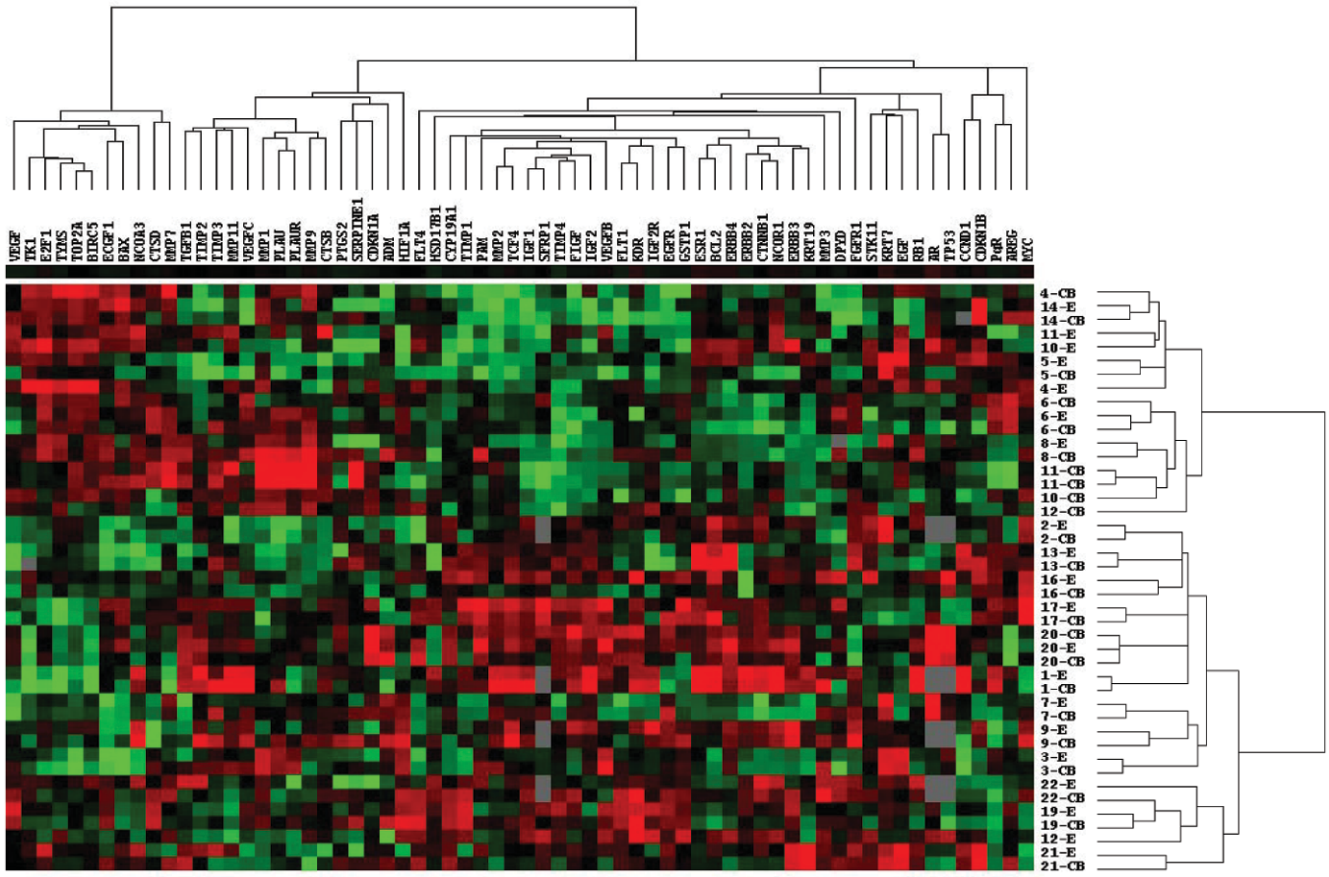
								Difference in gene expression between CB and ST (ΔCt [$Gene_{ST}$] – ΔCt [$Gene_{CB}$])		
Patient number	Age	Time between CB and ST (days)	Histology	Tumor diameter (cm)	Grade	Lymph node involvement	Correlation (r_s) between paired CB and ST	PAI-1	COX-2	ERBB2
1	68	7	Ductal	3.3	3	Negative	0.97	-0.39	-0.15	-0.08
2	59	8	Ductal	1.8	3	Negative	0.98	-1.68	-0.20	0.52
3	49	21	Ductal	3.3	2	Positive	0.98	-1.19	-1.03	0.62
4	57	13	Ductal	1.1	1	Positive	0.89	-0.63	0.61	-0.02
5	63	14	Ductal	9.0	2	Negative	0.97	-1.41	-1.80	1.40
6	60	21	Ductulo-lobular	2.7	1	Positive	0.98	-1.38	-2.41	0.23
7	80	12	Ductal	3.1	1	Positive	0.95	-2.62	-1.39	0.44
8	81	14	Lobular	2.8	2	Positive	0.94	-3.78	-3.73	0.29
9	34	14	Ductal	1.8	2	Negative	0.95	-3.16	-1.37	0.12
10	68	10	Ductal	0.9	2	Micrometastasis	0.87	-3.57	-7.27	1.01
11	65	16	Lobular	1.9	1	Negative	0.89	-5.16	-0.80	-0.02
12	80	13	Ductal	6.0	3	Positive	0.86	-2.16	-0.20	0.63
13	59	2	Apocrine	2.7	3	Negative	0.98	-1.18	-4.64	-0.27
14	73	12	Lobular	1.9	2	Positive	0.95	-4.60	-3.97	-1.85
15	61	1	Lobular	2.3	2	Positive	0.89	-1.88	-1.61	-0.94
16	73	22	Ductal	1.9	2	Negative	0.98	-1.28	-0.34	0.26
17	58	3	Ductal	9.0	2	Positive	0.96	-0.48	0.26	-0.42
19	68	13	Ductal	1.2	1	Negative	0.97	-2.18	-2.62	0.31
18[a]	67	11	Ductal	1.5	2	Micrometastasis	NA	NA	NA	NA
20	44	13	Ductulo-lobular	3.2	3	Positive	0.98	-0.12	0.01	0.19
21	71	20	Ductulo-lobular	1.7	3	Negative	0.95	-3.15	-1.69	0.26
22	53	5	Ductal	1.3	2	Negative	0.92	-2.81	-5.48	-0.08

[a]RNA in the core biopsy (CB) of this patient was degraded and, therefore, omitted from the study. COX, cyclooxygenase; NA, not available; PAI-1, Plasminogen Activator Inhibitor-1; ST, surgical tumor tissue specimen.

processes share molecular characteristics with an aggressive cancer phenotype, such as cell proliferation and survival, angiogenesis, and extracellular matrix remodeling; these biological hallmarks were shown to predict the clinical course in cancer [3,11-13]. Therefore, potential biological changes induced by CB require further study since they may have important consequences for clinical outcome prediction and treatment decisions as well as the interpretation of GEP changes investigated in neo-adjuvant studies.

We explored the molecular expression levels of 60 genes using quantitative real-time PCR (qrt-PCR), a highly sensitive and reproducible method, in paired CB and surgical samples. These genes were selected according to their known links to malignant cell behavior in breast cancer and their importance in major cancer hallmarks, such as proliferation, survival, invasiveness and angiogenic potential, and in the wound healing process [14]. Our first objective was to investigate whether the molecular profile of a CB is representative for the whole tumor. The second objective was to study if the repair process following CB alters the GEP and if this is influenced by the timeframe between the CB and the surgical excision.

Figure 1



Unsupervised hierarchical clustering of paired core biopsy and surgical tumor tissue samples (Spearman correlation, average linkage). Red color indicates high expression levels (low ΔCt) and green vice versa. Patients 6, 11 and 20 had two core biopsies taken that were analyzed separately. In four cases the paired gene expression profiles did not cluster together close to each other.

Materials and methods

Tumor specimen acquisition

Between June 2004 and June 2005, 22 consecutive breast cancer patients entered this study, for which both CBs and surgical tumor tissue specimens (STs) were available. This study was undertaken at the Women's University Hospital Basel, Switzerland, and approved by the local institutional review board (EKBB permission Nr. 81/04). Written informed consent was obtained from all patients.

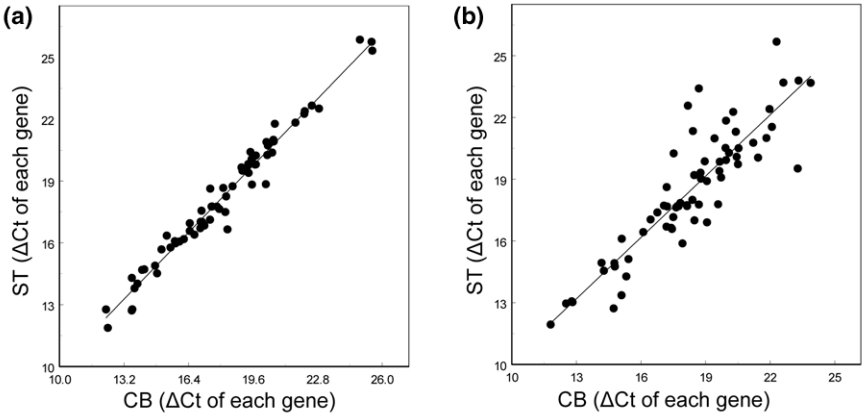
All CBs (14-gauge needle, Magnum® Core high speed, Bard Medica, Karlsruhe, Germany) were obtained under sonographic guidance (HDI 5'000 Sono CT®, Philips, Zurich, Switzerland) under local anesthesia through a skin incision in a sterile field. Five biopsy specimens were taken routinely for each patient. Two biopsy specimens were divided longitudinally; two halves to be used for molecular examination were, within one minute, stored in RNAlater®-solution (Qiagen, Basel, Switzerland), while the other two halves as well as the other three biopsies were immediately put in formaldehyde and sent for histological examination to the Institute of Pathology, University Hospital Basel.

After establishing the diagnosis of breast cancer, all patients underwent breast surgery with sentinel node lymphonodectomy or axillary lymphonodectomy if indicated. All surgical samples were examined by the same pathologist. If the tumor tissue was larger than 0.5 cm in diameter at the intra-operative frozen section, a representative piece containing more than 60% tumor cells was cryopreserved within five minutes and made available to the Stiftung Tumorbank Basel for molecular examination. The rest of the tumor tissue was embedded in paraffin for routine histological examination.

Reference study population

RNA expression levels of all 60 genes were detected using the same qrt-PCR method in 317 surgically excised breast cancer specimens [15] from patients undergoing primary surgery in 1992 to 1996 without previous examination by CB. All tissue samples were prepared by the pathologists as described above, all samples contained more than 60% tumor cells and

Figure 2



Correlation between the gene expression profiles (60 genes) of paired core biopsies (CB) and surgical tumor tissue specimens (ST). **(a)** Specimens from patient 2 (rs = 0.98), as representative for 82% of all cases. **(b)** Specimens from patient 12, as representative for less correlated paired gene expression profiles (rs = 0.86).

they were cryoperserved within five minutes. The Stiftung Tumorbank Basel was subsequently responsible for collection, storage at -70°C and analysis.

RNA extraction and qrt-PCR

Detailed procedures have been published elsewhere [16]. In brief, RNA was extracted using the RNeasy Mini Kit (Qiagen, Inc., Valencia, CA, USA), quantified and quality-checked on a Bioanalyzer 2100 (RNA 6000 Nano LabChip-Kit, Agilent Technologies, New Castle, DE, USA). High quality RNA samples were reverse-transcribed (10 mM DDT, 1 μg of hexamer primers, 2 U of MMLV Reverse Transcriptase (Invitrogen, Basel, Switzerland), 40 U of RNasin (Promega, Wallisellen, Switzerland), 0.5 mM of each dNTP (Promega), 1× reaction buffer). PCR primers were designed to be cDNA specific and ordered at GeneScan Europe (Freiburg, Germany). PCR was performed in 40 cycles on an ABI Prism 7000 using 2× SYBR Green I Master Mix (Applied Biosystems, Forster City, CA, USA) in a final volume of 25 μl. Relative quantities (ΔCt) were obtained by normalization against ribosomal 18S RNA, and standardization was achieved with Human Universal Standard RNA (Stratagene Europe, Amsterdam, The Netherlands). The 60 genes quantitatively assessed are listed in Table 1.

Statistical analysis

The same amount of RNA was used for the GEP analysis of each sample. For statistical analysis, ΔCt expression values of each gene were obtained by normalizing the raw gene values to 18S rRNA as a reference gene.

Cluster and TreeView programs were used to perform unsupervised hierarchical clustering of samples and genes (Spearman correlation, average linkage) [17]. Spearman correlation coefficients were calculated to compare the GEPs of all paired samples. Differentially expressed genes were identified with the paired two-sample *t*-test. The Mann-Whitney *U*-test was used to compare median expression values of genes among different subgroups of patients. All statistical analyses were carried out at 5% level of significance and performed with S-Plus software (Version 6.1, Insightful Corporation, Seattle, WA, USA).

Results

Patient characteristics

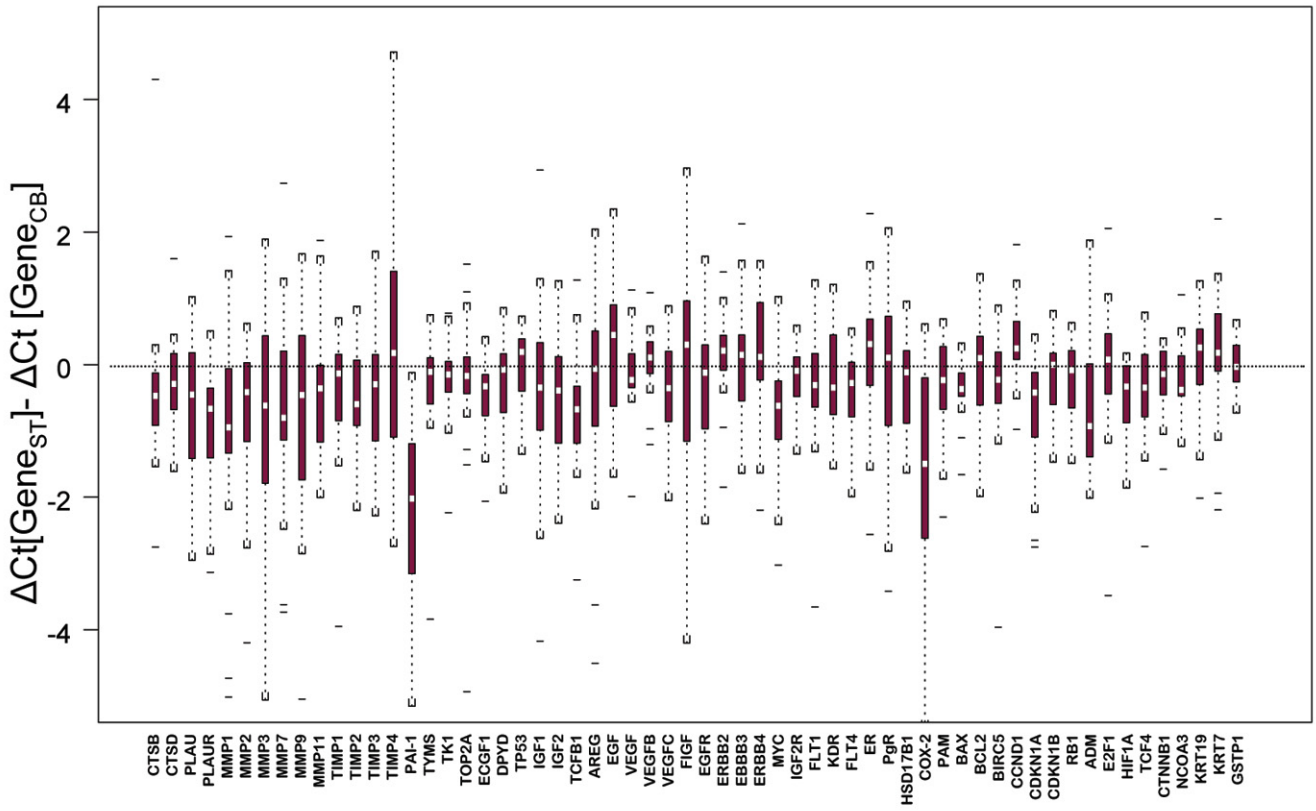
The mean age of the 22 patients was 63 years (range: 34 to 81 years). The period between CB and final surgery ranged from 1 to 23 days. The mean tumor diameter was 2.9 cm (range: 0.9 to 9 cm). Clinicopathological characteristics of each patient are listed in Table 1. Of note, one paired tissue sample was not further evaluated due to poor RNA quality, and RNA extracted from the two halves of the CB was not pooled in three cases but analyzed separately.

Comparison of the gene expression profile in paired samples

As shown in Figure 1, unsupervised hierarchical clustering revealed that paired CB and ST generally clustered together; in only four cases (patients 4, 10, 11, and 12) did the GEP of the CB not agglomerate with the profile of the respective ST. Interestingly, the two separate CBs taken from patient 11 were very similar to each other, although they differed from their ST GEP. The gene dendrogram of the cluster analysis also revealed that samples agglomerated in two main groups according to their respective estrogen receptor status, reconfirming the representative value of this study population.

Subsequent analysis of paired CBs and STs confirmed the high correlation between all samples (r_s from 0.86 to 0.98, all $p < 0.001$) for each patient. A scatter plot of two representative examples of a paired CB/ST is displayed in Figure 2. The differences in paired GEPs does not seem to be related to the

Figure 3



Box plots displaying the changes in the expression of each gene in the surgical specimens (ST) compared to the respective core biopsies (CB).

timeframe between CB and surgery or to any other clinicopathological parameters (Table 1).

Gene-by-gene analysis in core biopsies and paired surgical specimens

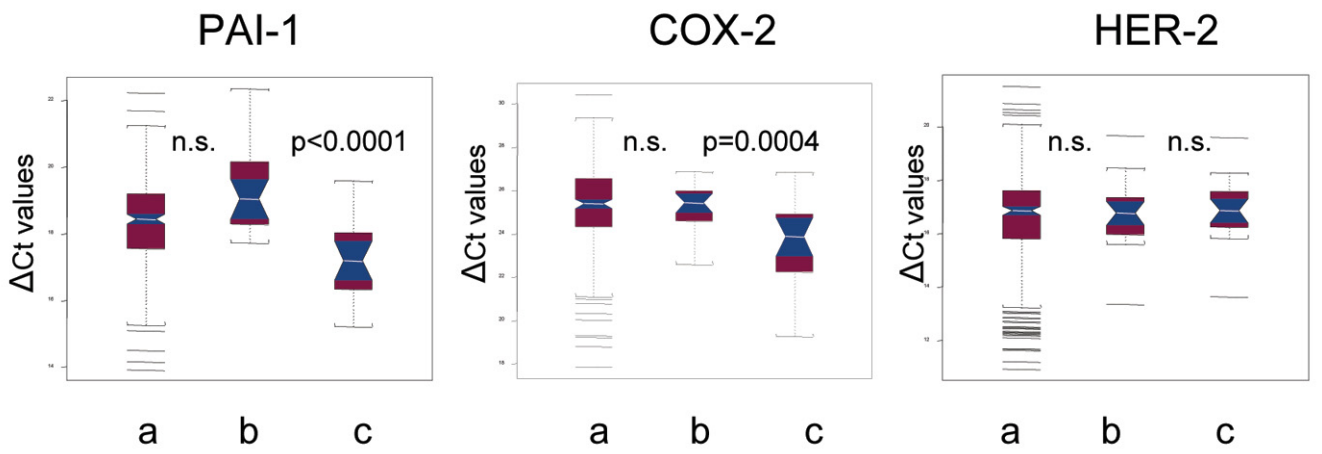
The comparison of the expression levels of individual genes by means of paired *t*-test showed no significant difference between CB and ST with the exception of four genes. Plasminogen activator inhibitor 1 (PAI-1; also known as SERPINE1) was significantly higher expressed (lower ΔCt values) in STs compared to CBs ($p < 0.001$, Table 1). Similar differences, although less pronounced, were observed for cyclooxygenase 2 (COX-2; also known as PTGS2; $p < 0.001$), urokinase plasminogen activator receptor (uPAR; also known as PLAUR; $p = 0.003$) and matrix metalloproteinase 1 (MMP1; $p = 0.03$). The increase in the expression of these genes was not related to the timeframe between CB and surgery. All other genes were very similarly expressed in paired CB/ST as shown in Figure 3. Table 1 lists differences in RNA expression values of PAI-1, COX-2 and ERBB2 for paired samples of each patient.

Histological re-examination of cryocuts of the surgical specimens revealed a certain amount of inflammation and fibrolysis. Whether these observations are due to cancerogenesis or to a *de novo* induced wound repair process can not be determined.

Comparison of the expression levels of selected genes with a reference study population of surgical tumor tissue specimens

To verify whether the higher expression levels of PAI-1 and COX-2 observed in STs could have been induced by the preceding CB procedure, we compared the expression levels of the same genes in an independent population of 317 primary breast cancer patients. These samples were investigated with the same qrt-PCR technique but were from patients from which no CB had been taken prior to surgery. As shown in Figure 4, the expression levels of PAI-1 and COX-2 measured in the independent STs without CB were found to be very similar to the levels detected in CBs and significantly different from those detected in STs after CB. Moreover, no variation at all

Figure 4



Notch box plots of the expression levels of plasminogen activator inhibitor (PAI)-1, cyclooxygenase (COX)-2 and HER-2 as detected **(a)** in 317 independent surgical specimens of primary breast cancer patients who did not undergo previous diagnosis by core biopsy, **(b)** in the core biopsy of the 21 patients entering this study, and **(c)** in the corresponding surgical excisions.

was observed for the remaining genes as illustrated by the expression levels of ERBB2 as an example.

Discussion

Ultrasound-guided CB is a well established method to diagnose breast cancer in women, since it is a reliable, and time- and cost-saving method. The clinical utility of the information gained by CB depends on whether the CB is representative for the whole tumor. Our data demonstrate that the quantitative expression levels of 60 genes detected in CBs were highly comparable to their paired STs in 17 out of the 21 cases investigated. Even in the cases where GEPs of a CB and ST did not agglomerate, the expression levels of the ER and progesterone receptor as well as ERBB2 measured in the CB were also representative for the whole tumor. This is important since today's therapy decisions are based on these markers obtained either by CB or ST (St Gallen consensus recommendations [18]). In addition, our results reconfirm previous observations reported by immunohistochemistry [19,20] or semi-qrt-PCR [21].

Tissue sampling by CB causes a local injury, inducing wound healing that is characterized by recruitment of inflammatory cells, stimulation of stromal and epithelial cell proliferation, cell migration and increased angiogenesis. Analysis on a gene-by-gene basis demonstrated higher expression levels of PAI-1, COX-2, uPAR and MMP1 in STs compared to their paired CBs, whereas no changes were observed for all other genes. These results are not surprising since proteinases (such as PAI-1 and uPAR) are known to be essentially involved in the wound healing process [22,23] and COX-2 plays roles in inflammation and angiogenesis [24-27].

However, many of these reparation processes show parallels with cancerogenesis [28-30]; while proteinases, their inhibitors, cyto-/chemokines and growth factors are essential for wound healing and tissue repair, they also play central roles in cancer progression. For example, uPA, uPAR and its inhibitor PAI-1 are responsible for the degradation and remodeling of the extracellular matrix, and are further involved in angiogenesis, cell adhesion and migration necessary for tumor cell invasion and metastasis [31,32]. COX-2 can be induced by cytokines and growth factors during the inflammatory repair process as well as in cancer [24-26,30] resulting in COX-2 overexpression observed in human malignancies [25-27].

Therefore, increased levels of these markers in the tumor specimen could suggest a more aggressive cancer phenotype. Indeed, elevated levels of uPA and PAI-1 are associated with poor clinical outcome in breast cancer and also have predictive value [33-35]. Moreover, a previously identified 'wound-response signature' turned out to be prognostic in several carcinomas, including breast cancer [12,13]. Although COX-2 has been associated with increased Vascular Endothelial Growth Factor A (VEGF), estrogen synthesis, proliferation, apoptosis and invasion [25,27,36], in our study, higher levels of COX-2 were not accompanied by changes in the expression of genes involved in these processes, indicating that the observed molecular alterations influence data interpretation but not tumor aggressiveness.

Conclusion

Our study demonstrates that expression levels of ER, progesterone receptor, ERBB2 and other genes relevant for the management of breast cancer as detected in CBs are representative for the whole tumor. However, increased expression levels of proteinases (e.g. PAI-1, uPAR, MMP1)

and COX-2 in STs compared to their paired CBs suggest induction of theses genes during the repair process following tissue injury caused by CBs. This observation is important since such molecular alterations may have an impact on the clinical interpretation of GEPs detected in STs with respect to the prediction of risk assessment and treatment response.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

RS and SE designed the study, presented it to the ethical committee and drafted the manuscript. RS performed the sonographic guided core biopsies and coordinated the clinical part of the study. VV and ML contributed to the selection of the genes, selected primers and supervised RNA extraction and qrt-PCR. VV and PU contributed to statistical analysis, data interpretation and manuscript revision. GS performed the pathological analysis of all samples, prepared the surgical samples for molecular analysis and participated in data interpretation. EW, UE and WH participated in designing the study and writing the manuscript. SE coordinated the molecular investigation and performed statistical analysis. All authors read and approved the final manuscript.

Acknowledgements

This work was in part supported by the generous donation of Mrs Bärbel Girgner. We are grateful to Monika Heinemann, Sabine Ehret and Ulrike Weissenstein for technical assistance and data management and to Linda Herberich for proof reading.

References

1. 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 USA* 2001, **98:**10869-10874.
2. Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R, Mohsin S, Osborne CK, Chamness GC, Allred DC, *et al.*: **Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer.** *Lancet* 2003, **362:**362-369.
3. Chang JC, Weiss H, Mohsin S: **Apoptosis induction by the core biopsy procedure.** *J Clin Oncol* 2005, **23:**7240.
4. 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:**1999-2009.
5. 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:**530-536.
6. Fornage BD, Sneige N, Edeiken BS: **Interventional breast sonography.** *Eur J Radiol* 2002, **42:**17-31.
7. Liberman L: **Percutaneous image-guided core breast biopsy.** *Radiol Clin North Am* 2002, **40:**483-500. vi
8. Parker SH, Burbank F, Jackman RJ, Aucreman CJ, Cardenosa G, Cink TM, Coscia JL Jr, Eklund GW, Evans WP 3rd, Garver PR, *et al.*: **Percutaneous large-core breast biopsy: a multi-institutional study.** *Radiology* 1994, **193:**359-364.
9. Philpotts LE: **Controversies in core-needle breast biopsy.** *Semin Roentgenol* 2001, **36:**270-283.
10. Symmans WF, Ayers M, Clark EA, Stec J, Hess KR, Sneige N, Buchholz TA, Krishnamurthy S, Ibrahim NK, Buzdar AU, *et al.*: **Total RNA yield and microarray gene expression profiles from fine-needle aspiration biopsy and core-needle biopsy samples of breast carcinoma.** *Cancer* 2003, **97:**2960-2971.
11. Menard S, Pupa SM, Campiglio M, Taglibue E, Balsari A, Fagnoni F, Costa A: **Apoptosis induction by trastuzumab: possible role of the core biopsy intervention.** *J Clin Oncol* 2005, **23:**7238-7240.
12. Chang HY, Sneddon JB, Alizadeh AA, Sood R, West RB, Montgomery K, Chi JT, van de Rijn M, Botstein D, Brown PO: **Gene expression signature of fibroblast serum response predicts human cancer progression: similarities between tumors and wounds.** *PLoS Biol* 2004, **2:**E7.
13. Chang HY, Nuyten DS, Sneddon JB, Hastie T, Tibshirani R, Sorlie T, Dai H, He YD, van't Veer LJ, Bartelink H, *et al.*: **Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival.** *Proc Natl Acad Sci USA* 2005, **102:**3738-3743.
14. Hanahan D, Weinberg RA: **The hallmarks of cancer.** *Cell* 2000, **100:**57-70.
15. Urban P, Vuaroqueaux V, Labuhn M, Delorenzi M, Wirapati P, Wight E, Senn H-J, Benz C, Eppenberger U, Eppenberger-Castori S: **Increased expression of urokinase-type plasminogen activator (uPA) mRNA determines adverse prognosis in ErbB2-positive primary breast cancer.** *J Clin Oncol* in press.
16. Labuhn M, Vuaroqueaux V, Fina F, Schaller A, Nanni-Metellus I, Küng W, Eppenberger-Castori S, Martin PM, Eppenberger U: **Simultaneous quantitative detection of relevant biomarkers in breast cancer by quantitative real-time PCR.** *Int J Biol Markers* 2006, **21:**30-39.
17. Eisen M, Spellman P, Brown P, Botstein D: **Cluster analysis and display of genome-wide expression patterns.** *Proc Natl Acad Sci USA* 1998, **95:**14863-14868.
18. Goldhirsch A, Glick JH, Gelber RD, Coates AS, Thurlimann B, Senn HJ: **Meeting highlights: international expert consensus on the primary therapy of early breast cancer 2005.** *Ann Oncol* 2005, **16:**1569-1583.
19. Cavaliere A, Sidoni A, Scheibel M, Bellezza G, Brachelente G, Vitali R, Bucciarelli E: **Biopathologic profile of breast cancer core biopsy: is it always a valid method?** *Cancer Lett* 2005, **218:**117-121.
20. Badoual C, Maruani A, Ghorra C, Lebas P, Avigdor S, Michenet P: **Pathological prognostic factors of invasive breast carcinoma in ultrasound-guided large core biopsies-correlation with subsequent surgical excisions.** *Breast* 2005, **14:**22-27.
21. Pohlodek K, Galbavy S, Bartosova M, Mucha V, Holoman K, Pastorek J, Pastorekova S: **Semi-quantitative RT-PCR assessment of molecular markers in breast large-core needle biopsies.** *Neoplasma* 2004, **51:**415-421.
22. Baker EA, Leaper DJ: **Proteinases, their inhibitors, and cytokine profiles in acute wound fluid.** *Wound Repair Regen* 2000, **8:**392-398.
23. Stadelmann WK, Digenis AG, Tobin GR: **Physiology and healing dynamics of chronic cutaneous wounds.** *Am J Surg* 1998, **176:**26S-38S.
24. Williams CS, Mann M, DuBois RN: **The role of cyclooxygenases in inflammation, cancer, and development.** *Oncogene* 1999, **18:**7908-7916.
25. Howe LR, Dannenberg AJ: **COX-2 inhibitors for the prevention of breast cancer.** *J Mammary Gland Biol Neoplasia* 2003, **8:**31-43.
26. Corcoran CA, He Q, Huang Y, Sheikh MS: **Cyclooxygenase-2 interacts with p53 and interferes with p53-dependent transcription and apoptosis.** *Oncogene* 2005, **24:**1634-1640.
27. Ma X, Yang Q, Wilson KT, Kundu N, Meltzer SJ, Fulton AM: **Promoter methylation regulates cyclooxygenase expression in breast cancer.** *Breast Cancer Res* 2004, **6:**R316-321.
28. Coussens LM, Werb Z: **Inflammation and cancer.** *Nature* 2002, **420:**860-867.
29. Dvorak HF: **Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing.** *N Engl J Med* 1986, **315:**1650-1659.
30. Balkwill F, Mantovani A: **Inflammation and cancer: back to Virchow?** *Lancet* 2001, **357:**539-545.
31. Harbeck N, Kates RE, Gauger K, Willems A, Kiechle M, Magdolen V, Schmitt M: **Urokinase-type plasminogen activator (uPA) and its inhibitor PAI-I: novel tumor-derived factors with a high prognostic and predictive impact in breast cancer.** *Thromb Haemost* 2004, **91:**450-456.

32. Duffy MJ, Duggan C: **The urokinase plasminogen activator system: a rich source of tumour markers for the individualised management of patients with cancer.** *Clin Biochem* 2004, **37:**541-548.
33. Decock J, Paridaens R, Cufer T: **Proteases and metastasis: clinical relevance nowadays?** *Curr Opin Oncol* 2005, **17:**545-550.
34. Cianfrocca M, Goldstein LJ: **Prognostic and predictive factors in early-stage breast cancer.** *Oncologist* 2004, **9:**606-616.
35. Look M, van Putten W, Duffy M, Harbeck N, Christensen IJ, Thomssen C, Kates R, Spyratos F, Ferno M, Eppenberger-Castori S, *et al.*: **Pooled analysis of prognostic impact of uPA and PAI-1 in breast cancer patients.** *Thromb Haemost* 2003, **90:**538-548.
36. Singh B, Berry JA, Shoher A, Ramakrishnan V, Lucci A: **COX-2 overexpression increases motility and invasion of breast cancer cells.** *Int J Oncol* 2005, **26:**1393-1399.

Research article

Open Access

Increased level of phosphorylated akt measured by chemiluminescence-linked immunosorbent assay is a predictor of poor prognosis in primary breast cancer overexpressing ErbB-2

Jonas Cicenas[1,2], Patrick Urban[1], Vincent Vuaroqueaux[3], Martin Labuhn[3], Willy Küng[2], Edward Wight[4], Mark Mayhew[5], Urs Eppenberger[2,3] and Serenella Eppenberger-Castori[1]

[1]Stiftung Tumorbank Basel, Basel, Switzerland
[2]University Clinics, Department of Research, Molecular Tumor Biology, Basel, Switzerland
[3]OncoScore AG, Riehen, Switzerland
[4]University Clinics, Department of Gynecology, Basel, Switzerland
[5]University of Virginia, Department of Anatomy and Cell Biology, East Carolina School of Medicine, Charlottesville, Virginia, USA

Corresponding author: Serenella Eppenberger-Castori, s.eppenberger@tumorbank.org

Received: 14 Oct 2004 Revisions requested: 16 Dec 2004 Revisions received: 9 Feb 2005 Accepted: 28 Feb 2005 Published: 24 Mar 2005

Breast Cancer Research 2005, **7**:R394-R401 (DOI 10.1186/bcr1015)
This article is online at: http://breast-cancer-research.com/content/7/4/R394
© 2005 Cicenas *et al.*; licensee BioMed Central Ltd.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Introduction Akt1, Akt2 and Akt3 kinases are downstream components of phosphoinositol 3-kinase derived signals from receptor tyrosine kinases, which influence cell growth, proliferation and survival. Akt2 overexpression and amplification have been described in breast, ovarian and pancreatic cancers. The present study was designed to investigate the prognostic significance of activated Akt in primary breast cancer and its association with other tumour biomarkers.

Methods Using a two-site chemiluminescence-linked immunosorbent assay, we measured the quantitative expression levels of total phosphorylated (P-S473) Akt (Akt1/Akt2/Akt3) on cytosol fractions obtained from fresh frozen tissue samples of 156 primary breast cancer patients.

Results Akt phosphorylation was not associated with nodal status or ErbB-2 protein expression levels. High levels of phosphorylated Akt correlated ($P < 0.01$) with poor prognosis, and the significance of this correlation increased ($P < 0.001$) in the subset of patients with ErbB-2 overexpressing tumours. In addition, phosphorylated Akt was found to be associated with mRNA expression levels of several proliferation markers (e.g. thymidylate synthase), measured using quantitative real-time RT-PCR.

Conclusion Our findings demonstrate that, in breast cancer patients, Akt activation is associated with tumour proliferation and poor prognosis, particularly in the subset of patients with ErbB2-overexpressing tumours.

Introduction

Akt/protein kinase B (PKB) is a serine/threonine kinase that is involved in mediating various biological responses, such as inhibition of apoptosis and stimulation of cell proliferation (for review [1,2]). Three mammalian isoforms are currently known [1]: Akt1/PKBα, Akt2/PKBβ and Akt3/PKBγ. Akt1 was first discovered as a cellular homologue of the viral oncogene v-Akt, which causes leukaemia in mice [3] and is the predominant isoform in most tissues. High expression of Akt2 has been observed in insulin-responsive tissues, whereas Akt3 has been shown to be predominantly expressed in brain and testis [2].

Phosphoinositol-3-phosphate (PIP3) is a product of phosphoinositol 3-kinase enzymatic activity and has been shown to be a prerequisite lipid modulator of Akt activity [4]. PIP3 has been described as a downstream component of a wide range of receptors, including the c-Met receptor [5], the epidermal growth factor receptor family [6], fibroblast growth factor receptor [7], insulin growth factor receptor [8] and platelet-derived growth factor receptor [9]. In addition, Akt activity can

CLISA = chemiluminescence-linked immunoassay; DFS = disease-free survival; EIA = enzyme immunoassay; ER = oestrogen receptor; P-Akt = phosphorylated Akt; PgR = progesterone receptor; PIP3 = phosphoinositol-3-phosphate; PKB = protein kinase B; RT-PCR = reverse transcription polymerase chain reaction; STB = Stiftung Tumorbank Basel.

be regulated by the PTEN tumour suppressor gene, which negatively regulates PIP3 levels (for review [10]). After PIP3 binding, Akt1 is activated by phosphorylation on two critical residues, namely threonine 308 (T308) and serine 473 (S473); similar activation residues (S472 and S474, respectively) are highly conserved in Akt2 and Akt3 (for review [1,2]). Several studies have found Akt2 to be amplified or overexpressed at the mRNA level in various tumour cell lines [11-13] and in a number of human malignancies, such as colon, pancreatic and breast cancers [14-16]. However, activation of Akt1, Akt2 and Akt3 by phosphorylation appears to be more clinically relevant than detection of Akt2 amplification or overexpression.

To date, several groups have investigated the phosphorylation of active Akt in breast, prostate, colon and pancreatic tumours by immunohistochemistry [14,17-22]. Under such conditions, phosphorylation structures may be disturbed by formalin fixation, rendering specific antigen sites inaccessible. Moreover, immunohistochemistry gives only semiquantitative results, limiting statistical analysis. Alternatively, enzyme immunoassays (EIAs) have the advantage that they yield highly reproducible and sensitive results of quantitative values.

In the present study we detected phosphorylated Akt (P-Akt) by means of a novel two-site chemiluminescence-linked immunoassay (CLISA) in fresh frozen primary tissue samples from 156 primary breast cancer patients. Because it was shown in previous immunohistochemistry studies that S473 P-Akt has prognostic significance [17-19], the aim of the present study was to measure levels of P-Akt continuously using CLISA and correlate these with survival and factors that are involved in tumourigenesis. Given that the antibody used in the reported immunohistochemistry studies recognized all Akt isoforms, we have developed an assay that allows specific quantitative detection of active Akt1, Akt2 and Akt3 when phosphorylated on their corresponding residues, namely S473, S472 and S474, respectively.

Materials and methods

Tumor and patient characteristics

Fresh material obtained during surgery was kept on ice and examined by a pathologist. Representative specimens with more than 60% tumour cells were sent to the Stiftung Tumorbank Basel (STB), immediately shock frozen and cryopreserved (-80°C). All activities of the STB are in accordance with an official Swiss permit, which guarantees patient confidentiality and respects ethical issues. For the present study, 156 samples of primary breast tumours were selected. Those samples overexpressing ErbB-2 (>500 U/mg total protein) were selected, based on ErbB-2 protein expression levels routinely detected using EIAs at the time of surgery by the STB [23]. EIA ErbB-2 positive samples correlate strongly with DAKO 3+ and with ErbB-2 amplification detected by fluorescent *in situ* hybridization (FISH; data not shown).

All patients underwent primary surgery before January 1996. Sixty-seven patients (43%) experienced disease recurrence within the median follow-up time of 57 months (range 27–88 months). Sixty-six patients (42%) were node negative, and 90 (58%) were node positive. Forty tumours (26%) were oestrogen receptor (ER)-α negative. Ninety-five patients (61%) had ErbB-2-negative (<500 U/mg total protein) and 61 patients (39%) had ErbB-2 positive tumours [23]. None of the patients received neoadjuvant therapy. Patient and tumour characteristics are summarized in Table 1.

Cell lines and tissue culture

MCF-7 breast cancer cells were cultured in IMEM-ZO (improved minimal essential medium with zinc option) supplemented with 5% foetal bovine serum, l-glutamine and antibiotics (penicillin/streptomycin) at 37°C in a 5% carbon dioxide incubator. For the phospho-standard preparation, subconfluent MCF-7 cells were serum starved for 48 hours in serum-free media, and were treated with NaF and Na_3VO_4 for 1 hour, and then with 10% foetal bovine serum for 10 min. Cells were lysed for 5 min on ice in EB lysis buffer (20 mmol/l Tris-HCl [pH 7.4], 0.5 mol/l NaCl, 10 mmol/l EDTA, 1% Triton X100, 20 mmol/l NaF, 20 mmol/l glycerophosphate, 2 mmol/l Na_3VO_4, proteinase inhibitor cocktail [Roche, Indianapolis, IN, USA]), centrifuged at 20,000 *g* for 5 min and supernatant was stored at -80°C.

Measurement of oestrogen receptor, progesterone receptor and ErbB-2 protein levels in tumour extracts by enzyme immunoassay

Tissue homogenates were prepared in accordance with standard procedures for tumour marker measurement using EIAs, as previously described [23]. In brief, the frozen tissues were pulverized in liquid nitrogen using a Micro-Dismembrator U (B Braun Melsungen AG, Melsungen, Germany). The powder was homogenized using a tissue homogenizer (Ultra-Turrax; Janke & Kunkel, IKA-Werke, Staufen, Germany) for 20 s in three volumes of ice-cold extraction buffer. The homogenate was centrifuged at 800 *g* for 30 min at 2°C, and the resulting supernatant re-centrifuged in an ultracentrifuge (Beckman Instruments, Fullerton, CA, USA) at 100,000 *g*. The resulting supernatants (cytosols) were used for measurement of the hormone receptors (ER, progesterone receptor [PgR]), and the membrane fractions were used for EIA measurement of membrane-associated ErbB-2. ER and PgR concentrations were measured from tumour cytosolic extracts by commercial quantitative ER and PgR EIA kits (Abbott Laboratories, Abbott Park, IL, USA) using a Quantum II photometer (Abbott Laboratories, Abbott Park, IL, USA). Quality control of ER and PgR measurements was carried out in collaboration with the Receptor Biomarker Group of the European Organization for Research and Treatment of Cancer. ErbB-2 receptor levels were determined on the particulate membrane fractions of tumour extracts using a commercial monoclonal antibody EIA kit, described by Eppenberger-Castori and coworkers [23].

Table 1

Clinicopathological characteristics of the patients

Feature	Number of patients (%)
Patients enrolled	156
Age (years):	
<40	12 (8)
40–60	85 (54)
>60	59 (38)
Histology type:	
Ductal	109 (70)
Lobular	17 (11)
Other	30 (19)
Tumour size:	
T1	49 (31)
T2	90 (58)
T3-T4	17 (11)
Lymph-node status	
Node negative	66 (42)
Node positive	90 (58)
Histopathological grade	
I + II	57 (37)
III	86 (55)
Not analyzed	13 (8)
Oestrogen receptor	
Positive (>20 fmol/mg)	116 (74)
Negative (≤ 20 fmol/mg)	40 (26)
Progesterone receptor	
Positive (>20 fmol/mg)	85 (54)
Negative (≤ 20 fmol/mg)	71 (46)

Immunoassay of phosphorylated Akt level

Neither antibody used in the CLISA discriminates between Akt isoforms. The catching antibody (anti-Akt/PKB, PH domain, clone SKB1; Upstate Biotechnology, Lake Placid, NY, USA) recognizes Akt1/PKBα, Akt2/PKBβ and Akt3/PKBγ (weak to none) based on immunoblot analysis using 100 ng recombinant fusion protein for each isoform, as reported by the manufacturer. The detecting phospho-specific (S473) Akt monoclonal antibody (4E2) detects endogenous levels of Akt1 only when phosphorylated at serine-473. This antibody also recognizes Akt2 (S472) and Akt3 (S474) if they are phosphorylated at the corresponding residues, according to the information obtained from the manufacturer (Cell Signaling Technology, Inc., Beverly, MA, USA). However, 4E2 does not recognize other Akt phosphorylation sites.

S473 phosphorylated Akt levels were measured using a novel two-site CLISA. Black 96-well microtitre plates (Nunc Black MaxiSorp Surface; Nalgen Nunc International, Rochester, NY, USA) were coated with coating antibody at a concentration of 3 mg/ml of coating buffer (phosphate-buffered saline with 0.6 mmol/l EDTA) in a volume of 100 µl/well and kept at 4°C overnight. To measure P-Akt, respective tumour extracts were prepared as described above in the presence of NaF and

Na_3VO_4. Before sample applications, the coated microtitre plates were washed five times with 200 µl/well washing buffer (25 mmol/l HEPES [pH 7.4], 300 mmol/l NaCl, 0.05% Tween-20) and then blocked for 2 hours at room temperature with 250 µl blocking buffer (25 mmol/l HEPES [pH 7.4], 300 mmol/l NaCl, 0.05% Tween-20, 3% TopBlock [Juro AG, Lucerne, Switzerland]). Blocked wells were washed five times with 200 µl washing buffer, and then 100 µl diluted tumour membrane extracts or reference material was added to the wells and incubated overnight at 4°C.

As a reference for each assay, an extract of MCF-7 cells, prepared as described above, was used. For use in the assay, MCF-7 cell extracts were sequentially diluted with sample dilution buffer (blocking buffer, proteinase inhibitor cocktail, NaF and Na_3VO_4) at ratios of 1×, 0.75×, 0.5×, 0.25×, 0.125× and 0.025×, and then 100 µl aliquots were incubated on each microtitre plate, together with tumour tissue extracts and negative controls (containing only dilution buffer). After incubation of the samples and reference material, wells were washed five times with 200 µl washing buffer at room temperature to eliminate unbound particles. Biotinylated detection antibody was added, followed by incubation for 2 hours at room temperature. Complexes were detected with horseradish peroxidase-conjugated streptavidin, diluted in conjugate diluents for 1 hour at room temperature. Horseradish peroxidase activity was detected using SuperSignal WestPico substrate (Pierce, Rockford, IL, USA) in a glow luminometer. The response data for diluted reference material was fitted, and the respective curve was used for the quantification of tumour extracts. The value of undiluted MCF-7 extracts was denominated as 1 U/ml.

Quantitative real-time RT-PCR for the detection of proliferation markers

RNA was extracted using RNeasy kit (Qiagen, Hilden, Germany). Quality and quantity were checked using a Bioanalyzer 2100 (Agilent, Palo Alto, CA, USA). All genes were examined using SYBR Green I methods with Taqman 7000 (Applied-Biosystems, Foster City, CA, USA). Relative quantification ($\Delta\Delta$Ct) was obtained by normalization with ribosomal 18S and a standardization step with Human Universal Standard RNA (Stratagene, La Jolla, CA, USA). Quantitative real-time RT-PCR results were expressed in arbitrary units of reverse transcribed RNA (U/µg rt-RNA).

Statistical methods

The statistical significance of the association between P-Akt and other dichotomous variables (e.g. node status) was assessed using Mann–Whitney U-test. Spearman rank correlation (r_s) was calculated to assess associations between continuous markers (e.g. ErbB-2 or tumour size and P-Akt protein expression levels). The continuous variable function of CLISA-determined P-Akt values was first tested for prognostic significance by univariate Cox regression. A cutoff or prognostic threshold value with respect to relapse-free survival was sought by means of classification and regression tree analysis [24,25]. Survival probabilities were calculated using the Kaplan–Meier method and compared by means of log-rank analysis [26]. The Cox proportional hazards regression model was also applied over multivariate analyses, with the associated likelihood ratio test used to assess test-of-trend differences. The results of multivariate Cox regression analysis were summarized in a table and expressed as relative risk for relapse.

Results

Distribution of phosphorylated Akt levels and its correlation with tumour characteristics

CLISA quantified P-Akt levels have a left-tailed distribution ranging from 0 to 1.08 U/mg total protein, with a median of 0.17 U/mg (mean 0.19 U/mg; Fig. 1) and could be transformed to normality by means of the 10th root. There was no correlation between P-Akt and ErbB-2 protein expression levels. In this set of primary breast cancer samples, we did not find any significant difference in P-Akt levels with respect to nodal status, tumour size, ER status or grading, nor any correlation between P-Akt levels and the continuous variables tumour size and ER level.

Prognostic significance of phosphorylated Akt levels

The prognostic value of P-Akt was investigated with respect to disease-free survival (DFS) in the patients overall (Fig. 2). Univariate Cox regression revealed a weak correlation between P-Akt levels and DFS ($P < 0.05$; likelihood ratio test). An optimal cutoff value for P-Akt (0.3 U/mg) was calculated using classification and regression tree analysis, dividing the patients into two subgroups: 21 patients (14%) patients expressed high levels of P-Akt (>0.31 U/mg total protein) and 135 patients (86%) expressed low levels of P-Akt. Subsequently, Kaplan–Meier survival curves stratified according to low and high P-Akt levels were plotted (Fig. 2). Sixty-seven per cent of patients (14 out of the 21) with high P-Akt levels relapsed, whereas only 36% (49 out of 135) with low P-Akt developed a relapse of disease within the period of observation ($P < 0.01$; log-rank test). The 5-year DFS was 33% in the high P-Akt group versus 60% in the low P-Akt group. The 5-year DFS in node-positive patients was 50% versus 68% in node-negative patients ($P < 0.05$; curves not shown).

Multivariate Cox analysis was performed including P-Akt and those additional variables that were found to have significant prognostic value in univariate Cox models (ER, ErbB-2 and node status, and tumour size and grading). In the tested multivariate model CLISA-determined elevated P-Akt level was an independent prognostic factor ($P = 0.02$), with a relative risk for breast cancer relapse of 2.09 (Table 2).

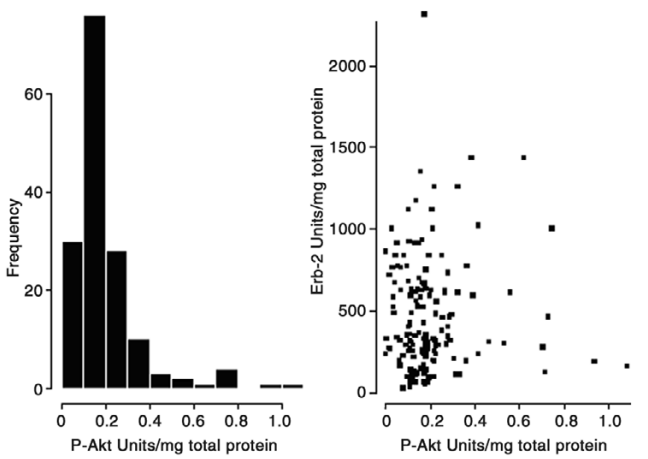
Table 2

Univariate and multivariate Cox analysis of relapse-free survival in patients with primary breast cancer

Factor	Univariate *P*	Multivariate *P*	Relative risk for relapse	95% CI
P-Akt	0.01	0.02	2.09	1.14–3.85
Node status	0.0003	0.09	1.33	0.95–1.85
ER status	0.03	0.17	0.67	0.38–1.19
ErbB-2 status	0.002	0.04	1.73	1.02–2.94
Grading	0.03	0.06	1.57	0.98–2.50
Tumour size	0.00005	0.02	1.51	1.07–2.13

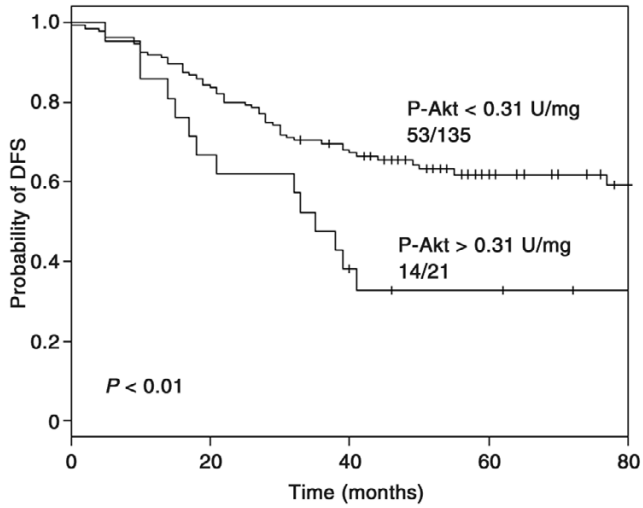
CI, confidence interval; ER, oestrogen receptor; P-Akt, phosphorylated Akt.

Figure 1



Chemiluminescence-linked immunoassay (CLISA)-quantified phosphorylated Akt (P-Akt) levels. **(a)** Histogram showing distribution of chemiluminescence-linked immunoassay (CLISA)-determined phosphorylated Akt (P-Akt) expression levels in 156 primary breast cancer samples. P-Akt levels ranged from 0 to 1.08 U/mg, with a median of 0.17 U/mg. **(b)** Scatter plot of P-Akt versus ErbB-2 expression levels. No correlation was found between the levels of P-Akt and ErbB-2.

Figure 2



Kaplan–Meier survival curves for patients overall. The curves are stratified by phosphorylated Akt (P-Akt) levels. Patients whose tumors express high levels of P-Akt exhibit a significantly worse outcome in terms of disease-free survival (DFS; $P < 0.01$).

Prognostic significance of phosphorylated Akt in ErbB2-overexpressing tumours

Although no correlation was found between P-Akt and ErbB-2 expression, the prognostic impact of P-Akt was greater in ErbB2-overexpressing tumours than in the samples overall. As shown in the Kaplan–Meier curves in Fig. 3a, patient prognosis decreased significantly when tumours expressed P-Akt levels higher than the median value ($P = 0.005$). This effect was even more pronounced when P-Akt levels exceeded the third quartile value ($P < 0.001$), which, together with the multivariate Cox-analysis, indicates that P-Akt has independent and additive prognostic value in combination with ErbB-2 (Fig. 3b).

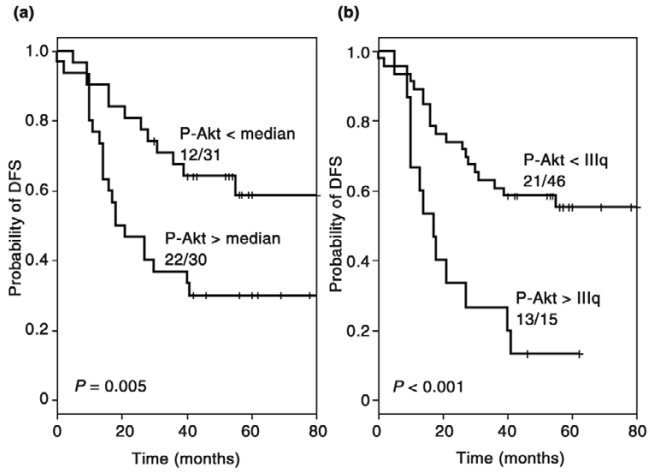
Correlation of phosphorylated Akt levels and mRNA expression of proliferation markers

Because involvement of P-Akt has been implicated in proliferation and apoptosis, we compared the quantitative P-Akt protein levels with the quantitative mRNA expression levels of genes involved in these biological processes. Using Spearman rank correlation, P-Akt levels were found to correlate with thymidylate synthase expression levels ($r_s = 0.38$; $P < 0.001$) and, to a lesser extent, with expression levels of thymidine kinase 1, survivin, topoisomerase IIα and the E2F transcription factor (Fig. 4, Table 3).

Discussion

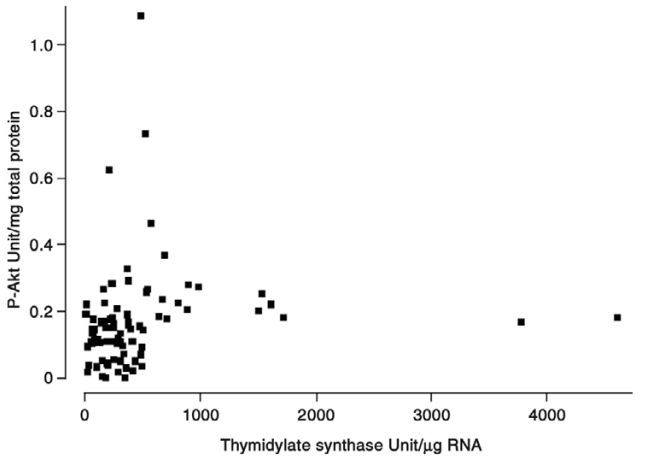
Correlations between elevated P-Akt and higher risk for relapsehas already been demonstrated by other investigators in certain subsets of patients, specifically patients who received adjuvant endocrine therapy [17], patients treated with radiotherapy [18] and patients with a node-negative disease [19]. Because ErbB-2 has been implicated in the activation of Akt [27], we investigated the association between P-Akt and ErbB-2 and its prognostic significance in tumours with known ErbB-2 expression levels. Our investigation re-con-

Figure 3



Kaplan–Meier survival curves for the subset of patients with ErbB-2 overexpressing tumours. The curves stratified by **(a)** median and **(b)** last quartile values of phosphorylated Akt (P-Akt). Patients whose tumours express high levels of P-Akt exhibit a significantly worse outcome in terms of disease-free survival (DFS; $P \leq 0.005$).

Figure 4



Scatter plot of phosphorylated Akt (P-Akt) versus thymidylate synthase (TS) mRNA expression. There is a good positive correlation ($r_s = 0.38$; $P < 0.001$) between the two factors.

firmed the prognostic value of elevated P-Akt levels, and demonstrated that P-Akt expression levels are independent of other prognostic parameters, such as tumour size, grading, and node, ER and ErbB-2 status.

The lack of correlation between protein levels of ErbB-2 and P-Akt may be explained by the fact that Akt is also activated by various receptor tyrosine kinases [5-9], and by G-protein-coupled receptors [28]. Additionally, it was also observed that loss of PTEN activity is frequent in breast cancer and accompanied by increased activation of Akt [29], confirming that Akt can be activated by stimuli other than ErbB-2. The prognostic significance of P-Akt levels is increased if combined with ErbB-2 overexpression, suggesting that coactivation of Akt and ErbB-2 may have a synergistic clinical impact.

Our study is the first report on P-Akt assessed by EIA using a phospho-specific antibody in breast cancer cytosols of cryopreserved tumour samples; the technique allowed us to obtain precise and quantitative results (for review [30]). In contrast to semiquantitative immunohistochemistry data, tumour marker profiles assessed by quantitative EIA are more sensitive and reproducible. EIA tests conducted with fresh frozen tissue extracts avoid the potential antigen damage due to formalin fixation, paraffin embedding and uncontrolled storage. Furthermore, the two-site (sandwich) CLISA assay used in this investigation ensures increased specificity as compared with single-antibody assays, such as immunohistochemistry and western blotting. In addition, chemiluminometric detection guarantees high sensitivity in the detection of antigen–antibody complex.

We assayed for P-Akt in total breast tumour lysates, and not in tissue samples obtained from microdissection, both because we wished to correlate the protein expression levels of ErbB-2 and P-Akt levels directly in cells extracted from human tumour samples, and because it has been demonstrated that the activation status of Akt varies considerably in tumours of the same histotype, but not between different histotypes of the same tumour [31]. The CLISA assay used in the study was based on homogenized samples, which can include some stromal and normal tissue cells. The STB tissue samples contained at least 60% tumour cells, as observed by the pathologist. In addition, samples were previously analyzed for ErbB-2, ER and PgR using both EIA assays, as well as immunohistochemistry and/or fluorescence *in situ* hybridization. Importantly, good correlation between the assays was observed [23], suggesting that homogenization of samples does not play a crucial role in the final result. As in other assays that measure phosphorylation levels, the role played by phosphatases should not be ignored. We used phosphatase inhibitors in all steps of CLISA, as well as sample dilution. There could be some degradation before P-Akt testing, but all samples were treated identically, and the study compared relative P-Akt levels among all tumours. Reference units (U) were used in order to establish a standard curve, but not to measure absolute P-Akt levels in separate samples.

Also of interest is the positive correlation between P-Akt and mRNA expression levels of tumour proliferation markers shown in the present study. Akt is known to promote cell cycle progression by modulating the expression [32] and stabilization of cyclin D_1 [33], which in turn activates the E2F transcription factor. Our results also reveal a significant correlation of P-Akt with E2F-1 transcription factor expression levels, as well

Table 3

Spearman rank correlation of quantitative P-Akt levels and quantitative mRNA expression levels of proliferation markers

Proliferation marker	r_s	*P*
Thymidylate synthase	0.38	<0.001
Thymidine kinase 1	0.23	<0.01
Survivin	0.22	<0.01
E2F	0.22	<0.01
Topoisomerase IIα	0.19	<0.05

r_s, Spearman correlation coefficient.

as with genes regulated by E2F, such as thymidylate synthase, thymidine kinase 1, survivin and topoisomerase IIα.

Conclusion

Using a highly sensitive and specific CLISA assay, we demonstrated that elevated P-Akt is a marker of poor prognosis (decreased DFS). The prognostic value of Akt phosphorylation is independent of other characteristics, including tumour size and grade, and node, ErbB-2 and ER status. In a subset of patients with ErbB-2 overexpressing tumours, we demonstrated that P-Akt levels are of particular prognostic significance. In addition, Akt phosphorylation correlated with elevated mRNA expression levels of tumour proliferation factors. Based on these findings, we suggest that P-Akt could play a predictive role with respect to Herceptin, topoisomerase IIα inhibitors and combination therapies using Akt inhibitors, which are currently in clinical trials and should primarily be assessed in patients with ErbB-2-overexpressing tumours.

Competing interests

The author(s) delcare that they have no competing interests.

Authors' contributions

JC carried out the development of CLISA assays, took measurements in breast cancer samples, participated in raw data analysis, participated in statistical analysis and drafted the manuscript. PU performed the statistical analysis. VV and ML carried out the RNA extraction and quantitative RT-PCR. WK participated in designing the study and participated in the raw data analysis. EW coordinated clinicians, providing tumour samples. MM helped in finalizing the manuscript. UE participated in designing the study and coordination, and helped to draft the manuscript. SE participated in coordinating the study and in statistical analysis, helped to draft the manuscript and coordinated the selection of samples from the STB. All authors read and approved the final manuscript.

Acknowledgements

This work was supported by a grant Nr. 31-059819.99/1 (U. Eppenberger) of the Swiss National Science Foundation and the Stiftung Tumorbank Basel (STB).

We thank Christine Wullschleger, Francoise David, Heidi Bodmer and Sabine Ehret for technical assistance, data management and tumour banking. We are indebted to A Almendral, M Anabitare, C Braschler, B von Castelberg, H Dieterich, D Fink, R Flury, R Gaudenz, K Lüscher, S Heinzl, M Mihatsch, H Moch, D Oertli, G Sauter, J Torhorst and M Zuber – clinicians and pathologists.

References

1. Vivanco I, Sawyers CL: **The phosphatidylinositol 3-kinase AKT pathway in human cancer.** *Nat Rev Cancer* 2002, **2:**489-501.
2. Hanada M, Feng J, Hemmings BA: **Structure, regulation and function of PKB/AKT: a major therapeutic target.** *Biochim Biophys Acta* 2004, **1697:**3-16.
3. Staal SP: **Molecular cloning of the akt oncogene and its human homologues AKT1 and AKT2: amplification of AKT1 in a primary human gastric adenocarcinoma.** *Proc Natl Acad Sci USA* 1987, **84:**5034-5037.
4. Franke TF, Kaplan DR, Cantley LC: **PI3K: downstream AKTion blocks apoptosis.** *Cell* 1997, **88:**435-437.
5. Bowers DC, Fan S, Walter KA, Abounader R, Williams JA, Rosen EM, Laterra J: **Scatter factor/hepatocyte growth factor protects against cytotoxic death in human glioblastoma via phosphatidylinositol 3-kinase- and AKT-dependent pathways.** *Cancer Res* 2000, **60:**4277-4283.
6. Hii CS, Moghadammi N, Dunbar A, Ferrante A: **Activation of the phosphatidylinositol 3-kinase-Akt/protein kinase B signaling pathway in arachidonic acid-stimulated human myeloid and endothelial cells: involvement of the ErbB receptor family.** *J Biol Chem* 2001, **276:**27246-27255.
7. Chen Y, Li X, Eswarakumar VP, Seger R, Lonai P: **Fibroblast growth factor (FGF) signaling through PI 3-kinase and Akt/PKB is required for embryoid body differentiation.** *Oncogene* 2000, **19:**3750-3756.
8. Duan C, Liimatta MB, Bottum OL: **Insulin-like growth factor (IGF)-I regulates IGF-binding protein-5 gene expression through the phosphatidylinositol 3-kinase, protein kinase B/Akt, and p70 S6 kinase signaling pathway.** *J Biol Chem* 1999, **274:**37147-37153.
9. Franke TF, Yang SI, Chan TO, Datta K, Kazlauskas A, Morrison DK, Kaplan DR, Tsichlis PN: **The protein kinase encoded by the Akt proto-oncogene is a target of the PDGF-activated phosphatidylinositol 3-kinase.** *Cell* 1995, **81:**727-736.
10. Cantley LC, Neel BG: **New insights into tumor suppression: PTEN suppresses tumor formation by restraining the phosphoinositide 3-kinase/AKT pathway.** *Proc Natl Acad Sci USA* 1999, **96:**4240-4245.
11. Miwa W, Yasuda J, Murakami Y, Yashima K, Sugano K, Sekine T, Kono A, Egawa S, Yamaguchi K, Hayashizaki Y, Sekiya T: **Isolation of DNA sequences amplified at chromosome 19q13.1-q13.2 including the AKT2 locus in human pancreatic cancer.** *Biochem Biophys Res Commun* 1996, **225:**968-974.
12. Cheng JQ, Godwin AK, Bellacosa A, Taguchi T, Franke TF, Hamilton TC, Tsichlis PN, Testa JR: **AKT2, a putative oncogene encoding a member of a subfamily of protein-serine/threonine kinases, is amplified in human ovarian carcinomas.** *Proc Natl Acad Sci USA* 1992, **89:**9267-9271.

13. Bellacosa A, de Feo D, Godwin AK, Bell DW, Cheng JQ, Altomare DA, Wan M, Dubeau L, Scambia G, Masciullo V, *et al.*: **Molecular alterations of the AKT2 oncogene in ovarian and breast carcinomas.** *Int J Cancer* 1995, **64:**280-285.
14. Roy HK, Olusola BF, Clemens DL, Karolski WJ, Ratashak A, Lynch HT, Smyrk TC: **AKT proto-oncogene overexpression is an early event during sporadic colon carcinogenesis.** *Carcinogenesis* 2002, **23:**201-205.
15. Ruggeri BA, Huang L, Wood M, Cheng JQ, Testa JR: **Amplification and overexpression of the AKT2 oncogene in a subset of human pancreatic ductal adenocarcinomas.** *Mol Carcinog* 1998, **21:**81-86.
16. Bacus SS, Altomare DA, Lyass L, Chin DM, Farrell MP, Gurova K, Gudkov A, Testa JR: **AKT2 is frequently upregulated in HER-2/neu-positive breast cancers and may contribute to tumor aggressiveness by enhancing cell survival.** *Oncogene* 2002, **21:**3532-3540.
17. Perez-Tenorio G, Stal O: **Activation of AKT/PKB in breast cancer predicts a worse outcome among endocrine treated patients.** *Br J Cancer* 2002, **86:**540-545.
18. Stal O, Perez-Tenorio G, Akerberg L, Olsson B, Nordenskjold B, Skoog L, Rutqvist LE: **Akt kinases in breast cancer and the results of adjuvant therapy.** *Breast Cancer Res* 2003, **5:**R37-R44.
19. Schmitz KJ, Otterbach F, Callies R, Levkau B, Holscher M, Hoffmann O, Grabellus F, Kimmig R, Schmid KW, Baba HA: **Prognostic relevance of activated Akt kinase in node-negative breast cancer: a clinicopathological study of 99 cases.** *Mod Pathol* 2004, **17:**15-21.
20. Malik SN, Brattain M, Ghosh PM, Troyer DA, Prihoda T, Bedolla R, Kreisberg JI: **Immunohistochemical demonstration of phospho-Akt in high Gleason grade prostate cancer.** *Clin Cancer Res* 2002, **8:**1168-1171.
21. Liao Y, Grobholz R, Abel U, Trojan L, Michel MS, Angel P, Mayer D: **Increase of AKT/PKB expression correlates with gleason pattern in human prostate cancer.** *Int J Cancer* 2003, **107:**676-680.
22. Yamamoto S, Tomita Y, Hoshida Y, Morooka T, Nagano H, Dono K, Umeshita K, Sakon M, Ishikawa O, Ohigashi H, Nakamori S, Monden M, Aozasa K: **Prognostic significance of activated Akt expression in pancreatic ductal adenocarcinoma.** *Clin Cancer Res* 2004, **10:**2846-2850.
23. Eppenberger-Castori S, Kueng W, Benz C, Caduff R, Varga Z, Bannwart F, Fink D, Dieterich H, Hohl M, Muller H, Paris K, Schoumacher F, Eppenberger U: **Prognostic and predictive significance of ErbB-2 breast tumor levels measured by enzyme immunoassay.** *J Clin Oncol* 2001, **19:**645-656.
24. Chambers JM, Hastie TJ: *Statistical Models in S* London: Chapman & Hall; 1971:414.
25. Breimann L, Friedman JH, Olsen RA, Stone CJ: *Classification and Regression Trees* Wadsworth: Belmont; 1984.
26. Mantel N: **Evaluation of survival data and the two new rank order statistics arising its consideration.** *Cancer Chemother Rep* 1966, **50:**163-170.
27. Hellyer NJ, Kim MS, Koland JG: **Heregulin-dependent activation of phosphoinositide 3-kinase and Akt via the ErbB2/ErbB3 co-receptor.** *J Biol Chem* 2001, **276:**42153-42161.
28. Murga C, Laguinge L, Wetzker R, Cuadrado A, Gutkind JS: **Activation of Akt/protein kinase B by G protein-coupled receptors. A role for alpha and beta gamma subunits of heterotrimeric G proteins acting through phosphatidylinositol-3-OH kinase gamma.** *J Biol Chem* 1998, **273:**19080-19085.
29. Mills GB, Lu Y, Fang X, Wang H, Eder A, Mao M, Swaby R, Cheng KW, Stokoe D, Siminovitch K, Jaffe R, Gray J: **The role of genetic abnormalities of PTEN and the phosphatidylinositol 3-kinase pathway in breast and ovarian tumorigenesis, prognosis, and therapy.** *Semin Oncol* 2001:125-141.
30. Ross JS, Fletcher JA: **HER-2/neu (c-erb-B2) gene and protein in breast cancer.** *Am J Clin Pathol* 1999:S53-S67.
31. Wulfkuhle JD, Aquino JA, Calvert VS, Fishman DA, Coukos G, Liotta LA, Petricoin EF 3rd: **Signal pathway profiling of ovarian cancer from human tissue specimens using reverse-phase protein microarrays.** *Proteomics* 2003, **3:**2085-2090.
32. Muise-Helmericks RC, Grimes HL, Bellacosa A, Malstrom SE, Tsichlis PN, Rosen N: **Cyclin D expression is controlled post-transcriptionally via a phosphatidylinositol 3-kinase/Akt-dependent pathway.** *J Biol Chem* 1998, **273:**29864-29872.
33. Graff JR, Konicek BW, McNulty AM, Wang Z, Houck K, Allen S, Paul JD, Hbaiu A, Goode RG, Sandusky GE, *et al.*: **Increased AKT activity contributes to prostate cancer progression by dramatically accelerating prostate tumor growth and diminishing p27Kip1 expression.** *J Biol Chem* 2000, **275:**24500-24505.

EUROPEAN JOURNAL OF CANCER 42 (2006) 636–645

available at www.sciencedirect.com

SCIENCE DIRECT®

journal homepage: www.ejconline.com

Phosphorylation of tyrosine 1248-ERBB2 measured by chemiluminescence-linked immunoassay is an independent predictor of poor prognosis in primary breast cancer patients

Jonas Cicenas[a,c], *Patrick Urban*[a,c], *Willy Küng*[c], *Vincent Vuaroqueaux*[b], *Martin Labuhn*[b], *Edward Wight*[d], *Urs Eppenberger*[b,c], *Serenella Eppenberger-Castori*[a,*]

[a]*Stiftung Tumorbank Basel, Lörracherstrasse 50, CH-4125 Riehen, Switzerland*
[b]*OncoScore AG, Riehen, Switzerland*
[c]*Molecular Tumor Biology, Department of Research, University Basel, Basel, Switzerland*
[d]*University Clinics, Kantonsspital Basel, Switzerland*

ARTICLE INFO

Article history:
Received 29 April 2005
Received in revised form
31 October 2005
Accepted 28 November 2005
Available online 18 January 2006

Keywords:
Phosphorylation
ERBB2 (HER2/Neu)
Immunoassay
Biomarker
Prognosis
Breast cancer

ABSTRACT

ERBB2 (HER2/Neu) gene amplification and overexpression is associated with increased risk of metastases and shorter survival in breast cancer. Tyrosine 1248 is a major phosphorylation site of ERBB2 and reflects the activation status of the receptor. The aim of this study was to investigate the relationships between quantitative levels of pY1248-ERBB2 (p-ERBB2) and the expression of epidermal growth factor receptor (EGFR)-family members, and whether p-ERBB2 could provide additional prognostic value compared with established prognostic markers. For this purpose we developed a highly sensitive chemiluminescence-linked immunoassay (CLISA) and detected p-ERBB2 levels in 70 primary breast cancer biopsies. Phosphorylated ERBB2 correlated with EGFR and ERBB2, and inversely with oestrogen receptor (ER), progesterone receptor (PgR) and ERBB4 expression levels. Additionally, p-ERBB2 was associated with poor clinical outcome in univariate and multivariate Cox regression analysis. Further studies are needed to evaluate the predictive value of p-ERBB2.
© 2005 Elsevier Ltd. All rights reserved.

1. Introduction

ERBB2 (HER2/Neu) is a member of the epidermal growth factor receptor (EGFR) family of receptor tyrosine kinases, which comprises EGFR (HER1, ERBB1), ERBB3 (HER3) and ERBB4 (HER4).[1–3] Upon ligand binding, homo- and heterodimeric complexes are formed, with ERBB2 as the preferred dimerisation partner.[4] This leads to autophosphorylation of specific tyrosine residues, activation of downstream signalling cascades and, finally, initiation of biological processes such as proliferation.[5] ERBB2 contains five major tyrosine autophosphorylation sites, including Y-1248.[6]

In primary breast cancer, ERBB2 is amplified and overexpressed in 15–30% of patients and has been associated with poor prognosis.[7–9] Trastuzumab (Herceptin™), a humanised monoclonal anti-ERBB2 antibody is the first clinically available oncogene-targeted therapeutic agent for treatment of solid tumours, and is approved for use in metastatic breast cancer patients.[10] First-line trastuzumab in combination with chemotherapy resulted in a 25% improvement in overall

* *Corresponding author:* Tel.: +41 61 645 9304; fax: +41 61 645 9308.
E-mail address: s.eppenberger@tumorbank.org (S. Eppenberger-Castori).
0959-8049/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2005.11.012

survival compared with chemotherapy alone. However, only up to 40% of patients respond to the therapy, suggesting that more accurate biomarkers are required to identify patients who are likely to respond to treatment such as trastuzumab. It was also reported that an inverse relationship exists between oestrogen receptor (ER) and ERBB2 expression, where ERBB2 overexpression is associated with decreased ER/progesterone receptor (PgR) levels and reduced sensitivity, possibly even resistance to endocrine therapy.[9,11,12]

ERBB2 gene amplification or overexpression per se may not reflect adequately the activated status of the ERBB2 receptor. It was hypothesised that the percentage of phosphorylated ERBB2, and thus activated receptor, could be different between tumours expressing similar amounts of ERBB2.[13]

The aim of the present study was to investigate the prognostic value of pY1248-ERBB2 detected with a newly developed chemiluminescence-linked immunoassay (CLISA), its association with protein and mRNA expression levels of the EGFR-family members including established prognostic markers in a set of 70 primary breast cancer patients.

2. Patients and methods

2.1. Patients and tumour characteristics

For all tumour samples the Stiftung Tumorbank Basel (STB) received a representative piece of fresh frozen tissue containing more than 65% tumour cells after surgery and pathological examination. Specimens were immediately processed or cryopreserved (−80 °C). For this study, 70 primary breast tumour samples were selected according to ERBB2 protein expression levels detected by enzyme immunoassay (EIA) at time of surgery. Tumours with ERBB2 protein levels >260 ng/mg total protein were considered positive, which corresponds to a previously published cut-off value of 500 U/mg total protein and correlates with the immunohistochemistry (IHC) DAKO 3+[8] as well as ERBB2 amplification detected by fluorescence in situ hybridisation (FISH) (Urban P, et al., submitted). ERBB2-negative tumours showed protein expression levels between 100 and 260 ng/mg. All patients underwent primary surgery before January 1996. Twenty-four patients (34%) relapsed within the median follow-up time of 55 months (range 30–89 months). Thirty-seven (53%) were nodal-positive, 50 (71%) were ER-positive and 40 (57%) patients were ERBB2-positive. None of the patients received neoadjuvant therapy.

STB is a non-profit organisation with an official Swiss permit that guarantees ethical issues and patient confidentiality. Patients and tumour characteristics are summarised in Table 1.

Table 1 – Tumour and patients characteristics

Feature	Number of patients (%)
Patients	70 (100)
Histology type	
Ductal	48 (69)
Lobular	11 (16)
Other	11 (16)
Tumour size	
T1	18 (26)
T2	42 (60)
T3–4	10 (14)
Lymph node status	
Node-negative	33 (47)
Node-positive	37 (53)
Histopathological grade	
I + II	27 (39)
III	34 (48)
Not analysed	9 (13)
Oestrogen receptor	
Positive (>20 fmol/mg)	50 (71)
Negative (⩽20 fmol/mg)	20 (29)
Median/mean (fmol/mg)	72/139
Progesterone receptor	
Positive (>20 fmol/mg)	38 (54)
Negative (⩽20 fmol/mg)	32 (46)
Median/mean (fmol/mg)	28/128
ERBB2	
Positive (>260 ng/mg)	40 (57)
Negative (<260 ng/mg)	30 (43)
Median/mean (ng/mg)	307/298

2.2. Cell lines and tissue culture

SKBr3 breast cancer cells were cultured in improved minimal essential medium with zinc option (IMEM-ZO) supplemented with 5% foetal bovine serum (FBS) and L-glutamine at 37 °C in a 5% CO_2 incubator. For the phospho-standard preparation sub-confluent SKBr3 cells were serum-starved for 48 h in serum-free medium, treated with NaF and Na_3VO_4 for 1 h, then with 10% FBS for 10 min. Cells were lysed in EB lysis buffer (0.5 M NaCl, 10 mM EDTA, pH 8, 1% Triton × 100, 20 mM Tris–Cl, pH 7.0, 20 mM NaF, 20 mM glycerophosphate, 2 mM Na_3VO_4, proteinase inhibitor cocktail, Roche) for 5 min on ice, centrifuged at 20,000*g* for 5 min and the supernatant stored at −80 °C.

2.3. Measurement of ER, PgR, ERBB2 and EGFR protein levels

Tissue homogenates were prepared in accordance with standard procedures for tumour marker EIA measurement, as described previously.[8] In brief, frozen tissues were powderised in liquid nitrogen (Micro-Dismembrator U, B. Braun AG, Melsungen, Germany) and homogenised (tissue homogeniser, Ultra-Turrax; Janke and Kunkel, IKA-Werke, Staufen, Germany) for 20 s in three volumes of ice-cold extraction buffer. The homogenate was centrifuged at 800*g* for 30 min at 2 °C, and the resulting supernatant recentrifuged in an ultracentrifuge (Beckman Instruments, Fullerton, CA, United States of America (USA)) at 100,000*g*. The resulting supernatants (cytosols) were used for measurement of the hormone receptors (ER, PgR by Abbott Laboratories, Abbott Park, IL, USA), while the membrane fractions were used for EIA measurement of ERBB2 (Oncogene Science Human HER-2/neu Quantitative ELISA Kit, Bayer, Leverkusen, Germany). Quantification of EGFR was done by radioligand binding assay (LBA) as described previously.[14] Quality control of ER and PgR measurements were carried out in collaboration with the Receptor

Biomarker Group of the European Organisation for Research and Treatment of Cancer (EORTC).

2.4. *pY1248-ERBB2 immunoassay*

p-ERBB2 levels were measured with a two-site CLISA. Black 96-well microtitre plates (Nunc Black MaxiSorp Surface; Nalgen Nunc International, Rochester, NY, USA) were coated with antihuman activated *Neu/c-ERBB2* antibody (#06-229, lot #15916; Upstate Biotechnology, Lake Placid, NY) at a concentration of 4 mg/ml of coating buffer (phosphate buffered saline (PBS) with 0.6 mM EDTA) in a volume of 100 μl/well and kept at 4 °C overnight. This antibody is virtually identical to the monoclonal antibody clone PN2A that recognises only pY1248-ERBB2.[13,15] Tumour extracts were prepared in the presence of Na_3VO_4. Before sample application, coated microtitre plates were washed five times with 200 μl/well of washing buffer (25 mM HEPES, pH 7.4, 300 mM NaCl, 0.05% Tween-20) and then blocked for 2 h at room temperature with 250 μl blocking buffer (25 mM HEPES, pH 7.4, 300 mM NaCl, 0.05% Tween-20, 3% TopBlock, Juro AG, Switzerland). The blocked wells were washed five times with 300 ml blocking buffer. Then 100 μl of the diluted membrane extracts or reference material was added to the wells and incubated overnight at 4 °C. SKBr3 cell extract was used as reference for each assay as described above. First, SKBr3 cell membrane extract was sequentially diluted with sample dilution buffer at ratios of 1×, 0.75×, 0.5×, 0.25×, 0.125× and 0.025×. Subsequently, 100 μl aliquots were incubated on each microtitre plate together with the tumour extracts or controls (dilution buffer only). After incubation, wells were washed five times with 300 μl washing buffer at room temperature to eliminate unbound particles. Biotinylated detection antibody (HER-2/Neu Microtiter ELISA kit, Oncogene Science) was added to the wells, incubated for 2 h at room temperature and complex detected with horseradish peroxidase (HRP)-conjugated streptavidin using SuperSignal WestPico substrate (Pierce) in a glow luminometer. A curve was fitted to the data of the reference dilution series and used for quantification of tumour extracts. The value of the undiluted SKBr3 extract was denominated as 100 U/ml. The standard curve and additional specification are shown in Fig. 1.

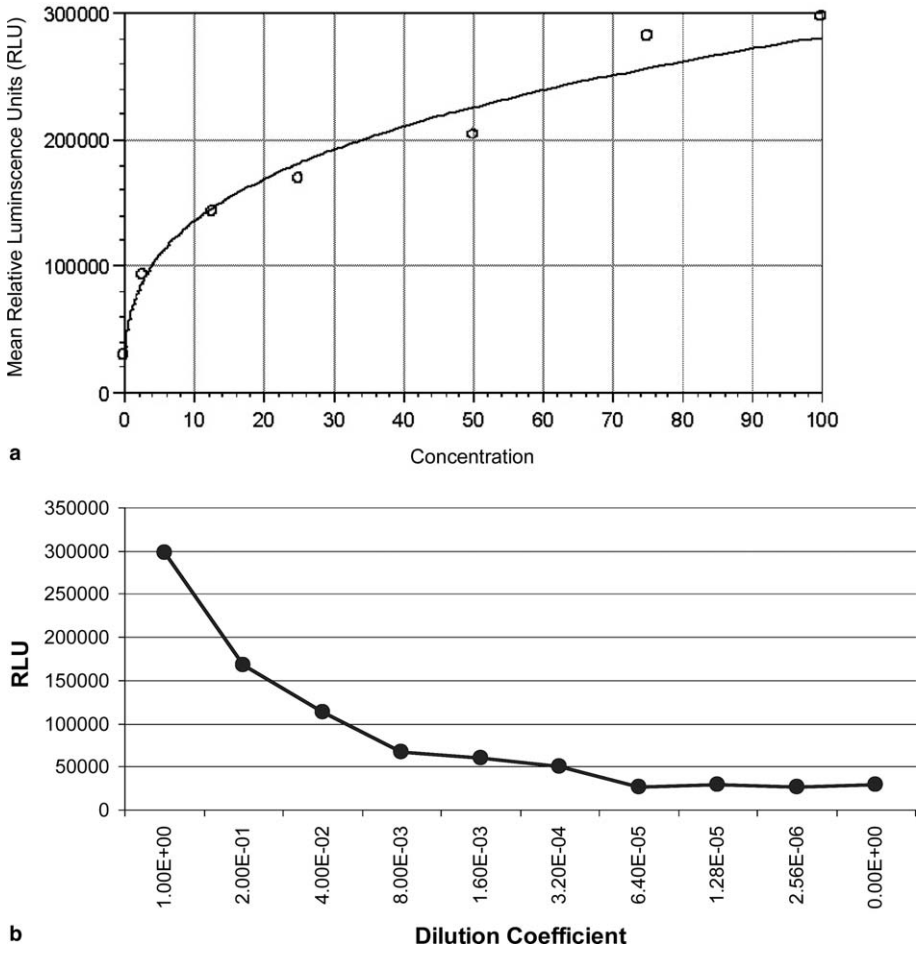


Fig. 1 – Supplementary assay specifications: (a) standard curve for pY1248-ERBB2 immunoassay as described in Section 2. A concentration of 100 (*x*-axis) corresponds to 100 μl undiluted standard (1×), a concentration of 50 to a dilution of 0.5×, etc. Standard curve was fitted using SoftMax software (Molecular Devices, CA, USA). Limit of detection (LOD) and limit of quantification (LOQ) was calculated from repeated measurements without analyte. LOD (mean + 2SDs) value was 31088 RLU or 0.07 U/ml, and LOQ (mean + 10SDs) was 39203 RLU or 0.41 U/ml, respectively. The undiluted value of SKBr3 standard was denominated 100 U/ml; (b) dilution curve for SKBr3 standard (see Section 2).

Table 2 – Primer sequences

Gene	RefSeq	Forward primer	Reverse primer
ESR1	NM_000125	CTTGCTCTTGGACAGGAACCA	CAAACTCCTCTCCCTGCAGATT
PgR	NM_000926	TGTCGAGCTCACAGCGTTTC	TACAGATGAAGTTGTTTGACAAGATCA
EGFR	NM_005228	GGACTATGTCCGGGAACACAA	CCAAGTAGTTCATGCCCTTTGC
ERBB2	NM_004448	CTGAACTGGTGTATGCAGATTGC	TTCCGAGCGGCCAAGTC
ERBB3	NM_001982	AATAAAAGGGCTATGAGGCGATACT	AGCTTCCTTAGCTCTGTCTCTTTGA
ERBB4	NM_005235	GTCCAGATAGCTAAGGGAATGATGTAC	CTAGCCCAAAATCTGTGATTTTCAC

PgR, progesterone receptor; EGFR, epidermal growth factor receptor.

2.5. *Quantitative real-time PCR (Qrt-PCR) for ER, PgR and EFGR-family*

Total RNA was extracted using RNeasy kit (Qiagen, Hilden, Germany). RNA quality and quantity was checked for all samples on Bioanalyzer 2100 (Agilent, Palo Alto, CA, USA). All genes were measured using SYBR Green I method and a Taqman 7000 (Applied-Biosystems, Foster City, CA, USA). Relative quantification (delta ct-value) was obtained by normalisation to the ribosomal 18S gene, and by a standardisation step using Human Universal Standard RNA (Stratagene, La Jolla, CA, USA). Quantitative real-time polymerase chain reactino (Qrt-PCR) results were expressed in arbitrary units of reverse transcribed RNA (U/ug rt-RNA). Primer sequences are listed in Table 2.

2.6. *Statistical methods*

Correlations between continuous values were assessed by the Spearman rank correlation coefficient (r_s). Statistical significance between p-ERBB2 and dichotomous variables was calculated using Mann–Whitney *U* test. In the present study patients were dichotomised and defined as p-ERBB2-negative/positive according to median value of p-ERBB2. Alternatively, an optimised cut-off value for p-ERBB2 with respect to prognosis was searched for by classification and regression trees (CART) analysis.[16] Relationships between categorical data were assessed using Fisher's exact test. The prognostic significance between p-ERBB2 and other variables was tested in univariate and multivariate Cox regression analysis and likelihood ratio test. Hazard rates and confidence intervals (CIs) were summarised in tables. Survival curves were estimated by the Kaplan–Meier method and statistical significance compared by means of log-rank test. All tests were performed using S-PLUS statistical software (Insightful, Version 6).

3. Results

3.1. *Distribution of p-ERBB2 and association with ERBB2*

p-ERBB2 levels quantified by CLISA ranged from 0 to 127 U/mg total protein, with a median of 1.096 U/mg (mean 7.49 U/mg). After log-transformation, values became almost normally distributed (Fig. 2). Correlation between p-ERBB2 and expression levels of ERBB2 was 0.62 for mRNA and 0.53 for protein (Fig. 3 and Table 3). p-ERBB2 levels were significantly higher

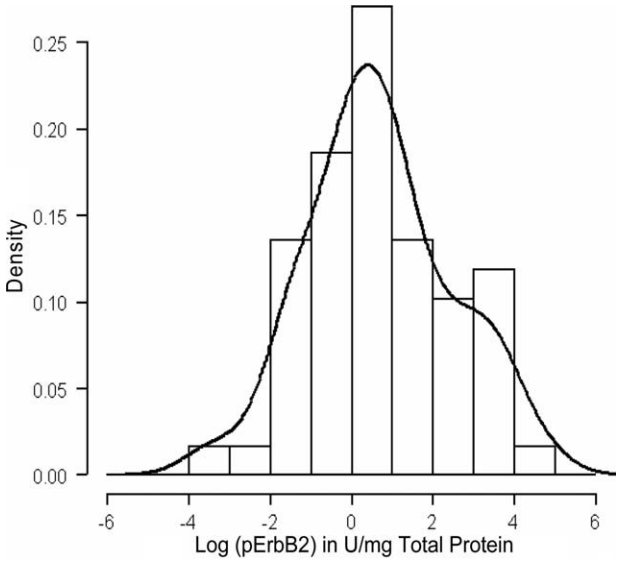


Fig. 2 – Distribution of log (p-ERBB2) levels in U/mg total protein; zero values were omitted.

in ERBB2-positive tumours than in ERBB2-negative tumours (Fig. 3). When taking the median value of p-ERBB2 as a cut-off for p-ERBB2 status, 8 out of 30 (27%) ERBB2-negative and 27 out of 40 (68%) ERBB2-positive tumours were p-ERBB2-positive ($P = 0.001$). Similar results were obtained when p-ERBB2 levels were compared with ERBB2 mRNA expression levels (Fig. 4). Of note, ERBB2-negative tumours included in the present study represent the upper third of all ERBB2-negative tumours (Fig. 5). No significant association was found in p-ERBB2 levels with respect to nodal status, either when analysing it as a continuous variable or after dichotomisation (Fig. 3).

3.2. *p-ERBB2 and expression of ER/PgR*

Quantitative p-ERBB2 expression levels correlated inversely with ER ($r_s = -0.54$) and PgR ($r_s = -0.46$) mRNA as well as protein expression level ($r_s = -0.67$ and $r_s = -0.45$, Table 3), and there was a significant difference in ER and PgR protein levels in p-ERBB2-negative as compared with p-ERBB2-positive tumours: median ER levels were almost 6-fold ($P < 0.001$) and PgR near 7-fold ($P = 0.005$) lower in p-ERBB2-positive tumours (Fig. 6). With respect to the hormone receptor status, 30 out of 35 (86%) p-ERBB2-negative tumours were ER-positive

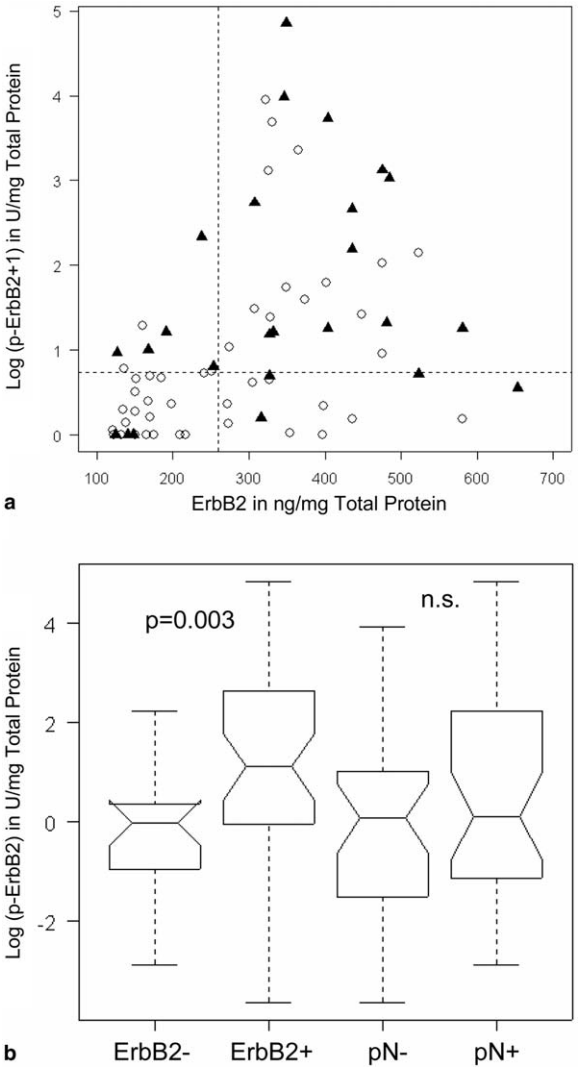


Fig. 3 – Scatter plot of p-ERBB2 (CLISA) versus ERBB2 (ELISA) protein expression levels. p-ERBB2 expression levels were log-transformed and one added as constant in order log-transformed values of zero correspond to zero values in the raw data. Tumours with ERBB2 >260 ng/mg and p-ERBB2 > median are considered positive (dashed lines): (a) circles and triangles indicate patients without and with relapse, respectively; (b) significant differences of p-ERBB2 levels in ERBB2-negative and ERBB2-positive tumours but no difference with respect to nodal status.

Table 3 – Correlation between p-ERBB2 levels and other quantitatively assessed markers (Spearman correlation coefficient r_s)

Correlation with	mRNA expression	Protein expression
ER	−0.54 ($P < 0.001$)	−0.67 ($P < 0.001$)
PgR	−0.46 ($P < 0.001$)	−0.45 ($P < 0.001$)
EGFR	0.26 ($P = 0.049$)	0.43 ($P = 0.005$)
ERBB2	0.62 ($P < 0.001$)	0.53 ($P < 0.001$)
ErBB3	−0.22 ($P = 0.080$)	n.d.
ErBB4	−0.47 ($P < 0.001$)	n.d.

n.d., not determined; ER, oestrogen receptor; PgR, progesterone receptor; EGFR, epidermal growth factor receptor.

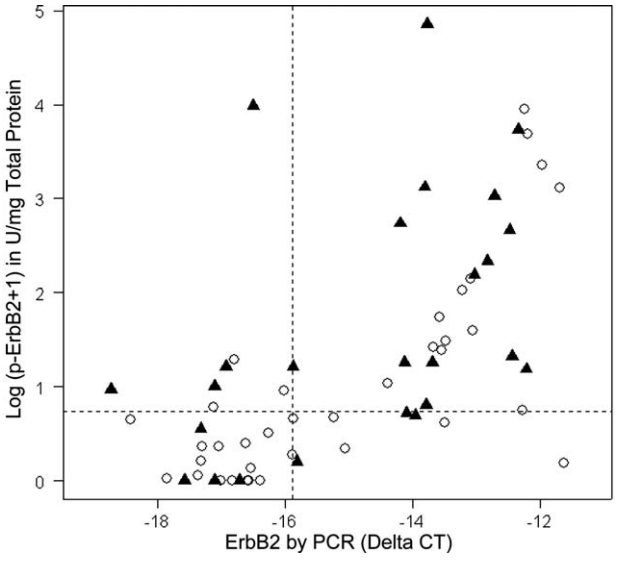


Fig. 4 – Scatter plot of p-ERBB2 (CLISA) versus ERBB2 expression levels measured by polymerase chain reaction (PCR) (Delta CT). p-ERBB2 expression levels were log-transformed and one added as constant so that log-transformed values of zero correspond to zero values in the raw data. Dashed lines indicate the cut-off value for ERBB2 status by PCR (M. Labuhn et al., submitted) and median expression value of p-ERBB2, respectively. Circles and triangles indicate patients without and with relapse, respectively.

(>20 fmol/mg) and 24 (69%) PgR-positive (>20 fmol/mg). In contrast, only 20 out of 35 (57%) p-ERBB2-positive tumours were ER-positive and 22 (43%) PgR-positive ($P = 0.016$), respectively.

3.3. *Correlation of p-ERBB2 with mRNA expression of EGFR-family members, and EGFR and ERBB2 protein expression*

p-ERBB2 levels were correlated with quantitative mRNA expression levels of all four EGFR-family members (Table 3). p-ERBB2 was positively correlated with EGFR ($r_s = 0.26$) and ERBB2 ($r_s = 0.62$) and inversely correlated with ERBB-4 ($r_s = -0.47$). A negative correlation between p-ERBB2 and ERBB3 mRNA was not statistically significant. Similar results were obtained for protein expression level of EGFR ($r_s = 0.43$) and ERBB2 ($r_s = 0.53$, Table 3). Distributions and statistical differences of the median mRNA expression levels for all EGRF-family members in p-ERBB2-negative versus p-ERBB2-positive tumours are summarised in Fig. 6. ERBB2 expression was significantly higher whereas ERBB4 expression was significantly lower in p-ERBB2-positive patients. Despite their correlation with p-ERBB2, EGFR and ERBB3 were not significantly differentially expressed between p-ERBB2-positive and -negative samples.

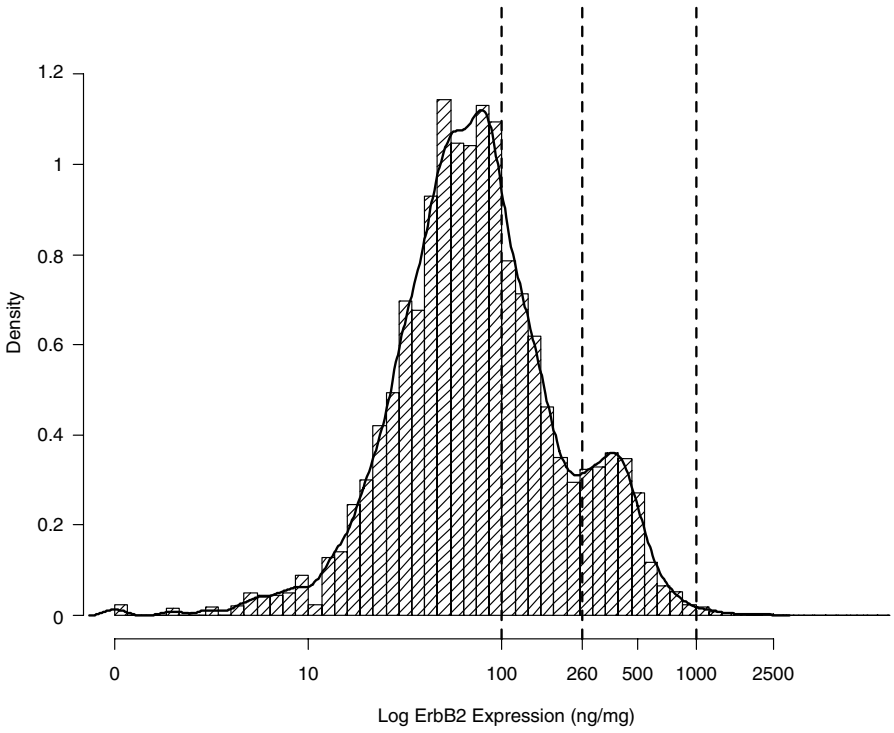


Fig. 5 – Distribution of ERBB2 protein expression levels (enzyme-linked immunosorbent assay (ELISA)) in over 3200 primary breast cancer patients.

3.4. *Prognostic significance of p-ERBB2*

p-ERBB2 was first tested in univariate Cox regression analysis, revealing significant correlation with patient disease-free survival (DFS) and overall survival (OS, Table 4). Notably, p-ERBB2 retained significant prognostic value in univariate Cox regression analysis in both ERBB2-negative and ERBB2-positive groups of patients. Subsequently, Kaplan–Meier survival analysis was performed for p-ERBB2-negative versus p-ERBB2-positive tumours (Fig. 7). Eighteen out of 35 (51%) patients with p-ERBB2-positive tumours developed disease recurrence, whereas this was the case for only 7 out of 35 (20%) in the p-ERBB2-negative group ($P = 0.004$). Five-year DFS was 45% (CI 31–67%) in the p-ERBB2-positive group and 82% (CI 70–96%) in the p-ERBB2-negative group. In OS analysis 9 out of 35 (26%) patients with p-ERBB2-positive tumours died compared with 4 out of 35 (11%) in p-ERBB2-negative tumours ($P = 0.079$, Fig. 7). Differences in survival remained significant when stratified according to nodal status. However, the results were more significant in nodal-positive than nodal-negative patients (Fig. 7). Moreover, p-ERBB2 status was significantly associated with DFS in ERBB2-negative tumours (Fig. 7). Five out of 8 p-ERBB2-positive/ERBB2-negative tumours relapsed. Finally, p-ERBB2 was an independent and significant prognostic factor in multivariate Cox regression analysis, which included ER, EGFR, ERBB2, tumour size (pT), lymph node status (pN) and age. The results are summarised in Table 4.

4. Discussion

This is the first study to measure quantitative levels of activated pY1248-ERBB2 applying a newly developed immunoassay, and investigating the relationship with mRNA and protein expression levels of the EGFR-family, established prognostic markers and survival. Compared with IHC, p-ERBB2 expression levels assessed by CLISA yield quantitative, highly sensitive and reproducible results. In addition, CLISA results obtained from fresh frozen tissue extracts avoid potential antigen damage due to formalin fixation, paraffin embedding and sample storage. However, phosphorylation reflects a dynamic process and potential alterations in phosphorylation levels require careful sample handling and the use of phosphatase inhibitors. Further, a two-site (sandwich) CLISA assay as used in this study ensures increased specificity compared with single-antibody assays such as IHC and Western blotting.

Several studies have investigated the role of phosphorylated ERBB2 in breast tumour samples using IHC.[13,15,17] We could reconfirm the association with ERBB2 status, poor prognosis and the inverse correlation with hormone receptor expression. Although there is good correlation between ERBB2 and its phosphorylation levels, we identified cases with high p-ERBB2 levels in ERBB2-negative tumours. This is in contrast to previously published data by Thor and colleagues,[13] which could not detect p-ERBB2 in ERBB2-negative tumours using IHC. A possible reason for these findings might be differences in assay sensitivity, this being higher in CLISA. We observed 8 (27%) of ERBB2-negative tumours to express p-ERBB2 values above the median, a cut-off that also revealed significant prognostic value in survival analysis. These were unlikely p-ERBB2-positive samples classified falsely as ERBB2-negative using protein-based ERBB2 status since the same number of p-ERBB2-positive cases was found using mRNA-based ERBB2 status determination (Fig. 4). We further report that p-ERBB2

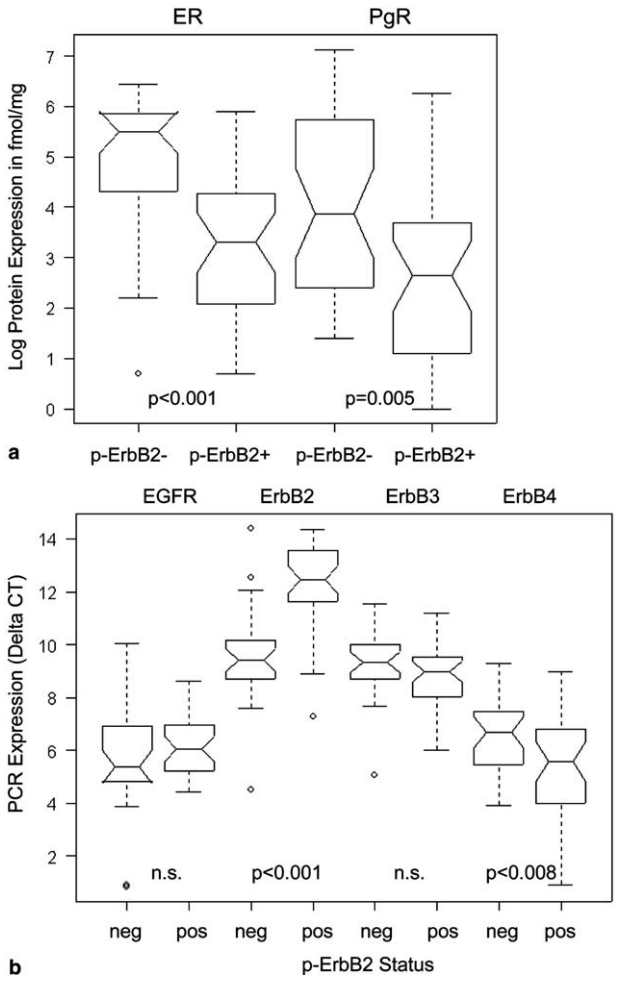


Fig. 6 – Box plot showing the differences in: (a) distribution of ER/PgR protein levels, and (b) mRNA expression levels of the epidermal growth factor receptor (EGFR)-family members in p-ERBB2-negative/-positive tumours.

was significantly associated with survival in both ERBB2-negative and ERBB2-positive patients. Finally, multivariate analysis including ERBB2, EGFR, ER, nodal status, age and tumour size demonstrated independent prognostic value for p-ERBB2, suggesting that p-ERBB2 is providing additional information despite being associated with ERBB2 – a finding which was also recognised by Thor and colleagues.[13] Besides p-ERBB2, nodal status (pN) and tumour size (pT) were independent prognostic markers in multivariate analysis. Grade was not included because of too many missing values.

We used the median value of p-ERBB2 to define p-ERBB2-negative/-positive tumours. Alternatively, we tested an optimised cut-off with respect to prognosis searched for by CART analysis, which classified 33 tumours in the high-risk group instead of 35 when using the median. This optimised cut-off, however, did not change the overall results. It is noteworthy that we selected the samples a priori according to ERBB2 protein expression levels and explicitly enriched the population with ERBB2-positive samples. The high correlation between ERBB2 and p-ERBB2 and the fact that there are approximately as many ERBB2-positive as ERBB2-negative patients in the population studied might explain why using the median as cut-off is reasonable in this setting. However, the number of p-ERBB2-positive tumours is expected to be significantly lower in a randomised situation. Moreover, our ERBB2-negative population had ERBB2 protein expression levels ranging from 100 to 260 ng/mg total protein. This correspond to the upper third of ERBB2 expression levels among ERBB2-negative patients when compared with the distribution of ERBB2 protein expression levels observed in a large study population (Fig. 5[8]). Thus, the percentage of p-ERBB2-positive tumours among ERBB2-negative patients is probably overestimated.

Because EGFR-family receptors have to homo- and/or hetero-dimerise to become activated, we sought to identify potential candidates for the preferred partner for ERBB2 in this process. We report correlation between p-ERBB2 and EGFR expression at both mRNA and protein level, suggesting a potential role for EGFR in ERBB2 phosphorylation and signalling towards a more aggressive phenotype. Indeed, it was proposed that ERBB2 in a heterodimer with EGFR is involved in signalling pathways required for a human breast cancer cell to become metastatic.[18,19] Despite this correlation with EGFR there was no significant difference in EGFR expression levels between p-ERBB2-negative and p-ERBB2-positive tumours. However, we observed that p-ERBB2 expression was significantly higher in tumours having high EGFR protein levels compared with tumours with low EGFR protein levels (Figs. 6 and 8). Conversely, ERBB3, ERBB4 and ER correlated negatively with p-ERBB2 and were expressed at significantly lower

Table 4 – Uni- and multi-variate Cox regression analysis against patient disease-free survival (DFS)

Factor	Univariate		Multivariate	
	P-value	HR (95% CI)	P-value	HR (95% CI)
p-ERBB2	0.004	3.4 (1.4–8.1)	0.010	4.1 (1.4–11)
ERBB2	0.041	1.9 (1.0–3.4)	n.s.	–
EGFR	n.s.	–	n.s.	–
ER	0.012	0.7 (0.6–0.9)	n.s.	–
pT	<0.001	1.9 (1.3–2.7)	0.008	2.1 (1.2–3.7)
pN	0.011	2.0 (1.2–3.3)	0.049	1.9 (1.0–3.7)
Age	n.s.	–	n.s.	–

n.s., not significant; ER, oestrogen receptor; EGFR, epidermal growth factor receptor; HR, hazard ratio; CI, confidence interval; pT, tumour size; pN, lymph node status.

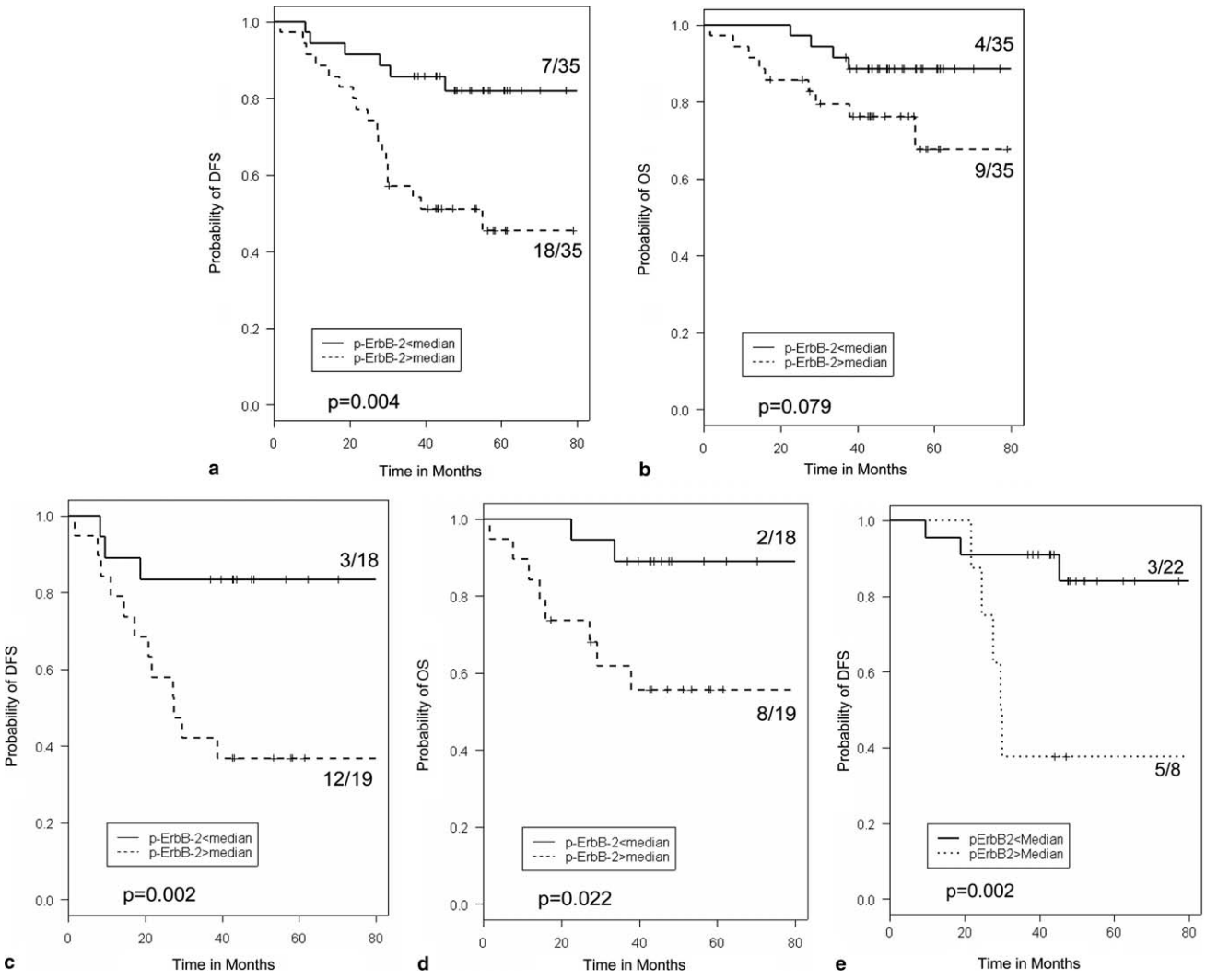


Fig. 7 – Kaplan–Meier survival curves for p-ERBB2. Patients were dichotomised according to the median expression value of p-ERBB2. Disease-free survival (DFS) and overall survival (OS) in (a, b) all patients, in (c, d) nodal-positive patients and (e) DFS in ERBB2-negative patients.

levels in tumours having high levels of p-ERBB2 (Fig. 6). Analogous findings were made in the subset of ERBB2-negative patients (Fig. 9). With respect to ERBB4 mRNA expression, it was shown to be associated with good prognosis in a number of studies,[20–22] whereas there is conflicting data about the role of ERBB3. ERBB3 together with ERBB2 promotes tumour cell proliferation in vitro,[23] but ERBB3 mRNA expression was also shown to correlate with good prognosis.[20,21] The latter findings are in accordance with our results, where ERBB3 and ERBB4 are good prognostic factors (data not shown).

In conclusion, we show that p-ERBB2, measured quantitatively by CLISA, is a marker of poor prognosis independent of ERBB2, ER/PgR, tumour size, nodal status and age. In addition, p-ERBB2 correlated with expression of EGFR and ERBB2, and inversely with ER, PgR, ERBB3 and ERBB4. Hormone receptors and ERBB4 were significantly lower in tumours expressing high p-ERBB2. These findings might be of interest with respect to the selection of appropriate treatment strategies. For trastuzumab, it was demonstrated that tumours expressing p-ERBB2 (IHC) have significantly longer survival when compared with tumours lacking p-ERBB2.[24] The positive correlation between p-ERBB2 and EGFR might suggest that p-ERBB2 may not only be a predictor of trastuzumab response but also predictive of novel EGFR-family targeted treatments such as ERBB2-dimerisation inhibitors and dual-specific tyrosine-kinase inhibitors.[25,26] The significantly lower levels of hormone receptors in p-ERBB2-positive patients might further suggest relative resistance to anti-ER targeted treatment.[9] However, since this study addressed mainly the prognostic value of p-ERBB2 and its relationship with other EGFR-family members, future research should further investigate the predictive value of p-ERBB2.

Conflict of interest statement

The authors declare no financial interest. Quantitative real-time PCR was performed at OncoScore AG.

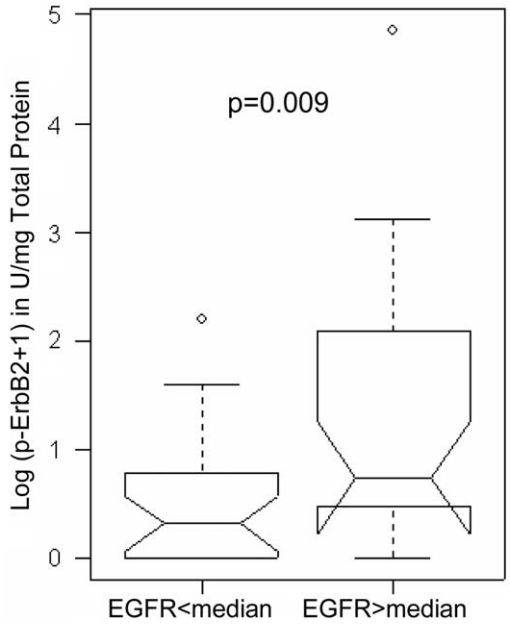


Fig. 8 – Distribution of p-ERBB2 expression levels in tumour expression epidermal growth factor receptor (EGFR) protein levels below and above the median value, respectively. p-ERBB2 expression levels were log-transformed and one added as constant so that log-transformed values of zero correspond to zero values in the raw data.

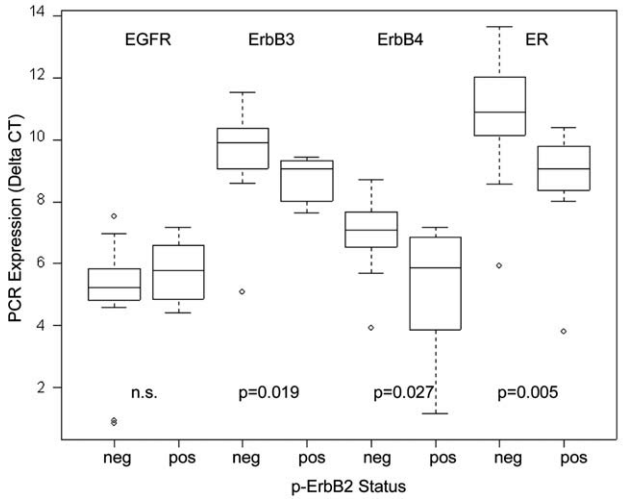


Fig. 9 – Distribution of epidermal growth factor receptor (EGFR)-family members and oestrogen receptor (ER) expression in ERBB2-negative patients with respect to their p-ERBB2 status.

Acknowledgements

This work was supported by a grant number 3100-059819.99/1 (U. Eppenberger) of the Swiss National Science Foundation and the Stiftung Tumorbank Basel (STB).

We are greatly indebted to Dr. Rainer Neumann of Bayer AG for providing some of the antibodies used in this study. We thank Christine Wullschleger, Francoise David, Heidi Bodmer and Sabine Ehret for technical assistance, data management and tumour banking. We are grateful to C. Braschler, B. von Castelberg, H. Dieterich, D. Fink, R. Flury, R. Gaudenz, K. Lüscher, S. Heinzl, M. Mihatsch, H. Moch, D. Oertli, G. Sauter and M. Zuber – clinicians and pathologists.

REFERENCES

1. Schechter AL, Hung MC, Vaidyanathan L, et al. The neu gene: an erbB-homologous gene distinct from and unlinked to the gene encoding the EGF receptor. *Science* 1985;**229**:976–8.
2. Bargmann CI, Hung MC, Weinberg RA. Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185. *Cell* 1986;**45**: 649–57.
3. Yarden Y, Sliwkowski MX. Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2001;**2**:127–37.
4. Tzahar E, Waterman H, Chen X, et al. A hierarchical network of interreceptor interactions determines signal transduction by Neu differentiation factor/neuregulin and epidermal growth factor. *Mol Cell Biol* 1996;**16**:5276–87.
5. Schlessinger J. Cell signaling by receptor tyrosine kinases. *Cell* 2000;**103**:211–25.
6. Ullrich A, Schlessinger J. Signal transduction by receptors with tyrosine kinase activity. *Cell* 1990;**61**:203–12.
7. Slamon DJ, Clark GM, Wong SG, et al. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. *Science* 1987;**235**:177–82.
8. Eppenberger-Castori S, Kueng W, Benz C, et al. Prognostic and predictive significance of ErbB-2 breast tumor levels measured by enzyme immunoassay. *J Clin Oncol* 2001;**19**:645–56.
9. Konecny G, Pauletti G, Pegram M, et al. Quantitative association between HER-2/neu and steroid hormone receptors in hormone receptor-positive primary breast cancer. *J Natl Cancer Inst* 2003;**95**:142–53.
10. Leyland-Jones B. Trastuzumab: hopes and realities. *Lancet Oncol* 2002;**3**:137–44.
11. Osborne CK, Bardou V, Hopp TA, et al. Role of the estrogen receptor coactivator AIB1 (SRC-3) and HER-2/neu in tamoxifen resistance in breast cancer. *J Natl Cancer Inst* 2003;**95**:353–61.
12. De Laurentiis M, Arpino G, Massarelli E, et al. A meta-analysis on the interaction between HER-2 expression and response to endocrine treatment in advanced breast cancer. *Clin Cancer Res* 2005;**11**:4741–8.
13. Thor AD, Liu S, Edgerton S, et al. Activation (tyrosine phosphorylation) of ErbB-2 (HER-2/neu): a study of incidence and correlation with outcome in breast cancer. *J Clin Oncol* 2000;**18**:3230–9.
14. Costa S, Stamm H, Almendral A, et al. Predictive value of EGF receptor in breast cancer. *Lancet* 1988;**2**:1258.
15. DiGiovanna MP, Stern DF. Activation state-specific monoclonal antibody detects tyrosine phosphorylated p185neu/erbB-2 in a subset of human breast tumors overexpressing this receptor. *Cancer Res* 1995;**55**:1946–55.
16. Breimann L, Friedman JH, Olsen RA. *Classification and regression trees*. Monterey, CA: Wadsworth and Brooks Publishing; 1984.
17. DiGiovanna MP, Chu P, Davison TL, et al. Active signaling by HER-2/neu in a subpopulation of HER-2/neu-overexpressing ductal carcinoma in situ: clinicopathological correlates. *Cancer Res* 2002;**62**:6667–73.
18. Brandt BH, Roetger A, Dittmar T, et al. c-erbB-2/EGFR as dominant heterodimerization partners determine a motogenic phenotype in human breast cancer cells. *FASEB J* 1999;**13**:1939–49.

19. Dittmar T, Husemann A, Schewe Y, et al. Induction of cancer cell migration by epidermal growth factor is initiated by specific phosphorylation of tyrosine 1248 of c-erbB-2 receptor via EGFR. *FASEB J* 2002;**16**:1823–5.
20. Knowlden JM, Gee JM, Seery LT, et al. c-erbB3 and c-erbB4 expression is a feature of the endocrine responsive phenotype in clinical breast cancer. *Oncogene* 1998;**17**:1949–57.
21. Pawlowski V, Revillion F, Hebbar M, et al. Prognostic value of the type I growth factor receptors in a large series of human primary breast cancers quantified with a real-time reverse transcription-polymerase chain reaction assay. *Clin Cancer Res* 2000;**6**:4217–25.
22. Suo Z, Risberg B, Kalsson MG, et al. EGFR family expression in breast carcinomas. c-erbB-2 and c-erbB-4 receptors have different effects on survival. *J Pathol* 2002;**196**:17–25.
23. Holbro T, Beerli RR, Maurer F, et al. The ErbB2/ErbB3 heterodimer functions as an oncogenic unit: ErbB2 requires ErbB3 to drive breast tumor cell proliferation. *Proc Natl Acad Sci USA* 2003;**100**:8933–8.
24. Hudelist G, Kostler WJ, Attems J, et al. Her-2/neu-triggered intracellular tyrosine kinase activation: in vivo relevance of ligand-independent activation mechanisms and impact upon the efficacy of trastuzumab-based treatment. *Br J Cancer* 2003;**89**:983–91.
25. Wakeling AE. Inhibitors of growth factor signalling. *Endocr Relat Cancer* 2005;**12**(Suppl. 1):S183–7.
26. Agus DB, Gordon MS, Taylor C, et al. Phase I clinical study of pertuzumab, a novel HER dimerization inhibitor, in patients with advanced cancer. *J Clin Oncol* 2005;**23**: 2534–43.