Homeobox B13 in breast cancer
– Prediction of tamoxifen benefit

Piiha-Lotta Jerevall
We dance round in a ring and suppose,
but the secret sits in the middle and knows.

Robert Frost
SUPERVISOR

Olle Stål, Professor
Department of Clinical and Experimental Medicine, Division of Surgery and Clinical Oncology, Faculty of Health Sciences, Linköping University

CO-SUPERVISOR

Agneta Jansson, Associate Professor
Department of Clinical and Experimental Medicine, Division of Surgery and Clinical Oncology, Faculty of Health Sciences, Linköping University

FACULTY OPPONENT

Håkan Hedman, Associate Professor
Department of Radiation Sciences, Division of Oncology, Faculty of Medicine, Umeå University

COMMITTEE BOARD

Sven Hammarström, Professor
Department of Clinical and Experimental Medicine, Division of Cellbiology, Faculty of Health Sciences, Linköping University

Karin Leandersson, Associate Professor
Cell and Experimental Pathology and Center for Molecular Pathology, Department of Laboratory Medicine, Lund University

Charlotta Enerbäck, Associate Professor
Department of Clinical and Experimental Medicine, Division of Cellbiology, Faculty of Health Sciences, Linköping University
ABSTRACT

A major issue in the management of breast cancer is to identify patients who are less likely to be cured after primary treatment and would benefit from adjuvant chemotherapy. Of great importance is also identification of patients with only local disease who traditionally would be given chemotherapy but would survive without. In this thesis we have validated the utility of the two-gene ratio HOXB13:IL17BR, which previously has been demonstrated to predict disease-free survival in tamoxifen-treated breast cancer patients. We have also studied the prognostic and predictive utility of a single gene as a biomarker in breast cancer medicine.

We could confirm that HOXB13:IL17BR may classify patients with different treatment benefit; only patients with a low value showed benefit from prolonged duration of tamoxifen therapy, whereas for the group with high ratios, the long-term recurrence rate did not improve with longer treatment duration.

The combination of HOXB13:IL17BR and the molecular grade index (MGI), another prognostic marker, has been shown to outperform either alone in predicting risk of breast cancer recurrence. We validated the prognostic utility of HOXB13:IL17BR+MGI in a large randomized patient cohort and found that this risk classification identified more than 50% of the tamoxifen-treated lymph node-negative patients as having a less than 3% risk of distant recurrence and breast cancer death. Furthermore, we developed and tested a continuous risk model of HOXB13:IL17BR+MGI called Breast Cancer Index (BCI), for estimation of recurrence risk at the individual level. Our study shows that BCI has the ability to identify more than 50% of patients with a low risk of recurrence more accurately than using traditional risk assessment. These results suggest that BCI may help clinicians to make better informed treatment decisions and spare toxic chemotherapy for a large group of breast cancer patients.

The protein expression of HOXB13 was also shown to be a valuable predictor in postmenopausal patients. High expression was associated with worse outcome after tamoxifen therapy. In a premenopausal cohort, patients with hormone receptor-positive tumors showed benefit from tamoxifen regardless of HOXB13 expression. Further analysis indicated that estrogen receptor β (ERβ) modified the performance of HOXB13 as a predictor of treatment effect and should be taken into account when identifying patients less likely to respond to the therapy given.

In conclusion, BCI identifies patients with a very low risk of distant recurrence. It may be utilized in the management of breast cancer patients to optimize the use of chemotherapy. HOXB13 protein expression may be used as a marker for tamoxifen benefit, but its performance in premenopausal patients might be modified by ERβ.
# TABLE OF CONTENTS

**Populärvetenskaplig sammanfattning** ....................................................... 3
**Original publications** ................................................................. 5
**Abbreviations** ................................................................................. 7
**Background** ......................................................................................... 9
  - The breast ................................................................................................. 10
  - The estrogen receptor ............................................................................. 11
  - The progesterone receptor ..................................................................... 13
  - Homeobox genes ................................................................................... 13
    - HoxB13 ................................................................................................ 15
    - Function of HOXB13 in normal and tumoral tissue ..................... 15
    - HOXB13 in breast cancer ................................................................ 16
  - Breast cancer ......................................................................................... 17
    - Epidemiology ......................................................................................... 17
    - Risk factors ............................................................................................ 18
  - Breast cancer therapy ........................................................................... 19
    - Surgery .................................................................................................. 19
    - Radiotherapy .......................................................................................... 19
    - Chemotherapy ........................................................................................ 19
    - Targeted therapy ................................................................................... 19
  - Tamoxifen resistance ............................................................................ 20
  - Prediction of patient outcome ............................................................. 21
    - Classical prognostic markers .......................................................... 22
    - Predictive factors .................................................................................. 22
  - Gene profiles as prognostic and predictive markers ....................... 23
    - The intrinsic subtypes of breast cancer ........................................... 24
    - MammaPrint® ....................................................................................... 24
    - Oncotype Dx® ....................................................................................... 25
Bröstcancer är en sjukdom vars förlopp kan skilja sig väldigt mellan olika individer. Tumörerna kan ha varierande biologi och därmed skiljer sig även prognosen och behandlingsbehovet mellan patienter. Med införandet av mammografikontroller, som leder till att man kan upptäcka tumörer i bröstet på ett tidigare stadium, har man fått bättre möjligheter att behandla och bota sjukdomen. Trots en långsamt ökning av antalet nya fall har överlevnadschanserna ökat, främst tack vare förfinade metoder för diagnos och bättre behandlingar. Det är ett ständigt pågående arbete att hitta markörer som kan användas för att särskilja mellan olika sorters bröttumörer och för att underlätta för den behandlande läkaren att optimera behandlingsstrategin för patienten. De verktyg som idag används för att förutsäga en patients prognos och sannolikhet för ett positivt behandlingssvar är relativt okänsliga, vilket blir uppenbart när patienter som tros ha en god prognos ändå får återfall i sin sjukdom. Dessutom finns det patienter med förmodad dålig prognos som får behandling med cellgifter trots att de antagligen har en väldigt låg recidivrisk. Detta är en av de stora utmaningarna inom bröstcancerforskning; vilka patienter förutsätts ha dålig nytta av hormonell behandling, vilka tros ha nytta av cellgiftsbehandling, och vilka hade klarat sig bra utan endera behandlingen?

Målet med denna avhandling var att undersöka och validera nyttan av den lovande biomarkören HOXB13, både ensamt och i kombination med ett annat prognostiskt verktyg, för att förutsäga behandlingsnytta och prognos för bröstcancerpatienter. Forskningen har utförts på arkiverade tumörer från patienter som har varit med i tidigare kliniska studier, för vilka vi har utförlig information om diagnos, behandling och utfall.

Vi visade att kombinationen av två genindex, HOXB13:IL17BR och MGI, kunde användas för att förutsäga den återfallsfria överlevnaden hos bröstcancerpatienter. Mer än 50% av de behandlade patienterna identifierades som lågriskpatienter, med lägre än 3% risk för återfall eller död i sin sjukdom. Vi vidareutvecklade även detta kombinationsindex så att det kan användas som ett hjälpmedel för risikopuskattning.
POPULÄRVETENSKAPLIG SAMMANFATTNING

Med detta verktyg, Breast Cancer Index, kan man på ett mer exakt sätt identifiera lågriskpatienter som inte behöver extra behandling i form av cellgifter.

Vi har också undersökt HOXB13, ett av de proteiner som ingår i Breast Cancer Index, och funnit att det är en indikator på hur stor nytta en patient har av hormonell behandling. Ett högt uttryck av proteinet gjorde att läkemedlet inte fungerade lika bra, och patienterna hade en högre risk att få återfall i sin sjukdom.
This thesis is based on the following original papers, which are referred to by their Roman numerals (I-IV).


<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF-1</td>
<td>Activation function-1</td>
</tr>
<tr>
<td>ASCO</td>
<td>American Society of Clinical Oncology</td>
</tr>
<tr>
<td>BCI</td>
<td>Breast Cancer Index</td>
</tr>
<tr>
<td>BCS</td>
<td>Breast cancer survival</td>
</tr>
<tr>
<td>BIG</td>
<td>Breast International Group</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence interval</td>
</tr>
<tr>
<td>CMF</td>
<td>Cyclophosphamide, methotrexate and fluorouracil</td>
</tr>
<tr>
<td>Ct</td>
<td>Cycle threshold</td>
</tr>
<tr>
<td>DMFS</td>
<td>Distant metastasis-free survival</td>
</tr>
<tr>
<td>DRFS</td>
<td>Distant recurrence-free survival</td>
</tr>
<tr>
<td>EGF</td>
<td>Epidermal growth factor</td>
</tr>
<tr>
<td>ER</td>
<td>Estrogen receptor</td>
</tr>
<tr>
<td>ERE</td>
<td>Estrogen response element</td>
</tr>
<tr>
<td>FDA</td>
<td>Food and Drug Administration</td>
</tr>
<tr>
<td>FFPE</td>
<td>Formalin-fixed paraffin-embedded</td>
</tr>
<tr>
<td>HER2</td>
<td>Human epidermal receptor 2</td>
</tr>
<tr>
<td>Hox</td>
<td>Homeobox</td>
</tr>
<tr>
<td>HR</td>
<td>Hazard ratio</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>IHC</td>
<td>Immunohistochemistry</td>
</tr>
<tr>
<td>IL17BR</td>
<td>Interleukin 17 receptor B</td>
</tr>
<tr>
<td>MGI</td>
<td>Molecular Grade Index</td>
</tr>
<tr>
<td>MINDACT</td>
<td>Microarray In Node negative and 1 to 3 positive lymph node disease may Avoid ChemoTherapy</td>
</tr>
<tr>
<td>MIQE</td>
<td>Minimum Information for Publication of Quantitative Real-Time PCR Experiments</td>
</tr>
<tr>
<td>NCCN</td>
<td>National Comprehensive Cancer Network</td>
</tr>
<tr>
<td>NHG</td>
<td>Nottingham grade, Elston grade</td>
</tr>
<tr>
<td>PR</td>
<td>Progesterone receptor</td>
</tr>
<tr>
<td>RFS</td>
<td>Recurrence-free survival</td>
</tr>
<tr>
<td>RRR</td>
<td>Recurrence rate ratio</td>
</tr>
<tr>
<td>RT-qPCR</td>
<td>Reverse transcriptase quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>TAILORx</td>
<td>Trial Assigning Individualized Options for Treatment</td>
</tr>
<tr>
<td>TRANSBIG</td>
<td>Translating molecular knowledge into early breast cancer management; building on the BIG network for improved treatment tailoring.</td>
</tr>
<tr>
<td>TMA</td>
<td>Tissue microarray</td>
</tr>
<tr>
<td>TNM</td>
<td>Tumor Node Metastasis</td>
</tr>
</tbody>
</table>
Breast cancer is a disease with high diversity, with multiple subtypes differing in both prognostic and therapeutic implications. In the past decade, intensive research on the subject has considerably improved the way that breast tumors are characterized, and we have also gained insight into the complex nature of the disease. Screening programs leading to earlier detection, in combination with all the efforts on the matter, have resulted in better possibilities to treat and cure the disease. Despite a slow, but steady increase in incidence rates, the survival rates have improved thanks to refined tools for diagnostics and better therapies.

In cancer medicine, the conventional approach to therapy has been to provide treatment according to the organ or tissue from which the tumor is derived. Modern research has completely changed this perspective, and knowledge of the multifaceted nature of different tumors originating from the same organ is slowly turning the clinical approach to a more personalized approach. Breast cancer is a good example of how a single diagnosis can display many different natures, and hence also require different treatment strategies depending on the inherent features of the individual tumor. There is an ongoing effort to find biomarkers that can be used to distinguish between and group different types of breast tumors in order to help clinicians make better informed decisions about optimal treatment strategy for the individual patient. The factors that are used today for prognosis and treatment prediction are relatively insensitive, which is apparent when patients expected to have a good prognosis still relapse. Today, modern technology is employed in order to develop and validate biomarkers that can be useful for predicting relapses and treatment benefit.

The aim of the research presented in this thesis was to look further into a proposed biomarker for breast cancer. We sought to develop and perform further validations of its potential as a treatment predictive marker for use in the breast cancer clinic to provide guidance in selecting the optimal treatment strategy and thereby avoiding over- and undertreatment.
BACKGROUND

Figure 1. Anatomy of the mammary gland (modified from Joakim Waltersson ©2009).

The breast

The breast is one of the organs in the female human body that is not fully developed at birth but continues to develop into adulthood. The development is initiated in the intrauterine life during embryogenesis and continues during fetal development, puberty and pregnancy (Russo et al., 2004). Only very primitive structures make up the breast in the newborn; small ducts and ductules which, under the influence of sex hormones during puberty, grow and divide into primary and secondary ducts. These elongate and branch into terminal ducts and alveolar buds, followed by formation of lobular structures, which complete the extensive ductal network that make up the breast tissue. Full differentiation of the mammae is not attained until the first full-time pregnancy (Russo et al., 2004). During gestation, the final development into a fully functional breast takes place through extensive growth and proliferation of epithelial cells and formation of the secreting units called alveoli that synthesize milk. Branching ducts extend from the areola in a treelike pattern, with small ductules terminating in lobules containing alveoli composed of the secretory units (Figure 1). Luminal epithelial cells line the ductal system, which is surrounded by myoepithelial cells in direct contact with the basal membrane (Anderson et al., 2004).

The breast in the premenopausal woman consists of approximately 15% epithelial cells, whereas corresponding number for the postmenopausal woman is 5%, due to replacement of epithelium by fat tissue (Hutson et al., 1985). The mature breast is composed of adipose and glandular tissue supported by fibers called Cooper’s ligament. The distribution of the two tissue types is usually similar in both breasts,
but shows wide variation between women (Ramsay et al., 2005). The breast tissue expresses receptors for the female sex hormones estradiol and progesterone, the key regulatory hormones of mammary gland proliferation.

The estrogen receptor

Proliferation of cells in the mammary gland is triggered by estrogens, the primary female sex hormones. Estrogens can influence growth, development and function, and these effects are mediated via estrogen receptors. In women, the major estrogens are estrone, estradiol and estriol, of which estradiol is the most predominant and potent substrate. Estradiol diffuses through the plasma membrane and its biological effects takes place after interaction with intracellular receptors. The estrogen receptor (ER) exists in two subtypes, ERα and ERβ, each encoded by separate genes (Walter et al., 1985, Kuiper et al., 1996) but show a high homology in their ligand- and DNA-binding domains. ERα is the predominant subtype in the breast, uterus, cervix and vagina (Couse et al., 1997), and is also present in the cardiovascular system and the brain. ERβ is more widely distributed throughout the body. The two subtypes also differ in their response to estrogen agonists and antagonists (Barkhem et al., 1998). Proliferation of the mammary gland occurs in the luminal compartment; the luminal epithelial cells lining the ducts and lobules are the main site for the ERα, although they also express ERβ (Joshi et al., 1986, Petersen et al., 1987, Speirs, 2002). The myoepithelial cells proliferate very rarely, and these cells usually express ERβ (Joshi et al., 1986, Speirs, 2002), which displays a more widespread distribution in the mammary gland as compared to ERα (Speirs, 2002).

Signaling via ERα can occur through several mechanisms, both genomic and extranuclear pathways, which can be ligand-dependent or ligand-independent (Figure 2). In the classical ligand-dependent mechanism of action, a conformational change of the receptor takes place upon binding of ligand, the receptor dimerizes and binds to estrogen response elements (EREs) in promoter regions of target genes (Klein-Hitpass et al., 1988). The contact between the receptor and DNA can be either direct or via cofactor proteins (McKenna et al., 1999), which stabilize the complex. The effect on target genes is either transcriptional activation or repression, depending on the cell and promoter context. The transcriptional activity is mediated via two activation domains in the receptor, the ligand-independent AF-1 (activation function-1) in the N-terminus, and a hormone-dependent AF-2 domain in the ligand-binding domain (Tora et al., 1989). Activation via these domains is dependent on cofactor recruitment.
In a nonclassical manner, ERα can induce transcription of target genes which lack EREs. In these cases, the activation is mediated by AP-1 sites, which is the binding site for the transcription factor complex Jun/Fos (Webb et al., 1999). Activation of ERα can also occur in the absence of estrogen. This can take place when the receptor is phosphorylated by downstream effects of cell surface receptors. Through an enhanced protein-protein interaction, binding to DNA is possible even in the absence of ligand (Jakacka et al., 2001). In addition, a subpopulation of the classical ERs that is located in the plasma membrane, can mediate nongenomic actions of estrogen, which employ intracellular calcium, production of cAMP, and activation of different signaling pathways (Björnström et al., 2005).

**Figure 2.** Classical and non-classical pathway for signaling via estrogen receptors. (ER, estrogen receptor; E, estrogen; CoF, cofactor; ERE, estrogen-response element; P, phosphorylation)
As a result of alternative splicing of the last coding exon, ESR2, the gene for ERβ, encodes at least five different C-terminal isoforms with different patterns of expression (Moore et al., 1998). Only ERβ1 is fully functional as a homodimer, whereas the other isoforms modulate ERβ1 transactivation (Leung et al., 2006). The proliferative effect of estradiol is mainly mediated by ERα; the function of ERβ appears to oppose the action of ERα. Signaling via ERβ and ERα occurs in a similar fashion, though ERα shows a greater transcriptional activity. It has been shown that the two receptors can form functional heterodimers capable of DNA binding and transcription of target genes (Cowley et al., 1997). In vitro studies have demonstrated that the β-subtype inhibits the transcriptional activity of ERα, suppresses proliferation and increases the estradiol-induced degradation of ERα, lowering the overall mRNA and protein levels of ERα (Reviewed in: Fox et al., 2008). In fact, ERβ is downregulated in breast tumor tissue as compared to normal breast epithelium and therefore it is likely that this receptor subtype plays a role in tumorigenesis.

The progesterone receptor

Progesterone, which is the other major female sex hormone besides estradiol, exerts its actions via the progesterone receptor (PR), which belongs to the same receptor family as the ERs. This receptor also has isoforms, PRA and PRB, encoded by the same gene. Upon ligand-binding, the receptor form homo- or heterodimers which bind to response elements in the promoters of target genes resulting in transcription. The two receptor variants have distinct functions, but the basis of these differences is not fully understood. The subtypes are under the control of two separate promoters (Kastner et al., 1990) which, in combination with the cell type, influence the effect of progesterone stimulation. PR is one of the target genes of estrogen-signaling; estradiol-stimulation results in upregulation of PR (Kastner et al., 1990, Vienonen et al., 2002), which is the most commonly used indicator of a functional ER-signaling (Lapidus et al., 1998). In general, cells in human tissue coexpress PRA and PRB in similar levels (Mote et al., 1999), but predominance of either of the receptors may occur. This is often the case in breast cancer, where the equality is altered early in carcinogenesis (Graham et al., 1995, Mote et al., 2002).

Homeobox genes

Homeobox (Hox) genes are a group of genes encoding nuclear proteins acting as transcription factors regulating essential mechanisms during development, modulating morphogenesis and cell differentiation. The proposal of existence of genes responsible for the spatial development of an embryo was formed already in the early 1900s (Bridges, 1921), but it was not until 1978 this gene family was
discovered in the fruit fly *Drosophila Melanogaster* (Lewis, 1978). It was found that mutations in these genes resulted in a dysregulated development with changes in body structure, such as formation of an additional pair of legs on the head instead of the antennae (Lewis, 1978). This phenomenon is called homeotic transformation, hence the name of the gene family.

The role of the Hox gene family in embryonic development has been thoroughly investigated; it is known to regulate several essential processes such as cell identity, cell differentiation and proliferation (Cillo *et al.*, 2001, Shah *et al.*, 2010). Apart from being crucial regulators of normal development, some Hox genes are also expressed in adult tissue, involved in control of essential cellular processes such as cell motility, apoptosis, receptor signaling, angiogenesis and cell-cell as well as cell-extracellular matrix interactions (Cillo *et al.*, 2001, Shah *et al.*, 2010). During vertebrate development, the Hox genes have a spatial and temporal collinearity – genes 3' to 5' in the cluster are expressed in the same order along the anterior-posterior axis of the developing animal. The same pattern applies to the temporal expression, with the genes successively expressed beginning with the 3' genes and ending with those located closer to the 5' end. Posterior prevalence, or posterior dominance, also applies to the Hox genes, meaning that the most posterior Hox gene expressed at a certain level along the body axis is the one that dictates the developmental program.

The homeotic complex in *Drosophila* has highly conserved homologues in most animals, all located in well assembled clusters along the chromosomes. In mammals, there are four unlinked Hox gene clusters called Hox A, Hox B, Hox C and Hox D, each on a different chromosome (Figure 3). In humans, the clusters are located on chromosome 7, 17, 12 and 2 (Scott, 1992), spanning approximately 200 kb each and contain 9-11 genes. Alignment of the clusters identifies thirteen paralog groups of genes, according to the homogeneity of the highly conserved homeodomains, with a different subset of the paralogs in each cluster (Krumlauf, 1994). To date, 39 human Hox genes have been identified (Scott, 1992, Zeltser *et al.*, 1996). Due to the coexistence of multiple Hox clusters, the principles of spatio-temporal expression of the Hox genes are more complex in mammals, which also reflect the more complex organization of internal and external organs as compared to *Drosophila*. The posterior prevalence is not directly applicable to mammal development. Instead, the morphogenetic program is controlled by the combination of functionally active Hox genes, the Hox code, rather than by individual dominant ones.
Figure 3: Schematic overview of the homeobox genes on the four loci.

**HoxB13**

The latest gene to be discovered in the Hox B cluster was HOXB13 (Zeltser et al., 1996). This gene is located approximately 70 kb upstream the rest of the Hox B cluster, with HOXB9 as the closest neighbor. Attempts to find the missing HOXB10-12 have been made, but it is suggested that these paralogs have been lost during evolution (Zeltser et al., 1996).

In the developing embryo, HOXB13 is expressed in the caudal part of the spinal cord, the hindgut and the urogenital sinus and is important for development of the prostate (Zeltser et al., 1996, Economides et al., 2003a). This gene is usually downregulated in adult tissue but one of the earliest studies of HOXB13 revealed an expression in skin, and an involvement of this gene in wound healing (Stelnicki et al., 1998, Kömüves et al., 2003). Subsequent mapping of HOXB13 in adult tissues has shown expression in prostate, distal colon, placenta, uterus, thymus, testis, bone marrow, kidney and salivary glands (Stelnicki et al., 1998, Sreenath et al., 1999, Economides et al., 2003a, Takahashi et al., 2004, Okuda et al., 2006).

**Function of HOXB13 in normal and tumoral tissue**

Not much is known about the normal function of HOXB13 in adult tissues, since most studies are focused on its involvement in carcinogenesis. Loss-of-function mutations in mice result in overgrowth of all major structures derived from the tail bud and a dysregulated development of the ventral prostate ducts along with an absence of
secretory proteins (Economides et al., 2003a, 2003b). Experimental studies on prostate cancer cell lines that normally do not express HOXB13 showed that ectopic expression had growth-suppressive effects possibly by driving the cells into terminal differentiation via inactivation of the β-actin-TCF-signaling pathway (Jung et al., 2004a). Similar mechanisms are observed in colorectal cancer cell lines, where this pathway usually is constitutively activated (Jung et al., 2005). Further studies have shown that HOXB13 suppresses the androgen-activated androgen receptor-signaling in prostate cancer cells, and inhibits cell growth (Jung et al., 2004b).

Overexpression of HOXB13 in an experimental system with rat epidermal keratinocytes illustrates a role for HOXB13 in suppressing proliferation, promoting apoptosis and epidermal differentiation, mechanisms similar to those seen in studies of prostate cancer cell lines (Mack et al., 2005). Evidence of an important role of HOXB13 in wound healing exists, and there is also a possibility that this gene could be involved in skin pathogenesis (Stelnicki et al., 1998, Mack et al., 2005).

As mentioned, several studies link HOXB13 to carcinogenesis, however, different mechanisms and signaling pathways are engaged in different tissues, since expression of the gene can be either beneficial or unfavorable depending on the tumor type. In a study examining the expression of different HOXB genes in normal cervical epithelium and carcinoma, López et al. (2006) detected RNA transcripts of HOXB13 in the tumoral tissue, but not in its normal counterpart. The same pattern has been shown in ovarian cancer (Yamashita et al., 2006). HOXB13 expression confers invasiveness in both endometrial and ovarian cancer cell lines and promotes ovarian tumor growth in vivo, whereas a forced downregulation is associated with a reduced invasive ability and proliferation (Zhao et al., 2005, Yamashita et al., 2006, Miao et al., 2007). On the other hand, in both prostate cancer and colon cancer, HOXB13 is a growth suppressant, and a reduced expression is connected to worse tumor grade and microvessel invasion in renal cell carcinoma (Jung et al., 2004a, 2004b, 2005, Okuda et al., 2006). It has also been suggested that HOXB13 acts as a tumor suppressor gene in renal cells, and that the loss of expression can be due to 5′-methylation of the gene (Okuda et al., 2006).

**HOXB13 in breast cancer**

A lot is known about the implications of different Hox genes in breast tumorigenesis, such as the effects on cell cycle regulation, proliferation, angiogenesis, metastasis and invasion (Reviewed in: Chen et al., 2003). The precise role of HOXB13 is however not extensively elucidated. It is known that HOXB13 is upregulated in several breast cancer cell lines and primary breast tumors as compared to normal breast cells
(Cantile et al., 2003, Svingen et al., 2003). Studies in the wake of the predictive two-gene expression ratio HOXB13:IL17BR correlates a high HOXB13 mRNA expression with tumor aggressiveness, worse patient outcome and dysregulated ER-signaling, which possibly is involved in tamoxifen resistance (Ma et al., 2004, Goetz et al., 2006, Ma et al., 2006, Jansen et al., 2007, Wang et al., 2007, Ma et al., 2008). On the molecular level, Ma and colleagues (2004) showed that ectopic expression enhances cell motility and invasion induced by epidermal growth factor (EGF). In a study of mRNA levels in breast tumors, HOXB13 expression was inversely correlated with expression of ERα, and it was suggested that HOXB13 is an ER-dependent estrogen-responsive gene; estradiol treatment of breast cancer cell lines reduced HOXB13 levels in ER-positive but not in ER-negative cells (Wang et al., 2007).

In contrast to the association between high HOXB13 expression and bad outcome for breast cancer patients, there are epigenetic studies pointing towards the inverse relationship. Two independent investigators have detected a hypermethylation of HOXB13, correlating to a decreased expression of its transcript as well as a shorter disease-free patient survival (Rodriguez et al., 2008, Tommasi et al., 2009).

Breast cancer

Epidemiology

Breast cancer is primarily a disease of developed countries, with incidences varying greatly worldwide. For women, cancer of the breast is the most common cancer at the global level, constituting 16% of all female cancers. Despite the predominance in the western world, a majority of all breast cancer-related deaths occurs in developing countries; of 517 000 female breast cancer-related deaths in 2004, a majority occurred in developing countries (World Health Organization, 2008). In Sweden, 7 380 new cases of female breast cancer were diagnosed in 2009, which corresponds to almost 30% of the cancer incidence in Swedish women (Socialstyrelsen - The National Board of Health and Welfare, 2010). The mean annual increase in incidence rate has been around 1.2% the last 20 years, but the mortality shows a slow but decreasing trend, with approximately 1 400 deaths reported in 2009 (Socialstyrelsen - The National Board of Health and Welfare, 2011). The relative 5- and 10-year survival rates are approximately 90% and 80%, respectively (Socialstyrelsen - The National Board of Health and Welfare, 2009). The observed incline in incidence rates may be related to changes in reproductive patterns, but also improved screening and higher detection rates. When mammographic screening programs were implemented and became generally available in 1986-1989, a transient peak of the breast cancer incidence was seen (Tejler et al., 2004). The improved survival after diagnosis can be
explained by earlier diagnosis and improvements in surgery, radiotherapy and medication.

Figure 4. Age-standardized incidence and mortality from breast cancer among women in the south-east health care region in Sweden (modified from Onkologiskt centrum 2009).

Risk factors

The major known risk factor for development of breast cancer is the cumulative exposure to estrogen and progesterone. An early menarche, late menopause, nulliparity and use of hormone-replacement therapies are hormonal variables known to increase an individual’s risk of developing breast tumors. The impact of pregnancies, however, is dual. Pregnancy at a young age reduces the life-time risk of developing a breast tumor, but there is a transient increased risk within the next decade following the pregnancy (Hankinson et al., 2004). This effect can probably be ascribed to the high levels of circulating hormones during pregnancy. The long-term effects are due to the hormone-initiated differentiation of breast epithelial cells which make them less susceptible to cancer initiation. Other factors suggested to affect the breast cancer risk are lifestyle factors such as alcohol intake, weight and diet (Smith-Warner et al., 1998, van den Brandt et al., 2000, Michels et al., 2007). Genetics also play a role in the risk of developing breast cancer, even though only a small proportion of all breast tumors are hereditary. Having a close relative with the disease increases the risk.
Breast cancer therapy

Surgery
The primary treatment modality for breast cancer is surgical removal of the tumor (Holmberg, 1995). In some cases, with small tumors without lymph node involvement, the cancer can be cured by surgery alone, but usually there is a need for additional treatment in order to prevent future relapse. Complete mastectomy (i.e. surgical removal of the whole breast) is used less frequently, since breast-conserving surgery in combination with radiotherapy has been shown to be an adequate alternative (Early Breast Cancer Trialists’ Collaborative Group, 1995, Wapnir et al., 2011). Indications for mastectomy are large or multicentric tumors, inflammatory breast cancer and previous breast irradiation, among others (Audretsch et al., 2006).

Radiotherapy
Radiation following surgery is the international standard treatment in order to reduce the risk of local relapse. Adjuvant radiotherapy is applied to all patients who have gone through breast-conserving surgery and a majority of mastectomized patients with affected lymph nodes, and has been shown to improve long-term survival (Early Breast Cancer Trialists' Collaborative Group, 2005b).

Chemotherapy
With the purpose to reduce the risk of undetected deposits of disease to develop into a clinical recurrence, adjuvant chemotherapy is administered and results in improved relapse-free and overall survival in the general breast cancer population. Polychemotherapy has proven more efficient than monotherapy (Early Breast Cancer Trialists' Collaborative Group, 2005a, Carrick et al., 2009), and anthracycline regimens are more beneficial than the CMF (cyclophosphamide, methotrexate and fluorouracil) combinations (Early Breast Cancer Trialists' Collaborative Group, 1998b, 2005a). Taxanes, such as paclitaxel and docetaxel, are well established drugs for metastatic breast cancer and are comparable to anthracyclines when administered as monotherapy. However, the coadministration of taxanes and anthracycline has been shown to improve distant recurrence rates as well as overall survival in the adjuvant setting (De Laurentiis et al., 2008).

Targeted therapy
Molecular targeted drugs selectively block specific signaling pathways in the cell, as opposed to chemo- and radiotherapy which function at the broad spectra. The targets are usually specific to cancer cells, and therefore the therapies may be more
effective but still less harmful to normal cells. One of the first molecular targets for cancer therapy was the ER; the selective ER modulator tamoxifen was originally developed as a contraceptive but turned out to be the standard treatment of care for ERα-positive breast cancer. Tamoxifen competes with estradiol in binding to ER, and upon binding the receptor dimerizes and binds to DNA, but transcription is inhibited since the AF-2 domain remains inactive and unable to recruit cofactors. The therapy is highly effective in improving long-term recurrence and mortality rates; five years of tamoxifen treatment gives a reduction of about 40% and 20-30%, respectively (Early Breast Cancer Trialists' Collaborative Group, 1998a, 2005a).

There are additional drugs on the market targeting ERα. Fulvestrant is also an ERα-ligand, but binding results in functional blockade and decreased ER-levels (Howell et al., 2000). The aromatase inhibitors anastrozole, exemestane and letrozole interfere with the ability of estrogen to promote the growth through ER by blocking the activity of the aromatase enzyme, which is necessary for the estrogen synthesis.

Approximately 15-30% of breast tumors have amplification of the gene encoding the human epidermal receptor 2 (HER2), resulting in an overexpression of the receptor protein on the cell surface. These tumors are usually associated with more aggressive behavior and worse outcome for the patients. Trastuzumab (Herceptin) is an antibody-based drug targeting the HER2 receptor. The mechanism of action is not fully understood, but effects upon binding to HER2 seems to be antibody-dependent cell cytotoxicity, disruption of ligand-independent receptor dimerization and downregulation of HER2 (Menard et al., 2003). Trastuzumab in combination with chemotherapy has become the standard treatment for HER2-amplified breast cancers, and the clinical efficacy is high, with significant reductions in risk of recurrence and death, both in early-stage and metastasized breast cancer (Piccart-Gebhart et al., 2005, Smith et al., 2007).

Tamoxifen resistance

As with many drugs, acquired resistance to tamoxifen therapy may develop over time. Also, a subset of ER-positive tumors demonstrates primary resistance, so called de novo resistance, and do not at all respond to the tamoxifen administered. Approximately one third of the patients undergoing five years of tamoxifen treatment will have recurrent disease within 15 years (Early Breast Cancer Trialists' Collaborative Group, 2005a). For metastatic disease, the response rate is just above 30% (Stuart et al., 1996).
There are several mechanisms proposed for tamoxifen resistance, e.g., overexpression of coactivators, underexpression of corepressors and activation of growth receptor pathways which may phosphorylate and activate ERα. Another much-disputed mechanism for resistance is inactive alleles of the enzyme CYP2D6 converting tamoxifen to its active metabolite endoxifen, mediating a decreased responsiveness to the treatment (Hoskins et al., 2009). However, studies of how the risk of recurrence relates to the patient’s CYP2D6 genotype have shown divergent results.

Endocrine resistance is a major clinical problem in the treatment of breast cancer and emphasizes the need for factors that can be used in the clinic to predict the efficacy of endocrine treatment more accurately than what is possible today.

Prediction of patient outcome

Due to the diversity of breast cancers, which is illustrated by a very inconsistent manifestation of disease between patients, it is important to gain information about the nature of the tumor already at the time of diagnosis. In general, a large proportion of patients have a good prognosis, with a survival rate above 75% ten years after diagnosis. The risk of developing metastases is at a peak during the first five years, but there is still a risk for late relapses after ten years. Prognostic and predictive markers are used in the clinical setting to provide valuable information about the tumor. In general, they help to determine whether a patient requires additional treatment, and aids in treatment selection. The prognosis for a patient is dependent on a number of different features of the tumor. Some of the variables can be classified as time-dependent, for example tumor stage (i.e. how far the tumor has spread), tumor size, invasion and tissue destruction. The remainders are innate biological features of the tumor, such as growth rate and hormone receptors. Predictive factors, as opposed to prognostic, are used to foretell the patient’s benefit of a specific therapeutic regimen, which is associated with tumor sensitivity or resistance to that therapy. It is important to note that some factors can be of prognostic, as well as predictive significance. Due to the nature of prognostic and predictive biomarkers, in order to truly validate a prognostic factor, studies in untreated cohorts of patients are warranted. Likewise, the validation of a predictive factor requires cohorts of patients from randomized clinical trials, or measures of treatment response in advanced breast cancer.
Classical prognostic markers

The most well-established clinically used prognostic marker in breast cancer is the TNM (Tumor Node Metastasis) stage. It is a measure of how far the tumor has spread and summarizes information about three important features of the tumor, namely size, lymph node status and distant metastasis. These variables are scored, and the tumor is classified into one of five different stages (0; in situ, I, II, III and IV), with the highest stage correlating to a worse prognosis.

Another prognostic factor for classification of breast tumors is the Elston grade, or Nottingham Grade index (NHG). This is a histological scoring system with three variables taken into account; tubule formation (degree of structural differentiation; how much of the tumor is composed of ductal structures), frequency of mitotic figures and nuclear pleomorphism (Bloom et al., 1957, Elston et al., 1991). Each variable is scored 1-3 and the combination of these yields a final grading of the aggressiveness of the tumor. Grade 1 tumors are well-differentiated and have better prognosis than tumors of grade 3 (poorly differentiated). NHG correlates well with 10-year survival, and has been shown to be an independent factor in prediction of prognosis (Elston et al., 1991, Razavi et al., 2005, Rakha et al., 2008).

The probably most essential biomarker in breast cancer is the hormone receptor ERα, which is expressed in roughly 70-80% of all breast tumors. The majority of these are also positive for PR, since this receptor is upregulated in response to ER-signaling. Presence of ER and PR is associated with better short-term survival. The main value of ERα, however, is to aid in treatment selection, since presence of this receptor is a prerequisite for endocrine therapy.

The HER2 receptor amplification/overexpression is associated with shorter patient survival (Slamon et al., 1987, Ross et al., 1998, Esteva et al., 2004). HER2-positivity has also been shown to correlate with an adverse outcome after tamoxifen therapy (Ross et al., 1998), while some aromatase inhibitors result in higher response rates (Ellis et al., 2001, Ellis et al., 2003).

Predictive factors

Due to the heterogeneity of breast cancer, some patients with local disease will be treated with unnecessary chemotherapy, whereas the relapse risk for patients with residual micrometastases may be underestimated and the treatment administered will be inadequate or ineffective. Since all drugs used in the treatment of cancer are associated with side effects, it is necessary to individualize treatment strategies by identification of patients who need additional therapy, and who also will respond to
treatment. This is one of the main challenges in breast cancer medicine; which patients are unlikely to respond to endocrine therapy, and who may benefit from chemotherapy. Likewise, there is an extensive interest in the possibility to identify the 40-50% of patients who would have survived without chemo- or hormonal therapy (Early Breast Cancer Trialists' Collaborative Group, 1998a, b). Worth mentioning is, that even though a massive research is going on to find and implement predictive markers in the clinical management of breast cancer, to date there are no markers that are optimal in prediction of benefit to a certain therapy. This is illustrated by patients likely to respond to a given drug still may relapse during the course of the therapy.

As already mentioned, ERα is a strong predictor of benefit from endocrine treatment. There is a correlation between the ERα level in the tumor and the tamoxifen response, so that patients with high-expressing tumors show a better response rate (Bezwoda et al., 1991, Harvey et al., 1999, Paik et al., 2004). The overall effects of tamoxifen in ER-poor or ER-negative tumors are small, if any (Early Breast Cancer Trialists' Collaborative Group, 1998a, 2005a). Although ER is an accepted predictor, PR seems to have an impact on the treatment outcome. ERα-positive but PR-negative breast tumors respond less well to tamoxifen (Bardou et al., 2003). Similar to ER, response to endocrine treatment is directly and positively related to PR levels; elevated PR correlates to increased probability of tamoxifen response and longer time to treatment failure (Ravdin et al., 1992, Elledge et al., 2000, Stendahl et al., 2006). Expression of ERβ is also suggested to affect the likelihood of response to endocrine therapy; increasing levels are associated with better disease outcome (Murphy et al., 2006).

Like many other factors, HER2, which is a quite recently introduced treatment predictive marker in breast cancer, has mixed prognostic and predictive significance. As already mentioned, a tumor overexpressing HER2 makes the patient eligible for trastuzumab therapy, whereas the response to endocrine treatment is decreased.

Gene profiles as prognostic and predictive markers

Recent high-profile molecular studies of breast cancer have demonstrated the importance of the genetic makeup of tumors and have shown associations between the molecular portrait of breast cancer and tumor biologic and clinical behavior. In fact, the latest St Gallen consensus guidelines warrant the use of validated multigene assays in cases when traditional tumor markers fail to clearly indicate the need for chemotherapy (Goldhirsch et al., 2009). Unlike targeted therapy, chemotherapy does not have any reliable predictive factors. Although a number of multiparameter assays
have been developed, and some also commercialized, the implementation of molecular profiling into clinical praxis is underdeveloped on the global scale. Another challenge for the integration of multigene predictors for chemotherapy is that in general, only the characteristics of the tumors are measured, and the patients’ characteristics such as drug metabolism are not taken into account. Nevertheless, combination of multiple genes, as compared to the use of single-gene markers, may provide more accurate information about the tumor biology.

**The intrinsic subtypes of breast cancer**

A discovery which has had a major impact in directing breast cancer research is the existence of molecularly distinct breast tumor subtypes with pervasive differences in their gene expression patterns (Perou *et al.*, 2000). Assessed with microarrays, the molecular portraits that classify the different subtypes reflect differences in the intrinsic biology, such as growth rate, activation of intracellular pathways and cellular composition (Perou *et al.*, 2000, Sorlie *et al.*, 2001, 2003, 2004). The luminal cancers are characterized by ER-positivity, and are further classified into luminal A and B, with low and high tumor grade, respectively. The HER2-enriched group display amplification and high expression of HER2, and are generally ER-negative. Tumors, which are usually referred to as triple-negative, since they are negative for ERα, PR and HER2, fall into the basal-like subtype. The normal breast-like tumors express genes known to be expressed by adipose tissue and non-epithelial cell types. Further development of the intrinsic taxonomy have revealed the claudin-low subtype, characterized by stem-cell like properties, triple-negativity and loss of genes involved in cell-cell adhesion (Hennessy *et al.*, 2009). The hereditary BRCA1-mutation is associated with the basal-like subtype (Sorlie *et al.*, 2003). The intrinsic subtypes also display various patient outcome; the basal-like and HER2 subtypes have shortest survival times and shortest relapse-free survival, whereas the luminal A subtype have the best prognosis (Sorlie *et al.*, 2001, 2003).

**MammaPrint®**

The 70-gene real-time RT-qPCR (reverse transcriptase quantitative polymerase chain reaction) based Amsterdam signature, named MammaPrint®, is cleared by the US Food and Drug Administration (FDA). It predicts prognosis for patients with node-negative, stage I or II disease with a tumor size of 5 cm or less (van't Veer *et al.*, 2002). The signature is based on the top 70 genes from microarray data covering approximately 25 000 genes involved in cell cycle regulation, invasion, metastasis and...
angiogenesis (van’t Veer et al., 2002), and estimates risk of breast cancer-recurrence at five and ten years by identifying patients’ risk profiles. For patients classified as having a low risk, hormonal therapy may suffice, whereas chemotherapy is recommended for high-risk patients. A low-risk patient without any additional adjuvant treatment is estimated to have a 10% risk of relapse within ten years, whereas the chance for a high-risk patient is about three times as high (Buyse et al., 2006). MammaPrint® has been validated in a large number of patients (van de Vijver et al., 2002, Buyse et al., 2006, Bueno-de-Mesquita et al., 2007). The analysis requires fresh-frozen tumor samples or tissues collected and stored in an RNA preservative solution.

Subsequent to the initial validations of MammaPrint®, the MINDACT study was initiated as a second, independent, European multicenter study launched by the Breast International Group (BIG)². The MINDACT trial is a prospective randomized study enrolling 6,000 patients, comparing MammaPrint® with common clinical-pathological criteria in the selection of patients eligible for adjuvant chemotherapy. Treatment decisions for patients in the study will be made based on random assignment to MammaPrint® or traditional clinicopathological risk assessment (Bogaerts et al., 2006).

**Oncotype Dx®**

Another genetic test that is often compared to MammaPrint® is the Oncotype Dx®. However, Oncotype Dx® was developed with a different purpose; to predict the risk of distant recurrence in ER-positive early-stage breast cancer patients undergoing tamoxifen therapy (Paik et al., 2004). The profile, composed of 21 genes (16 cancer-related genes, 5 reference genes), allocates a recurrence score between zero and 100, corresponding to the likelihood of recurrence within ten years from the initial diagnosis. Patients can then be stratified into three levels of risks on the basis of their recurrence score, with distant recurrence rates of 7%, 14% and 30%, respectively. Retrospective validations have extended the eligible patients to comprise not only

---

2 The Breast International Group (BIG) is a global network of groups that conduct joint large breast cancer clinical trials. TRANSBIG ("Translating molecular knowledge into early breast cancer management; building on the BIG network for improved treatment tailoring.") is a consortium launched by BIG to endorse international collaboration in translational research. The MINDACT (Microarray In Node negative and 1 to 3 positive lymph node disease may Avoid ChemoTherapy) trial is the main project launched under TRANSBIG. (http://www.breastinternationalgroup.org/)

2 Oncotype Dx® is a registered trademark of Genomic Health, Inc.
the group of node-negative ER-positive invasive tumors, but also postmenopausal women with ER-positive tumors and lymph node involvement (Sparano et al., 2008). Oncotype Dx® can be used to determine the likelihood that a patient will benefit from certain types of chemotherapy (Paik et al., 2006). Unlike MammaPrint®, this test does not require fresh tissue, but can be performed in formalin-fixed paraffin-embedded (FFPE) tissues.

The validation studies of Oncotype Dx® have been performed to establish its prognostic significance, although a high recurrence score also is suggested to predict chemotherapy benefit (Paik et al., 2006). The ability to guide treatment selection is further evaluated in the TAILORx (Trial Assigning Individualized Options For Treatment) clinical prospective trial (Sparano, 2006). This trial is the first to be set off as a result of a US program with the purpose to individualize cancer treatment by the integration of modern diagnostic tests into clinical decision making, and is planned to enroll at least 10 000 breast cancer patients. Clinical use of Oncotype Dx® is endorsed by the ASCO (American Society of Clinical Oncology) and NCCN (National Comprehensive Cancer Network) guidelines.

The Rotterdam signature

This signature, which is also known as the 76-gene assay, was developed as a prognostic microarray-based tool for node-negative patients independent of hormone receptor status (Wang et al., 2005). It has not yet been commercialized, but validation studies have confirmed both prognostic and predictive significance (Foekens et al., 2006, Desmedt et al., 2007, Yu et al., 2007). It has no genes in common with Oncotype Dx® and MammaPrint®, but likewise, the gene selection is biased towards proliferation-related genes. Like Oncotype Dx®, FFPE tissues are used in the analysis.

The two-gene expression ratio HOXB13:IL17BR

The HOXB13:IL17BR index is yet another predictive biomarker in the series of recently developed assays for breast cancer clinic. It is a real-time RT-qPCR-based assay performed on FFPE tissues, derived from a genome-wide microarray analysis of ER-positive breast cancers (Ma et al., 2004). The implication of HOXB13 in breast cancer is beginning to be understood, but the role of IL17BR (interleukin 17 receptor B) is not well explored. IL17BR encodes a cytokine receptor that may be involved in the immune response against tumor cells, but this still remains to be shown. The chromosomal location maps to 3p21, which is a region frequently lost in breast cancer, thus, low expression of IL17BR may correlate to the loss of putative tumor suppressor genes within the same region.
HOXB13:IL17BR shows differential expression in recurrent breast cancers as compared to those with no evidence of relapse, and is also associated with survival. It has been suggested as a negative predictive factor of tamoxifen benefit, both in the adjuvant and the metastatic setting, and is significant independent of standard clinical and pathological markers (Ma et al., 2004, Goetz et al., 2006, Ma et al., 2006, Jansen et al., 2007, Wang et al., 2007, Ma et al., 2008). A high index is proposed to be a marker of impaired ER-signaling and tamoxifen resistance (Wang et al., 2007).

**Molecular Grade Index (MGI)**

MGI is a gene expression assay comprised of five genes related to histological grade and tumor progression, which recapitulates tumor grade and can predict clinical outcome with high performance (Ma et al., 2008). This analysis differs from the other multigene assays as such that it does not involve complex weighting trained on clinical outcome; instead it is a molecular correlate of tumor grade. In the original study it was seen that MGI and HOXB13:IL17BR outperformed either alone in the prediction of outcome, and the use of the two indices in combination for stratification of patients into three levels of risk provided more accurate prognostic information. The tamoxifen-predictive capacity of HOXB13:IL17BR, in combination with MGI-prediction of pathologic response to chemotherapy (unpublished data; Ma et al., 2008), indicates that the combined index may be an important predictor of benefit from both endocrine and chemotherapy agents. A valuable advantage using this combinatorial approach for risk classification is that the large group of patients with intermediate grade II tumors, for whom there is always a debate whether to give or withhold chemotherapy, is reclassified into low, intermediate or high risk (Ma et al., 2008).

**St Gallen consensus guidelines and Adjuvant! Online**

The consensus recommendations from the St Gallen conference and Adjuvant! Online⁴ utilize clinicopathological factors such as patient age, ER status, tumor grade, tumor size, number of affected lymph nodes and comorbidity data for clinical risk assessment. In contrast to the commercially available gene tests for outcome prediction, Adjuvant! Online is a web-based software free to use in order to facilitate decision making for adjuvant therapy, computing the 10-year risk for breast cancer mortality as well as benefit from adjuvant therapy (Ravdin et al., 2001). There are some major limitations, though, since HER2, PR and proliferative status is not taken into account in the risk model.

⁴ http://www.adjuvantonline.com
BACKGROUND

The International St Gallen Breast Cancer Conference on primary therapy for early breast cancer biennially gathers experts in the breast cancer field from all over the world. The concluding ‘St Gallen Consensus’ provides a set of recommendations on how to optimally treat primary breast cancer, and has a global impact on the standard approach in the management of early breast cancer with an international recognition and use.

The major shortcomings with the St Gallen guidelines and Adjuvant! Online is the usage of histopathological features of the tumors. These factors are evaluated manually, and hence, intra- and interobserver variations as well as subjectivity exist, which limits the accuracy and reliability of the assessment.
The general aim of this thesis was to explore and validate the prognostic and treatment-predictive utility of HOXB13 and the two-gene ratio HOXB13:IL17BR in pre- and postmenopausal breast cancer patients.

Paper I
- To investigate the significance of the HOXB13:IL17BR index, and the two genes separately, in prediction of benefit from prolonged tamoxifen treatment in postmenopausal breast cancer.
- To examine prognostic effects in a cohort of systemically untreated premenopausal breast cancer patients.

Paper II
- To validate the combined index of HOXB13:IL17BR and MGI.
- To develop and test a continuous predictor based on the dichotomous combined index of HOXB13:IL17BR+MGI for risk prediction at the individual level.

Paper III
- To quantify the protein expression of HOXB13 and relate different expression levels to the tamoxifen benefit in postmenopausal breast cancer.

Paper IV
- To assess the level of HOXB13 protein in a cohort of premenopausal breast tumors, and to analyze the tamoxifen treatment-predictive value.
Ethical permissions

All studies in this thesis were ethically approved. Paper I and IV were approved by the local ethical committee at Lund and Linköping Universities. Paper II and Paper III were approved by the ethical committee at the Karolinska University Hospital in Stockholm.

Pre- and postmenopausal patients (Paper I and Paper IV)

The premenopausal patients included in Paper I and IV were enrolled in a trial comparing the effect of tamoxifen treatment versus no treatment in relation to recurrence-free survival (RFS) and overall survival in premenopausal patients (Rydén et al., 2005). The study was conducted by two study centers in the south-east and south regions of Sweden, with enrollment between 1986 and 1991. All patients were premenopausal or under 50 years of age with stage II (T2 N0 M0, T1 N1 M0 and T2 N1 M0) invasive breast cancer. A patient was considered premenopausal until one year after the last menstruation. Primary surgery was modified radical mastectomy or breast-conserving surgery with axillary lymph node-dissection. Surgery was followed by radiotherapy. Less than two percent of the patients received adjuvant systemic therapy other than tamoxifen; either CMF or goserelin. Patients in the south-east region were treated with 40 mg of tamoxifen, whereas the daily administration in the south region was 20 mg. In postmenopausal patients, the treatment effects of a daily dose of 20 mg or 40 mg tamoxifen have been reported to be similar (Early Breast Cancer Trialists' Collaborative Group, 2005a).

A total of 564 patients were enrolled in the original trial, of which 276 were allocated to tamoxifen and 288 allocated to control receiving no adjuvant endocrine treatment. Tumor samples from 487 of these women were available as FFPE specimens for the study in Paper IV. For Paper I, only patients from the untreated cohort were chosen; 93 freshly frozen tumor samples were analyzed.
PATIENTS

A postmenopausal cohort was also included in Paper I. Those patients were accrued to participate in a randomized trial of two versus five years of adjuvant tamoxifen for postmenopausal breast cancer (Swedish Breast Cancer Cooperative Group, 1996). The purpose of the original trial was to determine the optimal duration of adjuvant tamoxifen therapy. In the early 1980s, a multicentric randomized trial was initiated, involving five regional breast cancer study organizations. Between 1983 and 1991, a total of 3,887 patients were randomly assigned to tamoxifen treatment with different durations. The patients were less than 75 years of age with operable, axillary lymph node-negative or –positive, invasive breast cancer, and received primary surgery of either modified radical mastectomy or breast-conserving surgery in combination with axillary lymph node dissection. Only one of the study centers used adjuvant CMF, leaving 2.5% of the patients who were recurrence-free after two years having received chemotherapy. RNA samples from a total of 264 fresh-frozen tumors from the south-east study center were available for the analyses conducted in paper I.

Postmenopausal patients (Paper II and III)

In 1976 the Stockholm trial of adjuvant tamoxifen among postmenopausal women was initiated by the Stockholm Breast Cancer Study Group (Rutqvist et al., 2007). The purpose was to evaluate the effect of adjuvant tamoxifen in a wide variety of patients, hence the enrollment of both node-negative (low-risk) and node-positive (high-risk) patients. The main inclusion criteria were postmenopausal menstrual status, (more than six months since last menstruation), age below 71 years and a unilateral invasive breast cancer. A total of 2,738 women with invasive early-stage disease were included, of which 1,780 were considered as having a low risk of recurrence, bearing tumors ≤ 30 mm with no lymph node-involvement. The patients were treated with either breast-conserving surgery (including axillary dissection) with postoperative radiotherapy, or modified radical mastectomy. The randomization allocated the patients to two years of 40 mg tamoxifen daily, or no endocrine treatment. The control group was systemically untreated and did not receive any endocrine treatment or chemotherapy. The allocation of cancer patients to an untreated control group was possible at that time, since no convincing evidence about the survival benefit of tamoxifen therapy existed. However, the trial was closed for patient entry in 1990, after it was demonstrated that tamoxifen improved the RFS (Fisher et al., 1989).

A new trial was initiated in 1983, in which the tamoxifen-treated patients who were recurrence-free at two years were once again randomized; this time to either discontinue tamoxifen or continue for three additional years (Rutqvist et al., 1987).
As a result, the patients in the tamoxifen arm were treated for either two or five years. After a long follow-up period, the treatment duration did not significantly affect the benefit (Rutqvist et al., 2007).

For the analysis of protein expression in Paper III, tumor samples from 912 patients were available as FFPE sections on a tissue microarray. In Paper II, RNA was extracted from FFPE blocks from 808 of the tumors.
Real-time RT-qPCR (Paper I and Paper II)

The development of real-time RT-qPCR has made it possible to monitor and quantify the progress of a PCR-reaction in real-time. Starting with RNA extracted from tissue, complementary DNA (cDNA) is synthesized and put into the PCR reaction as template. The extension of DNA strands in the reaction is then followed by monitoring the increasing levels of fluorescence as the detached probe is cleaved and fluorescence is emitted. RT-qPCR shows both high sensitivity and good reproducibility.

A prerequisite for a successful RT-qPCR result is the use of highly specific primers and probes, as is the choice of endogenous control genes. Normalization of the results using endogenous control genes as internal references is a requirement for obtaining reliable results from the assay. This approach controls for variations in amplification efficiency, extraction yield, RNA quality and efficiency of the cDNA-synthesis, and enables comparisons of mRNA concentrations across different samples. However, normalization of expression data by only one internal control is strongly dependent of the integrity of the RNA (Fleige et al., 2006). The MIQE (Minimum Information for Publication of Quantitative Real-Time PCR Experiments) guidelines strongly recommends multiple endogenous controls (Bustin et al., 2009).

The standard curve method is a simple and commonly used method in RT-qPCR. However, care must be taken so that the Ct-values (cycle threshold) for all samples fall within the interval of the calibration curve, since the function of PCR efficiencies for concentrations outside the dynamic range covered by the calibration samples may be nonlinear.
RNA extraction (Paper I and Paper II)

For use in RT-qPCR, the integrity of the input RNA is pivotal. RNA is an unstable molecule and formalin-fixation as well as improper handling and storing are detrimental to the quality of the RNA. Presence of RNases in the environment is also crucial to pay regard to in order to obtain high-quality and high-integrity RNA samples. RNA extraction from FFPE samples are challenging since the RNA is usually heavily degraded, there are cross-links between nucleic acids and proteins, and the fixation process typically introduces base modifications. As described above; low quality RNA as template in RT-qPCR may strongly compromise the results. With the adoption of microarrays and RT-qPCR methods in clinical practice, the methods for isolating RNA from mainly FFPE tissue have been improved so that RNA of sufficient quality can be extracted.

Immunohistochemistry of tissue microarray slides (III-IV)

Immunohistochemistry (IHC) is a commonly used method for the detection of proteins in a tissue. Factors such as section type, fixatives, antigen retrieval and detection method all affect the staining success. The use of archived material is enabled through preservation of the cells and tissues, which is obtained by incubation of fresh tissue in a fixative such as formaldehyde. The fixation process may mask the antigens in the tissue; therefore some kind of antigen-retrieval is required during the preparation of the samples.

In order to receive reliable results using IHC, specific antibodies are required. Monoclonal antibodies are expected to have the best specificity, but there is always a risk of cross-reactivity and false negative results if the only epitope for the antibody is lost. There is also the risk of unspecific background staining if the optimal washing buffers and blocking reagents are not used. Antibody specificity is usually confirmed with western blotting, but since the conditions of the different steps in this analysis differs from the IHC procedure, one can always argue about the validity of those confirmations. Problems with background and unspecific staining of the membranes can often be overcome by modification of the washing and blocking buffers, in a similar way as is done for IHC. However, despite diverse conditions and setup of the two methods, the results from a western blotting analysis still give some indications about the specificity of the antibody.

Some of the advantages using IHC are low cost, no need for expensive equipment, and the simplicity of the technique. During the grading of stained slides it is possible to see the morphology of the immunostained cells, which enables an exact
METHODOLOGICAL CONCERNS AND COMMENTS

assessment of which cell type the protein is expressed in. The major disadvantage with immunostaining of tissue sections is the subjectivity of the grading procedure, therefore, multiple investigators grading blinded to outcome is preferable.

The use of tissue microarray (TMA) slides makes it possible to stain several hundred tumor samples simultaneously. This eliminates inter-assay variability which is always an issue when running large numbers of whole-tissue slides, but on the other hand, tumor heterogeneity could pose a problem. However, biopsies in triplicates from each tumor make TMA comparable to whole-section analysis, according to validation studies (Camp et al., 2000, Torhorst et al., 2001, Fernebro et al., 2002). TMA also keeps the need for tumor tissue to a minimum.

Statistics and clinical endpoints (Paper I-IV)

RFS was defined by the time of diagnosis to the first of the following events; local, regional or distant recurrence, or breast cancer-related death. Breast cancer survival (BCS) was the time elapsed from diagnosis to the date of death due to breast cancer. Distant metastasis-free survival (DMFS; Paper II) and distant recurrence-free survival (DRFS) was defined as the time from diagnosis to first distant metastasis. Any local or regional recurrences prior to distant metastasis were censored at the time of relapse.

The statistical analysis in this thesis was performed in the statistical software Statistica 9.1 (and older versions; Statsoft Scandinavia AB, Sweden) and the free software environment R (version 2.11.1)5. P-values ≤0.05 were considered statistically significant. Hazard and recurrence ratios were specified with 95% confidence interval (CI).

Relationships between grouped variables were analyzed with the $\chi^2$ test for contingency tables or the $\chi^2$ test for trend when appropriate. Spearman’s rank order correlation was used to assess the relationship between two variables (HOXB13 and IL17BR expression levels in Paper I).

Survival curves were calculated by Kaplan-Meier estimates, and associations of tumor characteristics with the clinical endpoints were assessed with the use of log-rank test and Cox proportional hazard regression. The proportional hazard assumption was verified by scaled Schoenfeld residuals. To evaluate whether tumor characteristics provided predictive or prognostic information independently of other parameters, multivariate Cox proportional hazard regression models were used. A multivariate Cox analysis was also used to test the interaction between different tumor variables

5 http://www.r-project.org
METHODOLOGICAL CONCERNS AND COMMENTS

and treatment; in these cases, the model included the covariate, treatment and an interaction variable.

The continuous risk model that was built in Paper II, combined HOXB13:IL17BR and MGI as continuous variables in the tamoxifen-treated ER-positive patients. The linearity of the two variables was checked by fitting a Cox proportional hazard model with restricted cubic splines. When significant nonlinearity was demonstrated, a polynomial function of HOXB13:IL17BR was used to approximate the restricted cubic spline, and the final model was selected by comparing Cox regression models using Akaike information criterion. The resulting predictor from the final Cox regression model was rescaled into the range of zero to ten, and categorized into three levels of risk. The cut-offs were chosen so that the proportions of low, intermediate and high risk groups were similar to those formed by the three categorical combination groups of HOXB13:IL17BR and MGI. The risk model was tested in the cohort of endocrine untreated patients in the trial.

The hazard ratio (HR) for the continuous risk score was calculated relative to a 5-unit increment, except for in the multivariate analysis of the developed risk model and Adjuvant! Online where HRs were calculated relative to an increment of their inter-quartile ranges to ensure a more accurate comparison.
SUMMARY OF THE STUDIES

The findings in this thesis are based on observations and analyses of tumor characteristics and clinical outcome of patients enrolled in clinical trials examining the beneficial effects of tamoxifen therapy. The tumors were removed surgically and then stored fresh-frozen or as FFPE specimens. Tumors still available by the date of initiation of this thesis work were the starting material for RNA extraction and protein expression analysis. Patient data and clinicopathological tumor characteristics assessed by the time of diagnosis and during other research projects were retrieved from previous studies.

Paper I

In Paper I we explored the index composed of the ratio between HOXB13 and IL17BR expression in breast tumors. The two-gene ratio as a predictor of disease-free survival in breast cancer patients was developed by Ma et al. (2004), and it was demonstrated that high HOXB13 and low IL17BR expression levels were associated with worse clinical outcome. These results were confirmed by two other studies (Goetz et al., 2006, Ma et al., 2006), but another validation study performed prior to these failed to validate the predictive significance. However, that study was performed in a fairly small cohort (58 patients) and comprised mainly lymph node-positive patients (Reid et al., 2005). The two successful validation studies proposed that the principal value of the two-gene ratio may be found in patients without metastatic spreading to the lymph nodes.

Previous studies have included patient uniformly treated with tamoxifen and thus cannot for certain prove if the ratio is treatment predictive, prognostic, or both. Only one of the cohorts in a previously published study was untreated, which indicated a prognostic value (Ma et al., 2006). The patients in our study were two cohorts of post- and premenopausal patients with different treatment regimens. The 93 premenopausal patients for whom the gene expression analysis was successful were endocrine untreated, whereas the 264 postmenopausal patients received two or five years of adjuvant tamoxifen. Therefore we could perform a more explicit
SUMMARY OF THE STUDIES

investigation of the predictive value of HOXB13:IL17BR. Results from the statistical analysis showed that HOXB13:IL17BR identified patients with different treatment benefit; a low value in combination with prolonged therapy was beneficial for the patients, recurrence rate ratio (RRR) was 0.39 (95% CI: 0.17-0.91, p=0.30). A high ratio was associated with decreased response to therapy, as there was no significant difference in RFS between the two treatment arms (RRR: 0.95, 95% CI: 0.43-2.1, p=0.90). Analyses of subgroups stratified by lymph node status indicated the validity of the two-gene ratio in either group.

In this study, HOXB13 was the main determinant of the ratio, and therefore we analyzed the predictive power of the individual genes separately. A low HOXB13 expression was beneficial for patients receiving tamoxifen (RRR: 0.37, 95% CI: 0.17-0.83, p=0.015), whereas the patients with higher tumor expression of the gene had similar recurrence rates independently of treatment duration (RRR: 1.1, 95% CI: 0.50-2.6, p=0.75). The expression level of IL17BR could not distinguish patients with more or less benefit.

Neither the ratio nor HOXB13 showed any prognostic value in the untreated cohort, univariate analysis of the two variables did not show any statistical significances (RRR for the ratio: 1.18, 95% CI: 0.98-1.42, p=0.084; RRR for HOXB13: 1.05, 95% CI: 0.84-1.31, p=0.69). However, IL17BR may be able to identify patients with worse outcome; RRR in univariate and multivariate analysis were 0.75 (95% CI: 0.57-0.99, p= 0.042) and 0.73 (95% CI: 0.53-1.01, p=0.058), respectively.

Taken together, these results demonstrated that a high HOXB13:IL17BR ratio or a high HOXB13 expression was associated with decreased benefit of prolonged tamoxifen treatment for breast cancer patients and hence, the two-gene ratio could potentially serve as a biomarker for prediction of recurrence after treatment with tamoxifen.

Paper II

The invaluable utility of prognostic and predictive gene expression signatures adjunctive to standard risk factors in breast cancer clinic has been demonstrated. For better informed treatment decision-making, novel tools based on microarray or RT-qPCR provide a way of identifying low- and high-risk patients. However, there will always be an uncertainty about how to deal with patients predicted to have an intermediate risk of recurrence, which demands a high precision in the risk assessment in order to reclassify patients from this large group to the other levels of risk, and hence get clearer indications about the most suitable treatment strategy.
Ma and colleagues (2008) demonstrated that the combined index of binary HOXB13:IL17BR and MGI were complementary prognostic factors outperforming either alone in predicting risk of recurrence in breast cancer patients. Using this combinatorial approach, a risk classification of distant metastasis could be made, stratifying patients into three risk groups. Our study, which is described in Paper II, was performed in collaboration with employees of bioTheranostics, Inc. (former Aviara Dx, Inc. and Arcturus Bioscience, Inc.) and the Molecular Pathology Research Unit at Massachusetts General Hospital. These are named inventors on a patent to use the HOXB13:IL17BR ratio for breast cancer prognosis. Herein, we utilized a cohort of 808 patients from the randomized Stockholm trial for validation of the prognostic utility of the combined index. RNA was extracted from FFPE tissue, and the subsequent gene expression analysis was successfully for a total of 769 cases, of which 588 were ER-positive and included in the validation analysis.

The dichotomous HOXB13:IL17BR+MGI index was significantly associated with DMFS and BCS in both the tamoxifen-treated and the endocrine untreated cohorts. Patients assigned to the low-risk group (more than 50% of the entire tamoxifen-cohort) had very low 10-year rates of distant recurrence and death (distant recurrence rate: 2.9%, 95% CI: 0.4-5.4; death rate: 2.3%, 95% CI: 0.1-4.5). For the other risk groups, the distant recurrence rates were increased but were similar between the two, with intermediate risk corresponding to a rate of 16.9% (95% CI: 7.2-25.6) and high risk 16.3% (95% CI: 6.0-25.5). To these groups, 23% respectively 18% of the patients were assigned. This risk classification of recurrence rates and death rates was confirmed with multivariate analysis to be independent of tumor size, grade, HER2 status and PR status. In the tamoxifen-untreated cohort, 50%, 27% and 23% of the patients were identified as having low, intermediate and high risk of distant metastasis. Combination of the two indices also demonstrated prognostic utility (p=0.0004).

We brought the combined index of HOXB13:IL17BR+MGI a step further and developed a continuous risk index to allow for individual risk assessment of the recurrence risk. In order to maximize the accuracy, the algorithm was trained to retain the entirety of the prognostic information available using the ER-positive tamoxifen-treated cohort. The algorithm, called Breast Cancer Index (BCI), assigns each patient an individual risk score between zero and ten, corresponding to a certain level of risk of distant recurrence. Using the same proportional values obtained with the combined index, the BCI scores were categorized into three levels of recurrence risk (low, intermediate and high).

The initial analysis in the training of BCI, almost 60% of the tamoxifen-treated patients were estimated to have a rate of distant recurrence of 1.7% (95% CI: 0-3.5)
SUMMARY OF THE STUDIES

and a death rate of 1.1% (95% CI: 0-2.6). The prognostic ability of BCI was further analyzed in a validation set comprising the endocrine untreated patients from the trial. In this cohort, a majority (53%) of the patients were assigned to low risk, leaving 27% and 20% in the other risk groups. The lowest rate of distant metastasis was 8.3% (95% CI: 4.7-14.4), increasing to 22.9% (95% CI: 14.5-23.2) and 28.8% (95% CI: 15.3-40.2) for the groups with worse outcome. As a comparison, the St Gallen guidelines were applied to the cohorts analyzed in this study. With those recommendations for risk classification, 78% of the tamoxifen-treated and 81% of the untreated patients were predicted to have intermediate risk of recurrence.

Further testing of the ability of BCI showed that it was a strong prognostic factor for distant recurrence and death, independent of tumor size, grade, PR status and HER2 status. Altogether, BCI was a strong prognostic for early breast cancer, and the results suggest the prognostic utility also extends into untreated patients.

Paper III

There have been several papers investigating and validating the utility of the two-gene ratio HOXB13:IL17BR for predicting benefit from tamoxifen and identifying patients who would benefit from additional chemotherapy (Ma et al., 2004, Goetz et al., 2006, Ma et al., 2006, Jansen et al., 2007, Wang et al., 2007, Ma et al., 2008). As we showed in Paper I, HOXB13 alone could predict recurrence in tamoxifen-treated patients, independently of IL17BR. Similar to the ratio, higher levels of HOXB13 were associated with a decreased response to therapy.

No studies have been published on the protein expression level of HOXB13 in breast tumors, and how it is related to treatment response and clinical outcome. With the large cohort from the randomized Stockholm trial available as FFPE tissue on TMAs, we had the opportunity to investigate the HOXB13 protein expression and its possible association with clinical outcome in a randomized material.

We used immunohistochemistry and a monoclonal antibody directed towards HOXB13 to determine the protein expression in the tumors. Validation of the antibody with western blot gave a single band corresponding to the expected protein weight. The assay was successful in a total of 866 cases, which were included in the survival analysis. Almost 30% of the tumors showed moderate to high staining, about 37% were weakly stained and 34% displayed no detectable expression.

HOXB13 protein was associated with tamoxifen response in our ER-positive study cohort. Both DRFS and BCS increased with tamoxifen therapy if tumor expression of HOXB13 was low (p=0.00002 and p=0.00008, respectively). HR for distant recurrences
was 0.38 (95% CI: 0.23-0.60, p=0.000048) for low HOXB13 expression, and 0.88 (95% CI: 0.47-1.65, p=0.69) for high, and were statistically different (p for interaction: 0.035). Similar numbers were obtained for HR for breast cancer death. Subgroup analysis of patients positive for ER, but without any PR expression, showed that HOXB13 could not predict the outcome after therapy in these patients. No benefit from the endocrine treatment could be detected, regardless of HOXB13 status. PR is the most commonly used indicator of functional ER-signaling (Lapidus et al., 1998), and our observations may reflect the possibility that those unresponsive tumors are more dependent on another signaling pathway than ER for proliferation. Consequently, HOXB13 has a predictive value in addition to PR.

In line with previous studies, HOXB13 did not provide any prognostic information for the untreated patient cohort. Groups with different expression levels had similar DRFS and BCS (HR for DRFS: 0.89, 95% CI: 0.58-1.36, p=0.59; HR for BCS: 0.76, 95% CI: 0.46-1.25, p=0.28).

In conclusion, this study is the first to report about HOXB13 protein expression and its utility as a predictor of response to tamoxifen treatment. In conformity with studies of HOXB13 mRNA, higher expression levels identify patients less likely to benefit from the therapy.

**Paper IV**

Most investigations of HOXB13 in breast cancer have been performed in patient cohorts with mixed age structure, with no regard to pre- and postmenopausal status. With the notion of HOXB13 being an ER-responsive gene, and that treatment with estradiol downregulates expression, in combination with the different hormone levels before and after menopause, it is likely that the expression pattern of HOXB13 also differs with regard to menopausal status. Our previous study on the two-gene ratio and the IHC-study of HOXB13 protein expression are, to our knowledge, the only with strict pre- and postmenopausal subsets.

Herein, we explored HOXB13 protein expression using IHC in a series of 487 premenopausal patients randomized to tamoxifen or no endocrine treatment. The procedure for protein quantification was identical to our previous study described in Paper III, and was successful in 367 cases. The distribution of staining intensities was: negative 28%; weak 48%, moderate 17% and strong 7%, which were similar to what we previously reported for the postmenopausal patients.

In line with previous studies, prognosis could not be predicted using HOXB13 data, as there were no significant differences in recurrence or death rate between groups.
with different protein expression (RFS HR: 0.97, 95% CI: 0.61-1.54, p=0.89; RFS HR: 0.84, 95% CI: 0.50-1.43, p=0.53). Interestingly, in this series of patients, HOXB13 was not a predictor of treatment effect; all patients receiving tamoxifen had a better outcome than untreated patients regardless of HOXB13 status (HR for recurrences for the HOXB13-low group: 0.69, 95% CI: 0.44-1.07, p=0.1; HR for recurrences for high HOXB13: 0.43, 95% CI: 0.18-1.01, p=0.054).

A strong correlation between HOXB13 and ERβ indicated a possible impact on the performance of HOXB13 as a predictor of outcome. Therefore, we stratified the patients on ERβ status for the subsequent analyses of survival. The results showed that ERβ may have a role as a modifier of HOXB13’s performance in identifying patients with lack of benefit from the treatment. Absence of this hormone receptor, in combination with a moderate to high HOXB13 level, made the tumors respond to tamoxifen, decreasing the HR for recurrences as compared to the group with low or no expression of HOXB13 (HR for high expression: 0.15, 95% CI: 0.02-1.27, p=0.08; HR for low HOXB13 expression: 1.24, 95% CI: 0.66-2.36, p=0.50).

For the tumors with detectable levels of ERβ, the treatment response was opposite to what was seen in the ERβ-negative group. Herein, there was a tendency that tamoxifen was only beneficial with a simultaneous lack of HOXB13 (HR for low HOXB13: 0.31, 95% CI: 0.09-1.08, p=0.065; HR for high HOXB13: 1.32, 95% CI: 0.37-4.70, p=0.67.

The interaction between HOXB13 and treatment effect was statistically significant in a multivariate model including tumor size, tumor grade and lymph node status, with a p-value of 0.04 for ERβ-negative tumors.

In conclusion, this initial study of HOXB13 protein expression in premenopausal breast cancer indicates ERβ being a modifier of predictive performance of HOXB13. Thus, ERβ status should be used, in addition to HOXB13, to identify premenopausal patients less likely to respond to tamoxifen therapy.
CONCLUDING REMARKS AND FUTURE ASPECTS

Breast cancer is probably the disease most extensively studied by gene expression profiling, but despite a steady increase in the number of discovered prognostic and predictive factors, there are only a few biomarkers adopted into clinical routine practice. A major issue in the management of breast cancer is to identify those patients who are less likely to be cured after primary treatment and probably would benefit from adjuvant chemotherapy. Likewise, identification of patients with only local disease who would be given chemotherapy but might have survived without is also of great importance. Advances in gene expression analysis have made it possible to refine disease classification, improve diagnostic and prognostic accuracy, develop new biological concepts and to identify new molecular targets for drugs.

The HOXB13:IL17BR expression signature in combination with MGI forms BCI, which has demonstrated prognostic as well as predictive utility in the clinical setting. However, independent validation studies and prospective clinical trials are needed in order to completely determine the benefit of BCI in treatment decision making for breast cancer patients. So far, we have seen that BCI has the potential of improving prognostic and therapeutic prediction, but additional tests are required.

Even with successful independent validations and subsequent implementation of a gene expression-based diagnostic test in the clinic, there are still some questions remaining; what is the role of the individual genes, and how are they related to carcinogenesis and treatment response? There are many studies on the function and effects of HOXB13 in different normal and cancerous tissue in general, but its role in breast cancer, and in tamoxifen resistance in particular, is not yet fully elucidated. Most of the correlations demonstrated remain unexplained. However, a few experimental studies have increased the understanding of HOXB13 functions, and studies performed in other tissues than the breast may provide clues to where to start searching for key features of this gene. Studies have shown that HOXB13 enhances EGF-induced cell migration and invasion (Ma et al., 2004), but the exact mechanism by which this is mediated is not known.
CONCLUDING REMARKS AND FUTURE ASPECTS

The causes of the increased HOXB13 levels seen in breast cancer are not known, but plausible explanations could be on the epigenetic level, as well as gene amplifications. These mechanisms still remain to be investigated. However, gene methylation studies by two independent investigators showed that HOXB13 can be hypermethylated in breast tumors, and this was associated with decreased gene expression and worse patient outcome (Rodríguez et al., 2008, Tommasi et al., 2009). These somewhat conflicting results make the question about epigenetics even more important. One could hypothesize that upregulation is a two-step process starting with epigenetic changes in early tumor development, followed by gene amplification or deregulation of signaling networks regulating the mRNA and protein expression of HOXB13.

The findings that HOXB13 may be a coactivating factor in ER-signaling (Jung et al., 2004b) and HOXB13 being an ER-responsive gene which is downregulated upon treatment with estradiol (Wang et al., 2007), are probably the initial keys to find out how HOXB13 confers tamoxifen resistance. This can also bring us one step closer towards an understanding of how the expression levels are regulated. Wang et al. (2007) provides an appealing hypothesis about HOXB13 regulation in breast cancer. With ER functioning as a downregulator of this gene, hormone-responsive tumors will have suppressed HOXB13 expression as long as signaling from the receptor is fully functional. If estrogen-signaling is impaired, which is a common feature of breast tumors, the suppression is lost and HOXB13 levels increase. Provided that HOXB13 may contribute to increased ER-activity in absence of estradiol, a natural consequence would be that the cells are unaffected by tamoxifen treatment.

In a study of prostate cancer cells, it was demonstrated that HOXB13 is directly regulated by the transcription factor FOXA1 (McMullin et al., 2010), a finding which may have important implications in breast cancer. FOXA1 functions as an important cofactor to ER in regulating gene networks in the breast (Carroll et al., 2005), and may act as the activating trigger of HOXB13 upregulation when ER-signaling is dysregulated. However, since premenopausal women naturally have higher systemic estradiol concentrations than postmenopausal women, this mechanism is more likely to affect tumors developing after menopause when the ovarian production of estradiol is ceased.

As we reported in Paper IV, ERβ was a modifier of the predictive performance of HOXB13, which suggests a mechanistic connection between HOXB13 and ERβ. ERβ is abundantly expressed in prostate, which is a more thoroughly investigated tissue regarding HOXB13, as compared to the breast. During prostate tumor development the expression of HOXB13 is commonly lost (Pasquali et al., 2001), a mechanism
which is also seen in breast carcinomas. Perhaps the knowledge about the function of HOXB13 in prostate cells can provide clues towards a greater understanding of this interplay in breast tissue.
GENERAL CONCLUSIONS

Modern technologies provide valuable tools for the clinical management of breast cancer; a natural step towards an era of personalized medicine. Prior to implementation of novel biomarkers and gene expression assays in the clinic, thorough validations are required. In this thesis, we have validated a promising novel gene assay for prediction of tamoxifen benefit and also studied the prognostic and predictive utility of a single biomarker in breast cancer medicine.

Our studies showed:

- The two-gene ratio HOXB13:IL17BR or HOXB13 mRNA expression alone identifies patients who are likely to benefit from tamoxifen treatment and therefore may be useful as treatment predictive markers.
- The continuous risk index, BCI, which is a further development of HOXB13:IL17BR and MGI, identifies a major proportion of lymph node-negative patients with a very low risk of distant recurrence. It has also significant prognostic value and may therefore help clinicians to make better informed treatment decisions in order to spare toxic chemotherapy for a large group of breast cancer patients.
- HOXB13 protein expression may be used as a predictive marker for tamoxifen benefit. However, ERβ seems to modify the performance of HOXB13 and should be taken into account when identifying premenopausal patients with less likelihood of response to therapy.
ACKNOWLEDGEMENTS

Först och främst, ett stort TACK till alla de bröstcancerpatienter som genom åren gett sitt samtycke till att vara med i de kliniska studier varifrån jag har hämtat de material jag har jobbat med. Utan ER skulle forskningen inte gå framåt!

Forskning är inte heller möjligt utan generösa bidragsgivare; Lions Club i Karlshamn, Svenska Läkaresällskapet, Cancerfonden samt Knut och Alice Wallenbergs stiftelse har alla stöttat mitt arbete genom personliga stipendier och reseanslag. Vi har även fått ekonomiskt stöd från Cancerfonden och Vetenskapsrådet.

Det är många som på ett eller annat sätt har bidragit till det arbete som ligger till grund för den här avhandlingen. Små, till synes obetydliga gester kan ha stor betydelse i långa loppet, och kan vara det som förändrar en hel dag till det bättre. Några personer vill jag tacka speciellt...

Olle Stål, min handledare. Om jag hade en bråkdel av all din kunskap skulle jag vara jätteglad; jag jobbar på det! Genom ditt oerhörda lugn i alla situationer och din tilltro till min förmåga har min doktorandperiod förflutit på bästa tänkbara sätt. Ditt handledarskap, som jag upplevt som ‘frihet under ansvar’, har gjort att jag sakta men säkert hittat min väg inom forskningen. Ibland dyker chanser upp mitt framför näsan på en, och vem hade kunnat tro att mitt första projekt skulle leda till ett samarbete som resulterade i en lång San Diego-vistelse, ett föredrag i San Antonio och till sist en publikation, och förhoppningsvis ännu mer i långa loppet. TACK för alla möjligheter jag har fått!

Agneta Jansson, Olles sidekick som handledare. Jag hade inte kunnat önska mig någon bättre bihandledare! Din effektivitet och framåtande smittar av sig och höjer standarden på mitt arbete. Jag är oerhört tacksam för allt du lärt mig, bl.a. vikten av en välorganiserad (?) labbok och ordning och reda på labbet. Din ambitionsnivå och driftighet (och givetvis ditt svarta bälte i shopping) gör dig till min stora förebild!

Bibbi. Den som har koll på allt och fixar allt. Utan dig skulle tillvaron på labbet inte vara lika enkel och smidig.
Tove Sivik. För värdefulla samtal om forskning, förhållningssätt och framåtanda. Med ambitioner och god vilja kommer man långt! DU kan gå hur långt du vill!


Mark Erlander at bioTheranostics, Inc., who made our collaboration possible, and for sharing your great knowledge about research and molecular diagnostics. (Yes, I WILL go to US for my postdoc!) Ranelle Salunga, Eden Estepa-Sabal, Keith McCormick and all the other employees at bioTheranostics who made my stay in San Diego the best possible. I enjoyed every single day! Nicole Kesty and Xiao-Jun Ma, for great collaboration during the manuscript writing process. It DID turn into a publication, eventually! Thao Ho and Yen Tran for excellent technical assistance with the never ending tumor samples, and nice company in the lab.

Mina medförfattare. En lång lista, men främst Sara Brommesson och Mårten Fernö.

Rolf & Ralf – Robert Carlsson och Roland Magnusson. ”Mina” studenter som jag handlett på labbet. Det är viktigt att kunna våga vatten!


Håkan Wiktander, Ammi Fornander och Pia Karlsson, för allt ni fixar på KEF!

Att vara doktorand på KEF innebär inte att man har ETT forskarutbildningsämne; det är en multidisciplinär utbildning där man lär sig om så vitt skilda ämnen som barnuppoftran, graviditet och förlössning, barnsjukdomar, trädgårdsarbete, matlagning, bakning, mode, litteratur och en hel massa annat! Tack alla ni underbara
medarbetare på KEF som förgyller fikastunder, lunchpauser och annan okynnestid som spenderas i fikarummet! Inga frågor är för svåra för snillen som spekulerar!

Och till sist, några av de som ger mig ett liv utanför universitetet...

Min PT på Worldclass, Daniel Winnerfeldt. Svett, mjölksyra och träningsvärk (och kramp!) i maximala doser! Jag uppskattar varenda sekund (nåja, efteråt i alla fall; ditt nöje, min pina)! Ditt tålamod och din förmåga att pressa mig till det yttersta (och dina ständiga påminnelser om att andas!) gör mig starkare än någonsin. Med dig bredvid mig i gymmet eller på ett box-pass är alla tankar på jobbet som bortblåsta; då är det bara Starkare, Snabbare och Högre som gäller!

Fina Familjen Eng; Camilla, Stefan, Jacob och Astrid. För middagar, luncher (ofta väldigt högljudda och livliga 😊), fikastunder, spelkvällar och alla andra gånger vi träffas som får mig att släppa jobbet för en stund. Vi har mycket att ta igen nu när boken äntligen är klar!

Min familj. Pappa, Mamma och alla mina syskon: Roy, Elisabet, Camilla, Sonny och Agneta. Ni har inte så bra koll på vad jag har sysslat med under dessa år ('någonting med cancer, celler och gener'), men jag vet att ni är stolta och det stärker mig! Och vet ni, NU ska jag inte vara student längre!

Per. Min fasta, trygga punkt i tillvaron.♡
REFERENCES


Bloom, HJ and Richardson, WW (1957) Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. Br J Cancer 11(3): 359-77.

REFERENCES


Harvey, JM, Clark, GM, Osborne, CK and Allred, DC (1999) Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for


REFERENCES


REFERENCES


REFERENCES


REFERENCES


REFERENCES


