Genetic and Epigenetic Characterisation of Breast Tumours

Holm, Karolina

2011

Genetic and Epigenetic Characterisation of Breast Tumours

Karolina Holm

LUND UNIVERSITY
Faculty of Medicine

Division of Oncology, Department of Clinical Sciences, Lund
Lund University, Sweden
2011

Faculty of Medicine,
Division of Oncology, Department of Clinical Sciences, Lund
Lund University

ISBN 978-91-86871-01-7
ISSN 1652-8220

Cover illustration by Kristina Persson
Printed by MediaTryck, Lund, 2011

To my family

Table of Contents

List of Papers

This thesis is based on the following original papers, referred to in the text by their Roman numerals:

I. Saal LH*, **Holm K***, Maurer M, Memeo L, Su T, Wang X, Yu JS, Malmström P, Mansukhani M, Enoksson J, Hibshoosh H, Borg Å, Parsons R. *PIK3CA mutations correlate with hormone receptors, node metastasis, and ERBB2, and are mutually exclusive with PTEN loss in human breast carcinoma*. Cancer Res. 2005 Apr 65(7):2554-9.
* These authors contributed equally.

II. **Holm K**, Staaf J, Jönsson G, Vallon-Christersson J, Gunnarsson H, Arason A, Magnusson L, Barkardottir RB, Hegardt C, Ringnér M, Borg Å. *Characterisation of amplification patterns and target genes at 11q13 in sporadic and familial breast cancer*. Manuscript.

III. **Holm K**, Hegardt C, Staaf J, Vallon-Christersson J, Jönsson G, Olsson H, Borg Å, Ringnér M. *Molecular subtypes of breast cancer are associated with characteristic DNA methylation patterns*. Breast Cancer Res. 2010;12(3):R36.

IV. **Holm K**, Grabau D, Lövgren K, Aradottir S, Bendahl P, Rydén L, Stål O, Malmström P, Fernö M, Hegardt C, Borg Å, Ringnér M. *Global H3K27 trimethylation and EZH2 abundance in breast tumour subtypes*. Manuscript.

Publications not included in the thesis:

- Saal LH, Johansson P, **Holm K**, Gruvberger-Saal SK, She QB, Maurer M, Koujak S, Ferrando AA, Malmström P, Memeo L, Isola J, Bendahl PO, Rosen N, Hibshoosh H, Ringnér M, Borg Å, Parsons R. *Poor prognosis in carcinoma is associated with a gene expression signature of aberrant PTEN tumor suppressor pathway activity*. Proc Natl Acad Sci U S A. 2007 May 1;104(18):7564-9.

- Saal LH, Gruvberger-Saal SK, Persson C, Lövgren K, Jumppanen M, Staaf J, Jönsson G, Pires MM, Maurer M, **Holm K**, Koujak S, Subramaniyam S, Vallon-Christersson J, Olsson H, Su T, Memeo L, Ludwig T, Ethier SP, Krogh M, Szabolcs M, Murty VV, Isola J, Hibshoosh H, Parsons R, Borg Å. *Recurrent gross mutations of the PTEN tumor suppressor gene in breast cancers with deficient DSB repair*. Nat Genet. 2008 Jan;40(1):102-7.

- Lundgren K, **Holm K**, Nordenskjöld B, Borg Å, Landberg G. *Gene products of chromosome 11q and their association with CCND1 gene amplification and tamoxifen resistance in premenopausal breast cancer*. Breast Cancer Res. 2008;10(5):R81.

- Staaf J, Ringnér M, Vallon-Christersson J, Jönsson G, Bendahl PO, **Holm K**, Arason A, Gunnarsson H, Hegardt C, Agnarsson BA, Luts L, Grabau D, Fernö M, Malmström PO, Johannsson OT, Loman N, Barkardottir RB, Borg Å. *Identification of subtypes in human epidermal growth factor receptor 2--positive breast cancer reveals a gene signature prognostic of outcome*. J Clin Oncol. 2010 Apr 10;28(11):1813-20.

- Jönsson G, Staaf J, Vallon-Christersson J, Ringnér M, **Holm K**, Hegardt C, Gunnarsson H, Fagerholm R, Strand C, Agnarsson BA, Kilpivaara O, Luts L, Heikkilä P, Aittomäki K, Blomqvist C, Loman N, Malmström P, Olsson H, Johannsson OT, Arason A, Nevanlinna H, Barkardottir RB, Borg Å. *Genomic subtypes of breast cancer identified by array-comparative genomic hybridization display distinct molecular and clinical characteristics*. Breast Cancer Res. 2010;12(3):R42.

Abbreviations

aCGH	Array comparative genomic hybridisation
ANOVA	Analysis of variance
BAC	Bacterial artificial chromosome
CCND1+	*CCND1*-amplified
CCND1-	*CCND1*-non-amplified
CDK	Cyclin dependant kinase
CpG	Cytosine followed by a guanosine
Cy3	Cyanine 3
Cy5	Cyanine 5
DDFS	Distant disease free survival
ddNTP	Dideoxynucleoside triphosphate
DNMT	DNA methyl transferase
dNTP	Deoxynucleoside triphosphate
ES	Embryonic stem
GEX	Gene expression
DAB	Diaminobenzidine
DMFS	Distant metastasis free survival
ER	Oestrogen receptor
FGA	Fraction of genome altered
H1K26	Histone 1, lysine 26
H3K27	Histone 3, lysine 27
H3K4me3	Trimethylation of lysine 4 on histone 3
H3K9me3	Trimethylation of lysine 9 on histone 3
H3K27me3	Trimethylation of lysine 27 on histone 3
hESC	Human embryonic stem cell
FISH	Fluorescence *in situ* hybridisation
IHC	Immunohistochemistry
lncRNA	Long non-coding RNA
MEV	MultiExperiment viewer
OS	Overall survival
PcG	Polycomb group
PI3K	Phosphatidylinositol 3-kinase
PIP2	Phosphatidylinositol-4,5-bisphosphate
PIP3	Phosphatidylinositol-3,4,5-triphosphate
PR	Progesterone receptor

PRC2	Polycomb repressive complex 2
RTK	Receptor tyrosine kinase
SAM	Significance analysis of microarrays
TMA	Tissue microarray
UV	Ultraviolet

Abstract

Breast tumours harbour a large amount of genetic and epigenetic alterations, which are associated with *e.g.* tumour aggressiveness, prognosis, and response to therapy. The heterogeneity of breast tumours is reflected in the identification of at least five molecular subtypes named basal-like, luminal A, luminal B, HER2-enriched and normal-like, which are believed to originate from different cell types and follow different progression pathways. In the current thesis, different genetic and epigenetic alterations of breast tumours have been studied and analysed in relation to the molecular subtypes and clinicopathologic variables. In Paper I, we studied genetic alterations of *PIK3CA* and *PTEN* as well as PTEN protein expression. We found frequent alterations in the two PI3K pathway components *PIK3CA* (26%) and PTEN (31%). The alterations in *PIK3CA* were associated with oestrogen receptor (ER) positivity, whereas PTEN predominantly was lost in ER-negative tumours. In Paper II, we analysed genomic aberrations at chromosome 11q13 in *CCND1* (11q13.3)-amplified breast tumours and identified cores proximal and distal of the *CCND1* locus that were frequently amplified. Additionally, we found that *CCND1* amplification and overexpression was most frequent in luminal B tumours. In Papers III and IV, we have focused on epigenetic studies of breast tumours. Using methylation arrays, we found specific methylation patterns and frequencies for luminal A, luminal B and basal-like tumours. Interestingly, a substantial amount of genes with subtype-specific expression appears to be regulated by DNA methylation. In addition, we found high gene expression of the Polycomb repressive complex 2 (PRC2)-member *EZH2* and low methylation frequency in basal-like tumours, indicating alternative epigenetic silencing mechanism in these tumours. EZH2 is the key member of PRC2 that catalyses the histone modification trimethylation of lysine 27 on histone 3 (H3K27me3). In Paper IV, we validated our previous findings of *EZH2* gene expression on protein level and found an identical pattern across the subtypes as well as identifying differential occurrence of H3K27me3 across the subtypes. Together these results add another layer to the heterogeneous nature of breast tumours.

Aims of the Thesis

The overall aim of the thesis has been to characterise breast tumours on different levels to add further layers of information to the genetic and epigenetic alterations in breast cancer.

Specific aims of the included papers were:

I To investigate the frequency and extent of PI3K pathway alterations in a large series of breast tumours and cell lines, and to correlate mutations to clinicopathologic characteristics in sporadic and hereditary breast tumours.

II To describe the amplification patterns on chromosome 11q13 in *CCND1*-amplified tumours in relation to breast cancer subtypes and clinicopathologic data.

III To investigate methylation patterns in the molecular subtypes of breast tumours and to what extent methylation affects the phenotypic behaviour of respective subtype.

IV To validate on protein level our findings of differential gene expression patterns of EZH2 in the molecular subtypes of breast tumours. Additionally, we wanted to depict the occurrence of H3K27me3 in relation to EZH2 expression as well as breast tumour subtypes.

Background

The Normal Breast

The task of the mammal breast is to produce milk to feed the offspring. From infancy to pre-puberty there is no difference between female and male breasts, anatomically. However, the female breast undergoes considerable changes in size, shape and structure during puberty, pregnancy, lactation and postmenopausal regression [1].

The breast consists of (1) milk lobules, (where the milk is produced), (2) milk ducts (leads milk to the nipple), (3) fatty tissue, (4) stroma, blood vessels, and lymph ducts (numbers referring to Figure 1). The milk ducts and lobules are in turn composed of two cell layers: the outer basal myoepithelial cell layer and the inner luminal epithelial cell layer facing the lumen. The myoepithelial cells have contractory functions while the luminal cells have secretory functions.

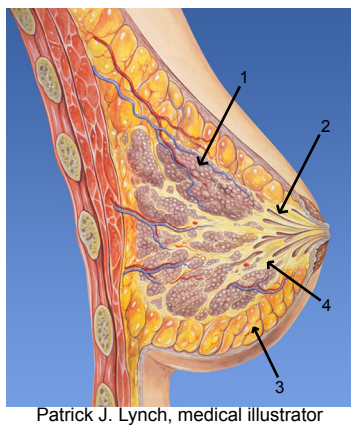

Figure 1. Breast anatomy. Illustration of the normal human female breast. Image adapted from P. Lynch, http://www.wikimedia.org.

The development of the breast is strongly dependent on hormones, especially oestrogens and progesterone. During puberty, lobules are formed as the ovaries start secreting oestrogens, but the complete formation of the breast is not finished until the end of the first full term pregnancy [1]. In the course of a pregnancy, the lobules and ducts differentiate and enlarge to produce and secrete milk when the offspring is born. After lactation, the milk-secreting lobules and ducts regress, however, until menopause more glandular tissue is seen in the breast than if no pregnancy and lactation had occurred [1]. Finally, after menopause, the breast undergo a regression as oestrogen and progesterone production cease.

Hallmarks of Cancer

Cancer is a disease of the genes with several hallmarks shared between tumours irrespective of the tissue from which the tumour develops [2]. The common core of cancer is the malfunctioning of circuits that guard normal cell proliferation. More than 100 types of cancer exist, which in turn can be divided into different subtypes within specific organs [2].

Hanahan and Weinberg propose in their classic review from 2000 six essential hallmarks that collectively dictate malignant growth: (1) self-sufficiency in growth signals, (2) insensitivity to growth-inhibitory signals, (3) evasion of apoptosis, (4) limitless replicative potential, (5) sustained angiogenesis, and (6) tissue invasion and metastasis [2]. Recently, the same authors published a review in which they updated the existing hallmarks as well as added two new hallmarks and two enabling characteristics, which have gained increasing importance during the progress of cancer research since the original publication (Figure 2) [3]. The two enabling characteristics are genomic instability and tumour-promoting inflammation, while the two new hallmarks are deregulation of cellular energetics and avoidance of immune destruction [3].

Normal cell growth is steered by growth signals that activate transmembrane receptors, and growth-inhibitory signals that manoeuvre the cell either into inactive states or into terminal differentiation. These processes are surveyed and defective cells, *e.g.* cells with DNA damages, undergo programmed cell death called apoptosis. Telomeres are the protective ends of chromosomes, which preserve the chromosome from deterioration. During each cell division the telomeres are shortened, a process that eventually would lead to DNA damages if cells could divide an unlimited number

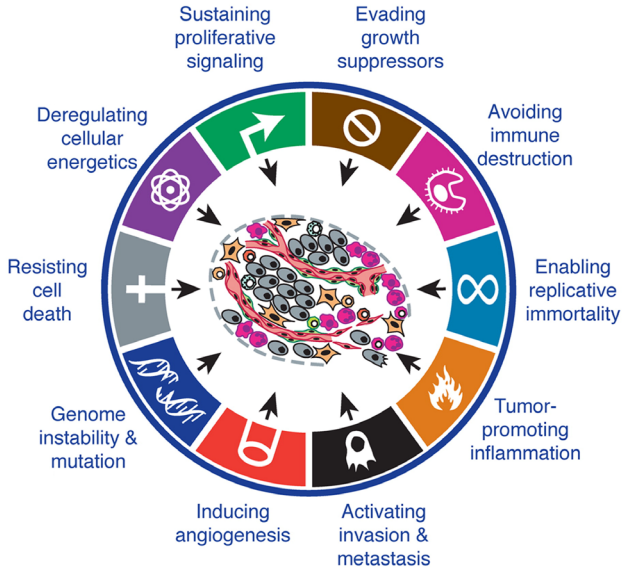

Figure 2. Hallmarks of cancer. Adapted from figure 6 in Hanahan *et al.* [3], reprinted with permission from Elsevier.

of times. Telomerase is the enzyme responsible for replenishing the telomeres but is normally not active in somatic cells. When cells have undergone a certain number of doublings they stop growing, a process called senescence. Furthermore, for a normal cell to grow it needs oxygen and nutrients brought to the cells by blood vessels. The final growth-limiting step for a normal cell is by contact inhibition since the cell division of normal cells is inhibited when they come in close contact with each other.

Tumour cells have, in many different ways, escaped these regulatory steps (numbers refer to the hallmark numbering). (1 and 2) Increased growth signalling can be obtained by endogenous overexpression of growth factor receptors and/or ligands, signalling to surrounding normal cells to supply cancer cells with excessive growth factors, or alternatively by aberration of downstream factors leading to constitutive proliferative signalling. (3) Escape of apoptosis is mediated in tumour cells by aberration of proapoptotic and/or antiapoptotic signalling. (4) Additionally, telomerase is frequently upregulated in tumour cells leading to endless replicative potential. (5) Solid tumours can, as they grow, recruit blood vessels and induce them to grow and support the tumour cells with oxygen and nutrients. (6) Finally, tumour cells have acquired the ability to penetrate blood vessels and the lymphatic system and spread to other parts of the body,

which is the ability that leads to the vast majority of cancer-related deaths.

The genomic instability of tumour cells is reflected in the accumulation of mutations, deletions, insertions and other rearrangements that are advantageous for tumour growth. However, as a response to the development of a tumour, the immune system is triggered to eradicate it. Tumour cells that evade destruction by the immune system can in fact exploit factors such as growth factors and extracellular matrix-modifying enzymes that were triggered as an immune response, for its own purposes and enhance tumour development. Finally, tumour development demands energy, therefore neoplastic cells carry not only deregulated cell proliferation but also deregulated energy metabolism.

Breast Cancer – Occurrence, Risk Factors and Treatment

Breast cancer is the most common cancer in women in Sweden, and the incidence has been steadily increasing in recent decades [4]. Approximately one in ten Swedish women will develop breast cancer before the age of 75 years, and around 7,000 women are diagnosed with breast cancer annually [4]. Breast cancer is more common in the Western world than in developing countries, most likely due to life-style factors but also due to screening programs that discover tumours, which possibly would have remained asymptomatic. Interestingly, people moving from a low-risk to a high-risk area usually increase their risk of developing breast cancer within one or two generations, indicating that environmental factors and adaptations to new life-styles affect the breast cancer risk [5]. Breast tumour development encompasses an intricate interplay between factors such as family history of the disease, life-style, increasing age, race, alcohol, smoking, oral contraceptives, hormone replacement therapy and reproductive factors such as early age at menarche, late age at fist full-term pregnancy and late menopause.

Today many patients are cured by surgery and adjuvant treatment in combination, however, a large fraction could potentially be cured by surgery alone but receive aggressive treatments with both harsh side effects and high costs. It is therefore essential to improve the methods of identifying the patients with low risk of recurrence and who currently are overtreated. Characteristics that determine the aggressiveness of a breast tumour are called prognostic factors and include age, TNM status, histological grade and expression of steroid receptors and HER2. Young patients usually have a more aggressive disease than older patients. Moreover, the proliferative capacity of the

tumour is sometimes measured by expression of Ki67. The TNM system classifies breast tumours based on tumour size (T), lymph node metastasis (N) and distant metastasis (M) into stage I-IV. Histological grade is defined by the combined evaluation of tubule formation, degree of nuclear pleomorphism, and mitotic count. Expression of oestrogen (ER) and progesterone (PR) steroid receptors is found in approximately two-thirds of all breast tumours and is associated with better prognosis, while around 15% show overexpression of HER2 [6], which is associated with a worse prognosis. ER, PR and HER2 are in addition to being prognostic factors also predictive factors as ER/PR predicts response to endocrine therapy and HER2 to HER2-targeted therapy.

The prognosis for breast cancer is highly dependent on disease stage at diagnosis and is generally fairly good. The 5-year survival rate is nearly 90% but still approximately 1,500 women die of their disease in Sweden every year [4]. The standard treatment includes surgery generally in combination with radiotherapy and trastuzumab, endocrine-, or chemotherapy. Small tumours are often removed by breast conserving surgery while large tumours as well as tumours spread to axillary lymph nodes demand mastectomy and removal of lymph nodes. After surgery, nearly all patients receive radiation and adjuvant systemic treatment *i.e.* endocrine or targeted therapy and chemotherapy. In some cases the patient requires neo-adjuvant therapy to shrink the size of the tumour before surgery.

The choice of neo-adjuvant and adjuvant systemic therapy is determined by hormonal receptor status, stage of the disease, HER2-expression and age of the patient. ER-positive patients generally receive endocrine treatment, *e.g.* tamoxifen, aromatase inhibitors or ovarian suppression. Tamoxifen acts by binding to the ER and thereby preventing the receptor from being activated by oestrogens, aromatase inhibitors disrupt the formation of oestrogen, thereby enhancing the inhibition of the ER pathway, and ovarian suppression hamper the production of oestrogen. Patients with HER2-amplified tumours can be treated with the monoclonal antibody trastuzumab (Herceptin®), which binds to the extracellular domain of the receptor and inhibits normal downstream signalling leading to *e.g.* proliferation as well as triggering of immune response. For patients with triple negative tumours, *i.e.* ER-, PR- and HER2-negative tumours, only chemotherapy is typically used as adjuvant systemic therapy.

In search for better prognostic and predictive tools many large gene expression studies have been conducted. Global gene expression analyses of breast tumours have revealed that breast tumours display gene expression patterns that can be connected

to for example prognosis [7-9], histologic grade [10], and familial status [11]. Some of the prognostic gene signatures have been developed into commercially available tests *e.g.* MammaPrint® (Agendia) and Onco*type* DX® (Genomic Health Inc.). Both are currently being evaluated in the Microarray In Node-negative and 1 to 3 positive lymph-node Disease may Avoid ChemoTherapy (MINDACT) [12], and the Trial Assigning IndividuaLized Options for Treatment (Rx) (TAILORx) [13] clinical trials, respectively. However, the performance of prognostic gene expression signatures have been somewhat criticised for not outperforming traditional clinical parameters or not performing well within all breast cancer subgroups [14, 15].

Breast Cancer Subtypes

In the early 2000's the microarray technology gained importance and the publications based on microarray results exploded. In 2000, Perou *et al*. published a paper in which they had identified four intrinsic molecular subtypes of breast tumours based on gene expression microarray data [16], and in 2001 [17] and 2003 [18] the same group further elaborated on the subtypes. Since then, other research groups have been able to validate and refine the subtypes and today at least five subtypes are defined: basal-like, HER2-enriched, luminal A, luminal B, and normal-like tumours [19, 20].

The molecular subtypes are believed to partly reflect the cell-type from which the tumour originates and to follow different tumour progression pathways. Indeed, it has been shown that the subtypes are associated with different prognosis, steroid receptor status, proliferation rates, hereditary backgrounds, CNAs, and sites for metastases [17-19, 21-24]. The majority of basal-like tumours are triple negative, *i.e.* no expression of ER, PR or HER2, express cytokeratins 5, 6 and 14, and are of a more stem-cell like phenotype than other subtypes [18, 25-27]. Additionally, the majority of *BRCA1*-mutated tumours belong to the basal-like subtype [18, 28-30]. HER2-enriched tumours can be ER-positive but the majority are ER-negative while luminal A and B tumours nearly always are ER-positive [16-20]. The major difference between luminal A and B tumours resides in their proliferative capacity with high proliferation in luminal B tumours and a shorter survival for patients with tumours of this type compared to patients with luminal A tumours [19]. Furthermore, *BRCA2*-mutated tumours are frequently classified as luminal B [28-30]. The subtypes with worst outcome are basal-like, HER2-enriched and luminal B. The normal-like subtype is the least well-defined subtype, and is named normal-like because of the expression of genes in common with

normal breast tissue specimens [16]. It is unclear whether the normal-like tumours represent a true breast cancer subtype or whether they merely reflect the degree of normal cells in the analysed sample.

Efforts have been made by several research groups to also on DNA-level characterise and subtype breast tumours [22, 28, 31-36]. Jönsson *et al.* used 32k BAC arrays to classify 359 breast tumours into basal-complex, 17q12, luminal-simple, luminal-complex, amplifier and mixed subtypes [28]. Four of the genomic subtypes displayed high association to the gene expression subtypes as decided by Hu *et al.* [19]. The majority of the basal-like tumours were classified as basal-complex and almost all HER2-enriched as 17q12. The luminal-simple subtype comprised almost only luminal A tumours while the majority of the luminal-complex tumours were classified as luminal B by gene expression [28]. Each subtype is characterised by their own set of amplified or deleted genes as well as patterns and amplitude of the aberrations leading to the notion that breast cancer is a profoundly heterogeneous disease also on the genomic level.

Examples of Genetic Aberrations in Breast Tumours

Genes involved in cancer development can either be categorised as oncogenes or tumour suppressor genes. While oncogenes act in favour of cell proliferation and frequently are abrogated by *e.g.* gain-of-function mutations or amplifications, tumour suppressor genes act in the opposite direction. Tumour suppressor genes can be further subdivided into gatekeepers and caretakers [37]. Gatekeepers regulate growth and caretakers maintain genomic stability. By *e.g.* deletions or epigenetic silencing mechanisms, frequently occurring in tumours, tumour suppressor genes can no longer control or pause the cell cycle, or steer malfunctioning cells into apoptosis leading to tumour development.

The *TP53* tumour suppressor gene encodes the p53 protein, which is a multifunctional transcription factor involved in many processes in response to cellular stress signals, and is also called "the guardian of the genome" [38]. Inactivating mutations of *TP53* are often associated with high risk of tumour development and occur in the majority of sporadic cancer types [39]. Approximately 20-40% of all breast cancers carry mutations in the gene [39]. In addition to inactivating mutations, p53 activity is impeded in many tumours by *e.g. TP53* binding proteins, transcription factors or by changes in upstream or downstream targets in the pathway [39].

Approximately 5-10% of all breast cancer cases occur in patients with a family history of the disease [40, 41]. Of these, mutations in one of the two major breast cancer susceptibility genes *BRCA1* or *BRCA2* [42, 43] account for approximately 50%. In Sweden and many other countries, especially in the Western world, *BRCA1* and *BRCA2* are screened for mutations in families with a history of breast and ovarian cancers. A mutation in either of the two genes generates a life-time risk of developing breast cancer of almost 85% [44-47]. *BRCA1* is located at chromosome 17q21 and has several roles such as involvement in DNA repair, cell-cycle checkpoint control, protein ubiquitylation and chromatin remodelling, which, if abrogated, can lead to cancer [44]. *BRCA2*, located at chromosome 13q12, is involved in homologous recombination, however, its full functions remain to be fully characterized [44]. Both genes function in a common pathway and are responsible for the integrity of the genome and maintenance of chromosomal instability rendering them important caretaker genes [44]. In hereditary tumours one allele is already non-functional due to germline mutations while the second allele generally must be lost through *e.g.* somatic deletion or epigenetic silencing for a tumour to develop [48].

HER2 is a transmembrane receptor tyrosine kinase (RTK) located at chromosome 17q12 that mediates cell growth, differentiation and survival and that is amplified in 15% of all breast tumours [6]. Overexpression of HER2 is a marker of poor prognosis, but is also a treatment-predictive factor since HER2-positive breast tumours can be treated using targeted therapy towards HER2 (trastuzumab).

CCND1, situated at chromosome 11q13.3, encodes the key cell cycle regulatory protein cyclin D1, which controls the transition from G_1 to S during the cell cycle [49]. Different cyclins regulate the cell cycle in collaboration with cyclin dependant kinases at specific points of the mitotic cycle (Figure 3). Cyclin D1 functions as a regulatory subunit of a complex with the two cyclin-dependant kinases CDK4 and CDK6. During cell cycle arrest the cell cycle inhibitors p15, p16, p18 and p19 enter the cell nucleus and bind to and inactivate cyclin D1-CDK4/6-complex [50]. Upon mitogenic stimulation cyclin D1 is released and binds to CDK4/6, which then phosphorylates the retinoblastoma protein (Rb) leading to release of E2F transcription factors [51]. E2F transcription factors, in turn, promote transcription of genes involved in cell cycle progression whereby the cell can enter S-phase (Figure 3; [52], reviewed in [53]).

CCND1 is amplified in 15-20% of breast tumours but overexpressed in up to 50% of all breast tumours, most often in ER-positive tumours [54-56]. Tumours

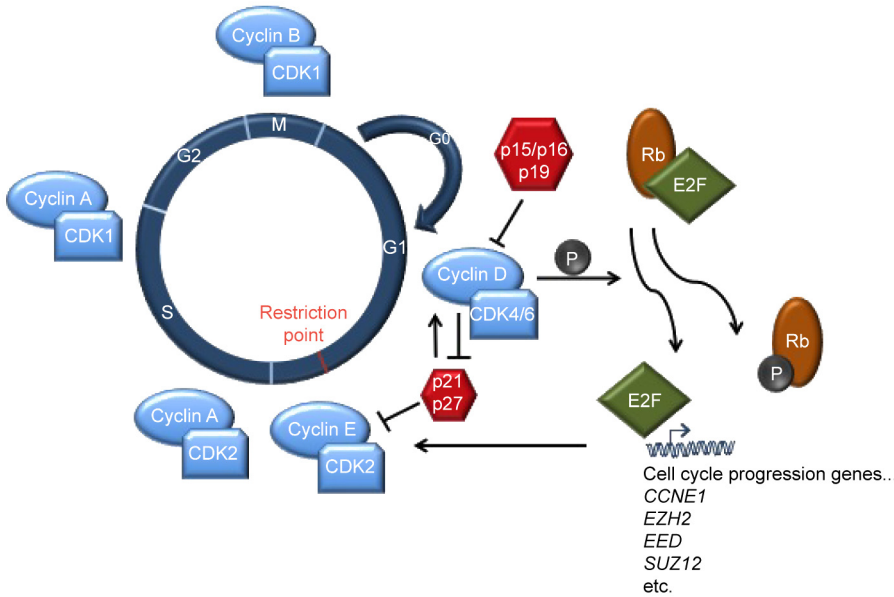

Figure 3. Cell cycle. Schematic image of the cell cycle focusing on cyclin D-related actions. The cyclins vary through the cell cycle and regulation of the G_1/S transition is performed by cyclin D1. Mitogenic signals promote assembly of cyclin D1 and the cyclin dependent kinases CDK4/6, which in turn phosphorylates Rb leading to the release of E2F. The E2F transcription factors can then promote transcription of genes involved in cell cycle progression as well as genes coding for the PRC2 components.

with *CCND1* amplification usually overexpress the gene, however in tumours lacking amplification other mechanisms contribute to the overexpression. It has been shown that the overexpression of cyclin D1 shortens the G_1-phase resulting in an increase of the number of cells progressing through G_1 into S-phase [57]. In addition to the cell cycle regulatory function of cyclin D1, it also has a CDK-independent function and may activate ER-mediated transcription in absence of both CDK and oestrogen and may enhance transcription in the presence of oestrogen, which may underlie its oncogenic role in ER-positive breast tumours [58].

PIK3CA encodes the phosphoinositide-3-kinase p110α catalytic subunit of the phosphatidylinositol 3-kinase (PI3K) holoenzyme (pathway discussed below). The gene has been shown to carry gain-of-function mutations as well as be amplified (located at chromosome 3q26) in several different cancer types [59-63]. More specifically, approximately 25% of all breast tumours have a somatic mutation in the gene [60, 61, 64]. Mutations, which have been shown to be oncogenic both *in vitro* and *in vivo,* lead to decreased apoptosis, increased tumour invasion, growth-factor independent

proliferation and increased angiogenesis [59, 65-67]. PI3Ks are heteromeric enzymes consisting of a catalytic and a regulatory subunit of which the catalytic enzymes can be subdivided into three classes, each class with its own protein structure, substrate specificity, tissue distribution and mechanism of activation (reviewed in [68]). *PIK3CA* belongs to class IA and encodes the catalytic subunit p110α, which binds to the regulatory subunit p85.

PTEN, a dual specificity phosphatase, is another important member of the PI3K pathway in which it has a role as a break and tumour suppressor. The gene is located at chromosome 10q23.31 and is deleted, methylated or carry loss-of-function mutations in several cancer types making *PTEN* the second most aberrated tumour suppressor gene after *TP53* [69, 70]. In breast cancers, approximately 5% of all tumours carry a mutation, however, the protein is lost in 20-48% of all breast cancers through other mechanisms such as epigenetic silencing or gross mutations and is associated with poor patient survival [71-75].

The PI3K Pathway

The PI3K pathway has a central role in tumour development in many cancer forms with deviations at several different levels of the pathway, resulting in it being one of the most commonly altered pathways in human cancer. PI3K pathway aberrations are estimated to account for up to 30% of all human cancers [76]. The pathway affects important tumorigenic hallmarks such as cell survival, proliferation, migration, metabolism, angiogenesis, and apoptosis (pathway reviewed in [70, 77, 78]). Both overexpression of RTKs, activation of oncogenes and inactivation of tumour suppressor genes are found in this pathway. More specifically, in breast tumours, the PI3K pathway is commonly activated by gain-of-function mutations in *PIK3CA*, loss of PTEN or overexpression of HER2 or EGFR.

Upon stimulation by a growth factor, RTKs are autophosphorylated, providing docking sites for the PI3K complex, which assembles at the cell membrane (Figure 4). The PI3K complex is then phosphorylated either directly by the RTK or via adaptor proteins such as IRS1 or IRS2. After activation, the p110α subunit can catalyse the phosphorylation of phosphatidylinositol-4,5-bisphosphate (PIP2) forming the lipid second messenger phosphatidylinositol-3,4,5-trisphosphate (PIP3). Alternatively, binding of RAS, activated through an RTK via adaptor proteins SHC, GRB2 and GAB2,

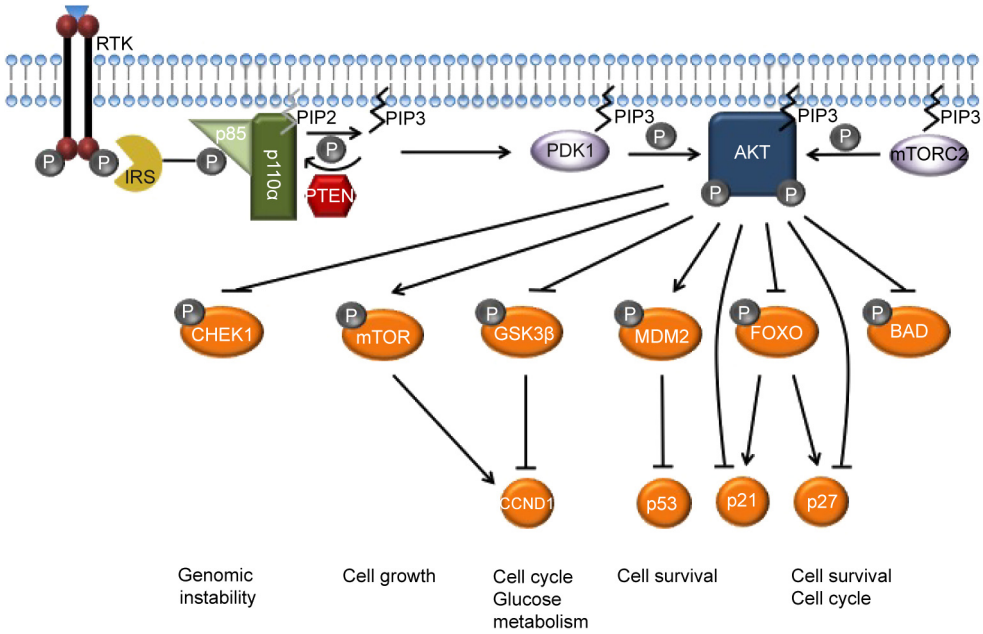

Figure 4. PI3K pathway. Overview of selected parts of the PI3K pathway. The PI3K complex activates the pathway as p110α phosphorylates PIP2 forming PIP3. PTEN acts as a break and tumour suppressor by dephosphorylating PIP3 back to PIP2. Activation of the pathway results in several processes that can be recognised as hallmarks of cancer. For details see text.

to the p110α subunit can lead to PIP3 formation by p110α without phosphorylation of the PI3K complex. PTEN acts as a break on this process by dephosphorylating PIP3 back to PIP2 (Figure 4).

After formation of PIP3 and subsequent phosphorylation of AKT via the adaptors PDK1 and mTORC2, an array of possibilities opens up. Through phosphorylation, AKT mediates both activation and inhibition of various target genes resulting in increased cellular growth, proliferation, glucose metabolism, cell motility and decreased apoptosis. For example, AKT regulates p27 through inhibition of FOXO and inhibition of TSC2-mediated stabilisation of p27 leaving CDK2 active, which increases proliferation (Figure 3) [79, 80]. In addition, inactivation of FOXO also results in downregulation of proapoptotic genes. Moreover, AKT inhibits GSK3β, which prevents export of cyclin D1 from the nucleus during S-phase and subsequent degradation resulting in continued proliferation [81]. By phosphorylation of MDM2, AKT promotes translocation of MDM2 into the nucleus where MDM2 targets p53

for degradation by exporting p53 into the cytoplasm where p53 becomes ubiquitinated and degraded, resulting in cell survival [82]. Through inactivation of BAD, which normally inactivates prosurvival factors such as Bcl-XL, cell survival is maintained and apoptosis avoided. Downstream activation of mTOR via downregulation of mTOR suppressors such as TSC1 and TSC2 leads to increased translation of cyclin D1, D3 and E, as well as MYC and results in an increment of pro-proliferation factors [83]. In addition, derangement of the PI3K pathway leads to genomic instability by aberration of CHEK1, which is a DNA damage-response cell cycle regulator [84]. Finally, the phosphorylation of GSK3β and FOXO results in increased uptake and production of glucose [85-87]. In summary, enhanced phosphorylation and activity of AKT results in several processes that can be recognised as hallmarks of cancer.

Breast Cancer Epigenetics

Epigenetics, literally "above genetics", is the study of changes to the phenotype that are caused by other mechanisms than those that alter the DNA sequence [88, 89]. These changes are heritable during cell division. Two modes of epigenetic regulation are DNA methylation and histone modification. While DNA methylation leads to stable and long-term repression of gene expression, histone modifications are of a more dynamic nature and can be changed upon stimulation. Furthermore, histone modifications can both increase or silence gene transcription.

DNA Methylation

During DNA methylation, a methyl group (CH_3) is added to the 5' position of the cytosine ring in a CpG dinucleotide by DNA methyltransferases (DNMTs). DNMT1 maintains previous DNA methylation patterns after DNA replication while DNMT3a and DNMT3b are responsible for *de novo* methylation, which is stimulated by DNMT3L [90-93].

CpGs occur less frequently in the genome than expected by overall GC content, which could be explained by the propensity of methylated cytosines to mutate to thymine [94]. However, CpGs are often found in the promoter regions of genes at higher frequencies than generally found throughout the genome forming CpG islands, or in regions of large repetitive sequences (*e.g.* centromeres and retrotransposon

elements) [95]. Different definitions of CpG islands exist of which one of the more stringent have been described by Takai *et al.* [96]. They define CpG island as a region with GC content of more than 55%, and an observed versus expected ratio of more than 0.65, and a minimum size of 500 bp [96]. Using this definition, about 70% of CpG islands are associated with human genes and more than half are within promoter regions [97].

When a CpG island becomes methylated the DNA becomes inaccessible and the promoter regions can no longer be recognised by the transcriptional machinery, resulting in gene silencing (Figure 5a). In normal cells, CpG islands generally are hypomethylated while the genome at large, including repetitive sequences, is methylated to prevent chromosome instability [98, 99]. In cancer cells, on the other hand, hypermethylation of CpG islands is frequent and accompanied by global genomic hypomethylation leading to silenced tumour suppressor genes and instable genomes [98].

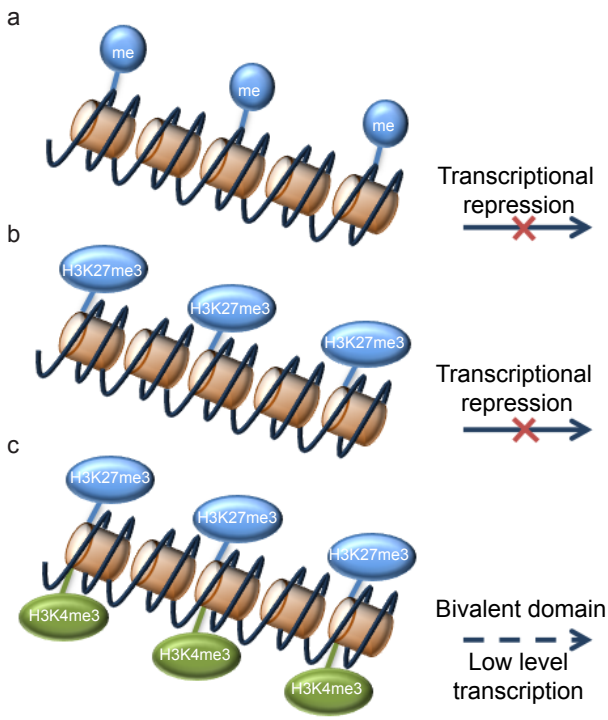

Figure 5. Examples of epigenetic modifications. a Transcriptional repression by DNA methylation. **b** Transcriptional repression by the histone modification H3K27me3. **c** Bivalent domain with both the activating H3K4me3 mark and the silencing H3K27me3 mark leading to low level transcription.

Histone Modifications

One hundred and forty seven base pairs of the DNA helix is wrapped around core histones (H2A, H2B, H3 and H4) as well as two linker histones (H1 and H5). Two of each of the core histones form an octamer, which is held together by H1, packing and organising the DNA into nucleosomes, which are the fundamental structural units of chromatin. In general, modifications of histones regulate gene expression by either altering the chromatin structure (loosening it for the transcriptional machinery to gain access to the DNA being transcribed), or by recruiting other regulatory proteins whereby transcription is either enhanced or prevented. Several mechanisms can modify the histones including methylation, acetylation, ribosylation, ubiquitination, SUMOylation, and phosphorylation [94]. Only histone modifications in the form of methylation have been studied in this thesis, and therefore the other mechanisms will not be discussed further.

Polycomb Repressive Complex and EZH2

Polycomb group (PcG) proteins can control chromatin compaction and maintain gene expression patterns of different cells that are determined before differentiation [100]. The two Polycomb repressive complexes 1 and 2 (PRC1 and PRC2, respectively) both contribute to chromatin compaction, and seem to be required to maintain gene repression [100]. However, only PRC2 has been studied in this thesis. PRC2 is responsible for the histone mark characteristic for PcG-mediated silencing: trimethylation of lysine 27 on histone 3 (H3K27me3).

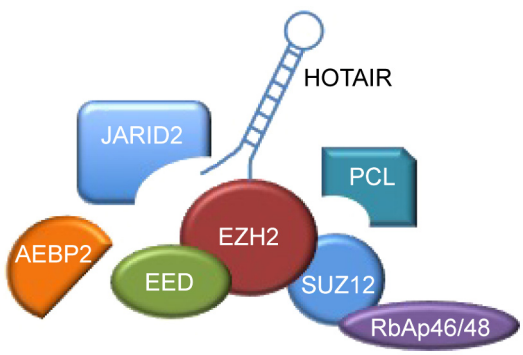

Figure 6. PRC2. The PRC2 complex includes EZH2, EED, SUZ12 and RbAp46/46. Other components such as AEBP2, JARID2, HOTAIR and PCL may akt as guides to a specific target sequence.

The PRC2 complex is involved in many processes such as differentiation, maintenance of cell identity and proliferation, and binds to response elements of target genes in a cell-type specific manner although with a preference for genes involved in developmental regulation such as the *HOX* genes [100-103]. The complex is built up by at least four components: EZH2, EED, SUZ12 (the transcription of which are regulated by E2F transcription factors; Figure 3) and RbAp46/48 (Figure 6) [100]. It has recently been shown that the complex contains additional peptides such as AEBP2, JARID2 and PCLs, which might aid in the recognition and binding of target sites, and affect the enzymatic activity of the complex [100, 104-108]. Additionally, long noncoding RNAs (lncRNA) such as HOTAIR, may serve as scaffolds and guide the complex to specific target genes (Figure 6) [100, 109].

Enhancer of zeste homologue 2 (EZH2) is the core member of PRC2 that catalyses the formation of H3K27me3. H3K27me3 is a histone mark that causes blocking of transcription factors and consequently gene silencing independent of promoter methylation (Figure 5b) [110]. Overexpression of EZH2 has been shown in several cancer types such as breast, melanoma, bladder, and prostate cancer [111-113], and has been found to correlate with aggressive phenotypes of breast and other cancer types [114, 115]. The expression of EZH2 in normal cells is cell cycle-regulated by pRb/E2F and the protein accumulates at the G_1/S transition (Figure 3) [116, 117]. However, in cancer cells this regulation is lost and overexpression results in induction of S-phase [116, 118].

Transcriptional Regulation under Epigenetic Control

Developmental genes are marked and silenced in embryonic stem (ES) cells by H3K27me3 in order to maintain ES cells in an undifferentiated state [103]. In addition, these genes are often marked with trimethylation of lysine 4 on histone 3 (H3K4me3), which is a mark for active transcription, forming bivalent domains (Figure 5c) [119]. The bivalent markings are characteristic of pluripotent cells and enable the cells to quickly respond to initiation of specific differentiation programs, reflecting the plasticity of pluripotent cells [119]. During differentiation, markings by large regions of either H3K4me3 or H3K27me3 decide and maintain lineage specific gene expression or repression [119]. Hawkins *et al.* compared the epigenomic landscapes of human embryonic stem cells (hESCs) to fetal lung fibroblasts (IMR90) and found expanded domains of H3K27me3 in the more differentiated IMR90 cells compared to the hESCs [120]. Moreover, the expansion seemed to maintain more stable gene silencing

in the differentiated cells [120]. Interestingly, they also found that the promoters of many genes are marked by combinations of repressive modifications such as DNA methylation in combination with H3K27me3; trimethylation of lysine 9 on histone 3 (H3K9me3) in combination with H3K27me3, or DNA methylation in combination with H3K9me3 [120]. Other groups have also shown that after differentiation, genes previously marked with H3K27me3 gain *de novo* DNA methylation [108, 121]. In agreement, it has been shown that genes with a low expression in normal cells undergo *de novo* DNA methylation in cancer cells by recruitment of DNMTs [122-124]. This suggests that a large fraction of *de novo* methylation in cancer cells could reflect an instructive mechanism from where the tumour originate instead of being the result of growth selection.

In tumours, the epigenetic regulators that control DNA methylation and histone modifications can be altered leading to irregular expression of tumour suppressor genes and oncogenes as well as other genes involved in maintaining cellular identity. Such alterations have been found for a number of epigenetic regulators [125-130].

Materials & Methods

Materials and methods are described in detail in the papers of the thesis, what follows is a summary and principles of methods:

Patients, Tumours and Cell Lines

Paper I

Two large tumour sets including almost 300 tumours as well as 51 breast cancer cell lines were included in Paper I. DNA from sporadic primary breast tumours or tumour cell pellets (n=162) together with normal tissues from a subset of the same patients were obtained from the South Swedish Breast Cancer Group collected at Lund University Hospital, Lund. DNA from breast tumour biopsies (n=133) was obtained along with normal tissue for a subset of the same patients from the Herbert Irving Comprehensive Cancer Center Tumor Bank at Columbia University, NY. The combined cohort contained tumours of all stages, a mixture of sporadic and hereditary cases, and patients had gone through different clinical therapies.

Paper II

DNA from 94 breast tumours was used for array comparative genomic hybridisation (aCGH). The tumours were obtained from the South Swedish Breast Cancer Group tissue bank at the Department of Oncology (n=55), Lund University Hospital, and from the Landspitali University Hospital, Reykjavik, Iceland (n=39). The tumours had been subtype-classified on both mRNA level according to Hu *et al*. and DNA level as previously described [19, 25, 28]. For FISH, we did imprints of ten frozen breast tumours selected from the Swedish subset of tumours.

Paper III

DNA from 189 fresh frozen primary breast tumours and four normal breast tissue samples were analysed for methylation status. The tumours were obtained from the South Swedish Breast Cancer Group tissue bank at the Department of Oncology, Lund University Hospital. The majority of the tumours had been subtype-classified according to Hu *et al*. [19] as described in [28].

Paper IV

Tissue microarrays (TMAs) with breast tumour tissue from more than 400 patients were used to analyse the expression of EZH2 and occurrence of H3K27me3 in relation to breast tumour subtypes and clinicopathologic variables. The TMAs originate from two well-characterised cohorts with long-term follow-up [131, 132]. The tumours were classified using ER/PR, HER2, and Ki67 as described in Paper IV. A subset of the two cohorts had earlier been classified into molecular subtypes according to Hu *et al*. as described [25, 28]. Additionally, the majority of the tumours in Paper III were also included in Paper IV, and were screened for a recurrent mutation of *EZH2*.

Sequencing

In Paper I we performed mutation screening of *PIK3CA* and *PTEN*, and in Paper IV of *EZH2*. All samples analysed for *PTEN*, *EZH2* and parts of the Swedish samples analysed for *PIK3CA* were sequenced using BigDye Terminator Cycle Sequencing (Applied Biosystems, Carlsbad, CA) and an ABI PRISM 3100 or 3130xl Genetic Analyzer (Applied Biosystems). The remaining samples sequenced for *PIK3CA* were sequenced by Agencourt Bioscience Corp. (Beverly, PA).

The sequencing reaction was performed on purified PCR products and involved incorporation of nucleotides (dNTPs) and fluorescently labelled dideoxynucleotides (ddNTPs). Incorporation of ddNTPs terminates the sequence, resulting in a pool of fragments with varied size and fluorescent labelling. The fragments were then separated according to size by capillary gel electrophoresis and detected using a fluorescence detecting system.

Sequence chromatograms were analysed using Sequencher (Gene Codes, Ann Arbor, MI) or Mutation Surveyor (Softgenetics, State College, PA). Mutations were verified by re-amplification and bi-directional sequencing. For the *PIK3CA* novel mutations, DNA from the corresponding normal patient tissue, when available, was also sequenced to confirm the mutation as a somatic change, and for a subset of tumours with mutations previously described, the corresponding normal DNA was also sequenced.

Immunohistochemistry

Immunohistochemistry (IHC) is a well-established and widely used method for analysis of protein expression in tissues. The principle is that an antibody specific to an epitope on a protein of interest is added to a microscope slide carrying the tissue that will be investigated. A secondary antibody conjugated with peroxidase is then added followed by diaminobenzidine (DAB), which produces a brown colour when oxidized by peroxidase. The nuclei are counterstained with haematoxylin, which gives a blue colour. This allows for analysis of both location of expression (*i.e.* nuclear/cytosol/membrane) as well as intensity and frequency of expression of the investigated protein. The development of TMAs allows for high-throughput analysis as several hundred representative tissue cores can be positioned on one slide.

IHC has been used in Paper I for analysis of PTEN (Neomarkers, Fremont, CA) and HER2 (ERBB2, DakoCytomation, Glostrup, Denmark), and in Paper IV for EZH2 (BD Transduction Laboratories, Franklin Lakes, NJ) and H3K27me3 (Abcam, Cambridge, MA).

Microarrays

Principle of Method

The advent of the microarray technology has revolutionised the amount of data that can be extracted from a sample in a single experiment. This has led to new challenges in data analysis, which will be discussed later. Using microarrays, we have analysed both transcription levels and mRNA abundance as well as copy number status of thousands

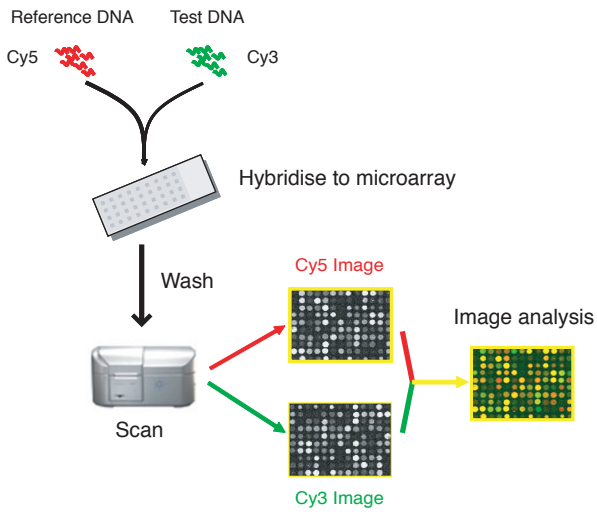

Figure 7. Schematic overview of the experimental procedure using microarrays. In two-colour microarrays a test and a reference sample labelled with different colours competitvely hybridise to an array with bound probes. The probes vary dependent on what is being analysed *e.g.* gene expression, copy number or methylation status. After washing, the slides are scanned with lasers of different wavelengths resulting in two images, which are subsequently merged, analysed and interpreted for each probe as *e.g.* over- or less expressed, amplified or deleted, and methylated or unmethylated in the analysed sample.

of genes. Global gene expression arrays and aCGH have been used in Paper II, and the results from gene expression arrays have been used in Papers III and IV.

Two-channel microarray technology is based on the competitive hybridisation of differentially labelled test and reference samples (Figure 7). Our samples were fluorescently labelled with Cy3 or Cy5. The test samples might originate from mRNA or DNA depending on whether gene expression or copy number is being measured. The pooled test and reference samples are then hybridised to thousands of probes precisely positioned on a microarray. For aCGH, we used a 32k BAC-clone set, (Gene expression omnibus (GEO) platform GPL4723 [133]) obtained from the BACPAC Resource Center at Children's Hospital, Oakland Research Institute, Oakland (CA, US), as probes. For gene expression we used 55k-arrays with oligonucleotides from the Human Genome Oligo Set version 2.1 and the version 2.1 upgrade (GEO platform GPL5345) as probes. All arrays were produced at the SCIBLU Genomics Centre at Lund University, Sweden [134] as described [135].

After removal of unbound or non-specifically bound fragments the microarray is scanned in a laser scanner registering the fluorescence of the two channels separately, one for the test sample (Cy3) and one for the reference sample (Cy5). The two images are then merged and intensity ratios are calculated (test/reference, Cy3/Cy5) for each probe (Figure 7). The ratios are then $\log_2$-transformed and positive values can be interpreted as more genetic material of a particular gene in the sample compared to the reference, and vice versa for negative values.

Data Analysis

The initial analysis of both gene expression and aCGH data includes spot quality control, background correction, and normalisation. During spot quality control spots with low signal-to-noise ratio, spots that have been marked as bad during image analysis and saturated spots are removed from further analysis. Background correction removes background fluorescence from true measurements, while normalisation adjusts for technical bias both within and between arrays. Normalisation accounts for the fact that the measured intensities not necessarily are proportional to the amount of starting material (mRNA/DNA) in the experiment since factors such as print-batches, different labelling efficiency or dye properties might affect intensities in a biased manner. Therefore normalisation is an essential step to take to be able to compare samples.

CGH-Plotter was used for segmentation of aCGH data (Paper II) [136], and was performed in BioArray Software Environment (BASE) [137]. During segmentation the $\log_2$-ratios for each BAC are compared with $\log_2$-ratios of neighbouring BACs to identify genomic breakpoints. We used sample adaptive $\log_2$-ratios as thresholds, derived from 200 kbp smoothed data, which take the noise in each sample into consideration when setting thresholds for gain and loss (described in [138]).

Hierarchical clustering, K-means clustering and Significance analysis of microarrays (SAM) [139] were used in Paper III. Hierarchical clustering was used to visualise similarities between samples. K-means clustering was used to investigate the robustness of the result from the hierarchical clustering varying the number of clusters from two to five. SAM analysis was performed as a two-class analysis (one subtype vs. all other samples) to identify significant CpGs for each subtype. Both clustering methods and SAM were performed in MultiExperiment Viewer (MEV) [140].

Fluorescence *in situ* Hybridisation (FISH)

FISH was performed in Paper II to validate results from aCGH and to investigate the potential occurrence of intratumour heterogeneous clonality. In principle, fluorescently labelled BAC clones are hybridised to known genomic sequences of chromosomes in either interphase or metaphase state. Using a microscope with a UV-lamp and appropriate filters, the fluorescence from the BAC-clones can be registered. Each BAC-clone generates a spot that can be registered and the number of copies can be calculated.

Methylation Arrays

In Paper III, we used Illumina GoldenGate Methylation Cancer Panel I (Illumina, San Diego, CA) to analyse the methylation status of more than 800 cancer-related genes. Briefly, genomic DNA is bisulfite-treated, during which unmethylated cytosines are converted to uracil while methylated cytosines stay unchanged. A cocktail of PCR-primers matching either the methylated unchanged sequence, or the unmethylated and therefore changed sequence, is then added to each sample together with universal primers tagged with an address sequence. After extension and ligation the templates are PCR-amplified with fluorescently labelled primers, one dye matching the methylated state and one the unmethylated state. The resulting products are then hybridised to a bead array with complementary address sequences unique to each interrogated CpG site. The fluorescence from each bead is registered in a laser scanner and methylation status is calculated as the ratio between fluorescence from one allele and the sum of fluorescence of both alleles and presented as β-values. The β-values range between 0 and 1, with 0 corresponding to completely unmethylated CpGs while 1 corresponds to completely methylated CpGs.

Statistical analyses

For survival analyses we used log-rank tests, and the data was visualised by Kaplan-Meier curves. The analyses were performed in R using the survival package [141]. Pearson or Spearman correlation was used for correlation analyses. For analysis of significant differences between sample groups we used t-test, Wilcoxon test, binomial

test, χ^2-test or Fisher's exact test for two-category comparisons, and analysis of variance (ANOVA) when comparing multiple categories. Data was visualised by boxplots. All statistical analyses were performed in R except for χ^2-tests [141]. P-values were adjusted for multiple comparisons using Bonferroni correction.

Experimental Considerations

When working with breast tumour samples it is important to bear in mind that the samples comprise a mixture of cells and not only tumour cells. Normal breast epithelial and stromal cells will most likely also be present in the sample. This is important to remember when working with *e.g.* gene expression, copy number and methylation data. Gene expression results will not only show genes over- or less expressed in breast tumour cells but the expression profile will also be affected by the presence of cells from surrounding tissues, which, in reality, also might be of interest since the behaviour of the surrounding stroma might affect the growth of the tumour. For copy number data the genomic profiles of the tumours will be "diluted" by the unchanged profiles from normal cells. That is the reason to why a hemizygous deletion rarely reaches a $\log_2$-ratio of -1. The situation is the same for methylation data, in our case presented as β-values ranging between 0 and 1. These values cannot be interpreted as absolute but should be used in comparison with other β-values. Moreover, the values will be "diluted" by the methylation states of the genes in normal cells.

Results & Discussion

PI3K Pathway Alterations (Paper I)

We performed mutational analysis in 292 primary breast tumours and 50 cell lines on all of the ten exons of *PIK3CA* that previously had been shown to carry mutations [64]. In addition, we screened all nine exons in *PTEN* for mutations in the Swedish tumours as well as analysed protein expression by IHC.

PIK3CA Mutations in Breast Tumours

We found a *PIK3CA* mutation frequency of 26% in the primary tumours and 28% in the cell lines; similar frequencies have been found in other reports [60, 61, 64, 142-144]. The mutations were clustered in the helical (exon 9) and kinase domains (exon 20) with H1047R, E545K and E542K being the most frequent mutations. In breast tumours, mutations in exon 20 predominate the mutation spectrum, while in colorectal cancers mutations in exon 9 are more frequent [59]. In addition, we reported mutations in exon 7 (C420R, C2 domain) as a third hotspot, which later has been reproduced [145]. In two tumours and one cell line, we found double mutations in *PIK3CA* (exon 20 in combination with either exon 7 or 9), indicating that some mutations are less potent, a second hit in the alternate allele, or a multiclonal tumour.

Functional Effect of Site of Mutation

Several functional studies have shown different potency for mutations in the non-kinase and kinase domains. In *PIK3CA*-mutated breast cancer cell lines, P539R, E542K, E545K, and H1047R have been shown to harbour strong oncogenic properties as they increase kinase activity, promote cell invasion, show abilities to growth factor independent as well as anchorage independent growth, and change the morphology of the cells [59, 146]. However, helical and kinase domain mutations seem to trigger

the PI3K pathway trough different mechanisms. Gain-of-function via helical domain mutations is independent of binding to p85 but requires interaction with RAS-GTP. Contrary, kinase domain mutations are highly dependent on p85 and do not require RAS-GTP binding [147]. The different potency of kinase and helical domain mutations was made visible in another paper by Saal *et al.* in which a PTEN-loss gene expression signature was constructed [74]. Hierarchical clustering demonstrated that the majority of tumours with kinase domain mutations in *PIK3CA* clustered with other samples that had a loss of PTEN as measured by IHC, while tumours with mutations in other domains clustered of *PIK3CA* clustered with tumours with retained PTEN [74].

PIK3CA and PTEN Aberrations

For PTEN, the majority of the tumours displayed positive staining and 5% carried a mutation, which is in line with earlier publications [148, 149]. Correlation analysis between PTEN protein expression and *PIK3CA* mutational status revealed that the majority of the *PIK3CA* mutations occurred in tumours with expression of PTEN (P=0.0066, χ^2-test). When we combined PTEN protein expression data with mutation data, the inverse correlation between loss of PTEN and *PIK3CA* mutation status was even more prominent (P=0.0037, χ^2-test). Interestingly, two tumours carried mutations in both *PIK3CA* and *PTEN*, however, neither of the *PIK3CA* mutations belong to the most potent mutations, and are potentially not robust enough to activate the PI3K pathway in presence of intact PTEN. As *PIK3CA* and PTEN catalyse the opposite reaction in the PI3K pathway, aberration of any of the proteins could lead to the same end result, namely increase of PIP3. Given this, it is not unexpected that the overlap between aberrations is low, since the growth advantage that is given by aberration of any of the two factors could be enough to enhance tumour cell growth.

Mutational Status, Clinicopathologic Factors, and Molecular Subtypes

Next, we wanted to investigate the relationship between *PIK3CA* mutational status and clinicopathologic factors. We found *PIK3CA* mutations to be significantly correlated to positive lymph node status, ER- and PR-positivity (P=0.0375, P=0.0001, and P=0.0063, respectively, χ^2-test). Other authors have shown confirmatory results [143, 144, 150], whereas some smaller studies have failed to find such correlations [60, 151].

Since Paper I was published, HER2 has been re-evaluated in the Columbia tumour set using chromogenic *in situ* hybridisation (CISH), which resulted in a significant proportion of tumours changing from positive to negative HER2 status (Saal *et al.*, manuscript in preparation). Given that we did not find any correlation between *PIK3CA* mutation and HER2-positivity in the Swedish set and the new results in the Columbia set, we conclude that there is no association between the two factors, which is corroborated by other authors [144, 145, 150].

Interestingly, it has been shown that patients with any mutation in *PIK3CA* [150] or specifically exon 20 mutation [152] have a better survival than patients with wild type *PIK3CA*. This could be explained by the fact that the majority of the *PIK3CA*-mutated tumours are ER-positive and patients with ER-positive tumours generally have a better outcome. However, we did not find any correlations between *PIK3CA* mutational status and survival, not even when stratifying for cohort, stage, lymph node or ER status. Consequently, it would be interesting to compare survival between patients with *PIK3CA*-mutated and non-mutated tumours stratified by molecular subtype, which would divide the ER-positive tumours essentially into luminal A and B. In a study by Li *et al.*, they found *PIK3CA* mutations to be associated with positive steroid receptor status, larger tumour size and poor survival [145]. When they performed survival analysis among ER-positive tumours, they found shorter survival among *PIK3CA*-mutated patients than for patients with wild type *PIK3CA* indicating the importance of identifying in which tumour subtypes a biomarker such as *PIK3CA* mutation status, and potentially type of mutation, might add important information.

The given observations of *PIK3CA* mutations in ER-positive tumours and PTEN deviations in ER-negative tumours are interesting and lead us to the conclusion that the PI3K pathway is activated by different mechanisms in different breast tumour subtypes. Unfortunately, the tumours included in Paper I have not been classified into the molecular subtypes. However, considering known relations between the molecular subtypes and clinicopathologic features such as ER/PR, HER2, and node status, it is tempting to speculate that *PIK3CA* mutations potentially affect tumours classified as luminal B or A (or both) whereas PTEN deviations have been shown to affect tumours classified as basal-like [72].

Our and others results are interesting therapeutically for the development of targeted therapies towards the PI3K pathway. Of all breast tumours, without stratification of *e.g.* subtype, around 15% carry an amplification of *HER2*, 5-6% have an amplification

of *EGFR*, 2-3% carry a mutation in *AKT1*, 25-30% have an alteration (mutation or amplification) of *PIK3CA*, and around 25-30% carry an alteration of PTEN. In addition, other components further downstream can also be altered to activate the pathway. Since some alterations co-exist, it is believed that the PI3K pathway is activated in more than 50% of all breast tumours. In the light of the frequent aberrations in this pathway it would be of benefit for many patients with effective therapies towards specific sites of the PI3K pathway, alternatively towards a common effector downstream of the altered proteins.

Copy Number Aberrations at 11q13 (Paper II)

We performed aCGH to characterise and fine map copy number aberrations at 11q13 in 94 *CCND1*-amplified (*CCND1*+) primary sporadic, hereditary and familial breast tumours, as well as analysed associations to clinicopathologic factors. We used a previously published set of 281 *CCND1*-non-amplified (*CCND1*-) breast tumours for comparisons [28].

CCND1 Status and Clinicopathologic Factors

In accordance with what has been shown before, we found that the majority of the *CCND1*+ samples were ER/PR-positive [54, 55, 153]. In addition, we found that luminal B was the most frequent subtype and luminal A came second. This is not unexpected given the association to ER status and proliferation for both *CCND1*+ samples and luminal subtypes, *i.e.* cyclin D1 is a proproliferation protein and tumours in the luminal B subtype are more proliferative than tumours in the luminal A subtype [19].

Very few of the *CCND1*+ tumours were *BRCA1*-mutated and none were *BRCA2*-mutated. The low frequency of *BRCA1*-mutated tumours is not unexpected given their ER-negative and basal-like phenotype, and has been shown before [11, 154]. Contrary, *BRCA2*-mutated tumours often express both *CCND1* transcript and protein [11, 155].

The majority of the samples included in Paper II, were also included in a paper by Jönsson *et al.* in which aCGH was performed to identify genomic subtypes of breast cancer [28]. Jönsson *et al*. identified six genomic subtypes named 17q12, basal-complex,

luminal-simple, luminal-complex, amplifier, and mixed subtypes, which harboured different genomic alterations in regard to both site and pattern. The subtypes displayed different clinical behaviour and concordance to the molecular subtypes derived from gene expression data [28]. Almost half (43%) of the *CCND1*+ samples were classified as luminal-complex [28]. A subtype characterised by presence of ER/PR, *BRCA2*-mutated tumours, poor survival and high association to especially luminal B, but also luminal A subtype tumours.

When investigating the gene expression of *CCND1* across all molecular subtypes using a publicly available breast tumour material in GOBO [156, 157], we found highest expression of *CCND1* in luminal B tumours. We then used our own tumour material of *CCND1*+ and *CCND1*- samples classified as luminal B, and compared gene expression of *CCND1* between the two states, and found highest expression of *CCND1* in the *CCND1*+ samples.

Patterns of Copy Number Alterations at 11q13 and Correlation to Gene Expression

We identified six cores that were amplified at different frequencies and combinations covering parts of 11q13.1 to 11q14.1. The largest core was core 3, covering *CCND1*, two cores were situated proximal of core 3, and three cores were distal of core 3. The subtype with the most number of cores amplified was basal-like in accordance with its global pattern of genomic instability [28].

By comparing gene expression status for every gene, for which we had gene expression data, in every core, with copy number data for the same genes, we could study to what extent copy number dosage affects gene expression. Moreover, gene-by-gene, we compared gene expression values among samples with amplification of a core to the samples in the *CCND1*- set. To sum up, nearly all genes were higher expressed among amplified cases than non-amplified.

EMSY, located at 11q13.5, encodes a protein that relocalises in the nucleus upon DNA damage, it binds and represses the transactivation domain of BRCA2 and seems to serve as a surrogate for BRCA2-loss in sporadic breast tumours [158]. The gene is not included in any of the cores since its amplification frequency did not reach our cut off of 90%. This is not surprising since it has been shown to be amplified independently of *CCND1* and *CCND1* amplification was an inclusion criteria for Paper II [158-161]. In the current study, 37% of the tumours had a co-amplification of *EMSY*. All

co-amplified tumours were sporadic, *BRCA1*-mutated or familial while none were *BRCA2*-mutated. Its absence in *BRCA2*-mutated tumours corroborates its function as a *BRCA2*-loss surrogate [158]. Furthermore, *EMSY* was the gene on 11q13 with the highest correlation between gene copy number status and transcription level. This high correlation has been shown before [158, 159].

CCND1 Amplification and Survival

We compared overall survival (OS) between patients with *CCND1*+ and *CCND1*- tumours, respectively, and found a significantly shorter survival for patients with amplification of *CCND1 (P*=9×10^{-4}, log-rank test). Stratification of the tumours based on histological grade and molecular subtypes displayed shorter survival for patients with histological grade 2 or luminal A tumours with CCND1 amplification (*P*=0.0007 and *P*=0.01, respectively, log-rank test). Tendencies of shorter survival for patients with *CCND1*+ tumours in luminal B, HER2-enriched, and normal-like subtypes were found, although nothing was found significant and sample numbers were small.

An interesting finding is that within the luminal A subtype, the survival detriment of having a *CCND1*+ tumour seem to occur more than five years after diagnosis. Generally, 10-20% of all breast cancer patients relapse within 5 years after diagnosis. Another 15% relapse later, however these patients are difficult to identify beforehand. In Paper II we find that in tumours of luminal A subtype, *CCND1* status is potentially important for the identification of late recurrences. Patients with tumours characterised as luminal A have a fairly good diagnosis (ER-positive and low proliferation), however exceptions exist, and these patients need to be identified. If our results can be repeated using other tumour materials and by other research groups, *CCND1* status could potentially be used as a biomarker for which patients should receive aggressive treatments and be carefully monitored many years after their initial diagnosis.

Methylation Profiling (Paper III)

Using 189 well-characterised primary breast tumours and methylation microarrays containing more than 1,500 CpGs corresponding to more than 800 cancer-related genes, we performed a methylation analysis and related our results to the molecular

subtypes of breast tumours. Additionally, for the majority of the tumours both gene expression and aCGH data was available [28].

Methylation Profiles and the Molecular Subtypes

Unsupervised hierarchical clustering resulted in three main branches. Intriguingly, these three branches were highly correlated with three of the molecular subtypes of breast tumours: luminal B, luminal A, and basal-like (P=0.0002, P=0.0004 and P=6×10^{-22}, respectively, Fisher's exact test). In addition, the majority of the HER2-enriched tumours were found in the luminal B-associated cluster (P=0.03, Fisher's exact test), while normal-like tumours were found in all clusters.

Survival analysis on the three identified clusters displayed expected results with best outcome in the luminal A-associated cluster and shortest survival in the basal-like-associated cluster (P=0.05, log-rank test) [17, 19]. Also for fraction of genome altered (FGA, representing the percentage of BAC clones subjected to gain or loss for each sample) and S-phase fraction we found expected results *i.e.* both factors were high in the basal-like-associated cluster and low in the luminal A-associated cluster (P=4×10^{-14} and P=4×10^{-9}, respectively, ANOVA) [162].

Significant differences in methylation frequencies were found for the molecular subtypes, family status, hormone receptor status and tissue (normal breast vs. breast tumour) with highest frequencies in luminal B, *BRCA2*-mutated, hormone receptor positive and breast tumour tissue, respectively. *BRCA2*-mutated tumours are often hormone receptor positive and classified as luminal B, consistent with our results [28]. Methylation frequencies were significantly different across all subtypes and the frequency was lowest in normal-like, basal-like and HER2-enriched tumours (P=2×10^{-7}, ANOVA). Our findings of differential methylation patterns and frequencies for the different subtypes have been reproduced by others using both the same and other types of methylation analyses [163-165].

Methylation and Gene Expression

The overall inverse correlation between gene expression and methylation status for genes with both types of data available was high (P=2×10^{-35}, binomial test). To investigate to what extent genes with subtype-specific gene expression are regulated by methylation

we compared the gene set generated by Hu *et al*. [19] for subtype classifications with our genes with subtype-specific methylation. We conclude that around 25% of genes with subtype-specific expression could be regulated by methylation. It will be interesting to see if this holds true when larger methylation analyses in regard to both sample numbers and CpGs will be performed in the future.

PRC2 and the Molecular Subtypes

To explore whether genes in any of the subtypes with lower methylation frequencies could be silenced by histone modifications generated by PRC2, we investigated the gene expression of its core member *EZH2*. Intriguingly, we found significant differences between subtypes with highest expression in basal-like and HER2-enriched tumours (P=1×10^{-31}, ANOVA), which also had low methylation frequencies, indicating that genes might indeed be silenced by PRC2-mediated formation of H3K27me3. The high expression of EZH2 in basal-like tumours has been shown before [115, 166, 167].

We then used a PRC2 target gene list identified by Lee *et al*. using ChIP arrays on ES cells [102] to investigate expression levels of PRC2 target genes in the subtypes. We found that both basal-like and luminal B tumours harbour low expression of these genes (P=5×10^{-18}, ANOVA), however, methylation levels for the same genes were high only among luminal B tumours (P=0.004, t-test). Further, we investigated the expression of three PRC2 target gene sets derived from tumour cells. The first set represents target genes in less differentiated cells and gave similar results in our material as for the ES PRC2 gene set [168]. The second and the third sets represent target genes derived from more differentiated cells [101, 169] and for these genes the expression was low in luminal B and high in basal-like tumours, accordingly the methylation levels were low in basal-like and high in luminal B tumours. Interestingly, the PRC2 target genes derived from the more differentiated cells had a tendency to be more methylated in luminal B than luminal A tumours. Our results indicate unique PRC2 occupation patterns for the different subtypes. Indeed, during lineage specification, PRC2 becomes displaced from pro-differentiation genes while being recruited to other sets of genes in a cell-type specific manner [101, 170], which is in line with the differential PRC2 patterns in basal-like and luminal breast tumours. Since PRC2 can attract DNMTs leading to methylation and subsequent gene silencing, this could be the explanation to the increased DNA methylation in luminal B tumours [122].

Molecular Subtypes and Different Mechanisms of Gene Silencing

The subtypes might have a common core of genes silenced by PRC2 as well as additional subtype-specific set of genes. This is what Squazzo *et al*. have shown when comparing targets in adult tumour cells (MCF7) with embryonic tumours [101]. Moreover, it has been shown that PRC2 target genes in ES cells, become hypermethylated in cancer cells [111, 176-179] .

Genes with low expression in normal cells undergo *de novo* methylation in tumours [123], which could explain the higher methylation frequencies observed in luminal B tumours. The low methylation levels of PRC2 targets and the potential gene silencing in basal-like tumours by trimethylation of H3K27 would be consistent with the stem-cell-like character of such tumours, since the mechanism is important for maintaining stem cells in an undifferentiated state [103]. To further explore this, it would be intriguing to perform methylation analysis as well as analysis of histone modifications of isolated normal and cancerous luminal cells since also basal-like tumours are believed to originate from aberrant luminal progenitor cells [26].

Finally, we propose a model in which basal-like tumours develop from cells in a progenitor state with genes silenced by PRC2-mediated trimethylation of H3K27. During differentiation PRC2 is displaced and genes specifying lineage differentiation are transcribed, a state matching luminal A tumours. In cancer cells, altered gene silencing may take place by the recruitment of DNMTs by PRC2 to mediate more stable silencing of PRC2 target genes by promoter methylation. These characteristics match our findings in luminal B tumours.

The reasons behind the altered patterns of DNA methylation and histone modifications remain to be described. The explanation could potentially be found in mutations (or other alterations) in genes that control the two states such as members of the PRC2 complex or other histone methyltransferases, histone demethylases or DNA methyltransferases. Indeed, articles are now frequently being published describing alterations in such epigenetic regulators in different tumour types [125-130].

Occurrence of EZH2 and H3K27me3 (Paper IV)

Immunohistochemical analysis of EZH2 and H3K27me3 was performed on more than 400 tumours placed on six TMAs.

EZH2, H3K27me3, Clinicopathologic Factors and Survival

For both EZH2 and H3K27me3 we found significant associations to ER, PR, histological grade, and S-phase fraction, as well as tumour size for H3K27me3. Intriguingly, the expression patterns were inversed with high expression of EZH2 in tumours that were ER/PR-negative, histological grade 3 and had large S-phase fraction, while occurrence of H3K27me3 was high in tumours that were ER/PR-positive, histological grade 1 and 2, and had low S-phase fraction, as well as in small tumours.

High expression of EZH2 was associated with shorter distant disease free survival (DDFS), which is in line with what has been shown before for both gene and protein expression [112, 114, 115, 166]. For H3K27me3, the pattern was reversed with better survival for patients with high score and decreasing survival time with decreasing score. Wei *et al*. showed that loss of trimethylation at H3K27 was a predictor of poor survival in both breast, ovarian and pancreatic cancers [171], which is compatible with our results.

EZH2, H3K27me3, the Molecular Subtypes and Tumour Development

To validate our findings of high gene expression of *EZH2* in basal-like tumours from Paper III, we used both a publicly available breast tumour gene expression material compiled in GOBO [156, 157], and the protein expression as measured by IHC. We conclude that EZH2 expression is mainly driven by mRNA levels as we found identical expression patterns across the subtypes for both gene and protein expression.

Interestingly, high expression of EZH2 did not lead to abundance of H3K27me3. Basal-like, triple negative, and HER2-positive tumours have highest expression of EZH2, but also least occurrence of trimethylated H3K27. This would indicate that excess of EZH2 potentially alters the formation of the polycomb complex, the preferred histone target, or have other functions.

An additional effect of overexpression of EZH2 has been shown by Gonzalez *et al*. regarding regulation of BRCA1 and resulting in increased proliferation [118, 172]. In the normal breast BRCA1 regulates the G_2/M transition in the cell cycle by decreasing the levels of Cdc25C, which, if activated, increases the activities of Cyclin B/CDK1 and results in entry into mitosis (Figure 3). However, overexpression of EZH2 leads to exportation of BRCA1 to the cytoplasm via phosphorylation of AKT1 in ER-negative cells, which in turn leads to both activation of Cdc25C, Cyclin B/CDK1 and mitosis, as well as genomic instability since BRCA1 no longer can control DNA repair in the nucleus [118, 172]. These are all phenotypic marks of basal-like tumours. The exact mechanism by which EZH2 regulates BRCA1 is not yet known other than indications that it is not transcriptional and not by direct binding between EZH2 and BRCA1 [172]. What causes overexpression of EZH2 is not entirely known either, but factors such as amplification of its chromosomal region (7q36.1) [173], or deletion of microRNA-101, which is a negative regulator of *EZH2* [174] have been described. Furthermore, the transcriptional regulator of EZH2, pRB, is deregulated in 20-30% of all breast tumours [175]. An event that does not affect gene transcription but protein activity is a recurrent gain-of-function mutation leading to increased trimethylation of H3K27 that has been described in lymphomas [125, 126]. However, the mutation has not been described in any epithelial malignancies yet, nor could we find the mutation in our screen of 182 breast tumours.

Tumours classified as luminal A or ER+/HER2-/Ki67low harboured low expression of EZH2 but still had rather high occurrence of H3K27me3. This could possibly be explained by the distinct chromatin modification landscapes that have been studied in embryonic and fibroblast cell lines [120]. Fibroblasts have larger regions of the two silencing marks H3K27me3 and H3K9me3 in comparison with embryonic stem cells [120]. This could potentially explain the differences between the more differentiated breast cancer subtypes and the less differentiated basal-like or triple negative tumours.

In summary, EZH2 is a potential target for treatment since it is overexpressed in a substantial fraction of breast tumours and downregulation decreases mitotic activity *in vivo* as well as breast tumour growth [118] - a treatment that could be of benefit for many patients with ER-negative tumours.

Conclusions and Future Perspectives

The main results can be summarised as follows:

- Alterations in the PI3K pathway are central in breast tumour development. Derangements occur at several levels of the pathway, correlate with different clinicopathologic parameters, and lead to poor patient survival. Mutations in *PIK3CA* activate the pathway in ER-positive tumours, while loss of PTEN activate the pathway in ER-negative tumours. Large sequencing projects will in the future aid in creating a gathered picture of all alterations to the pathway in both individual tumours as well as tumour subtypes.

- *CCND1* is amplified and overexpressed in predominantly luminal B but also luminal A tumours. Copy number status of *CCND1* can divide both histological grade 2 and luminal A tumours into groups of patients with good or poor survival. The genomic alterations on chromosome 11q13 are diverse with several regions amplified in different combinations. Future studies will reveal whether *CCND1* status could be used as a biomarker for patients with tumours classified as luminal A or histological grade 2, to distinguish aggressive disease and patients needing adjuvant treatment.

- The molecular subtypes of breast tumours harbour different DNA methylation patterns, especially basal-like, luminal A and luminal B tumours. PRC2 target genes are silenced in both luminal B and basal-like tumours. In luminal B tumours they are silenced by DNA methylation, and in basal-like tumours they could potentially be silenced by trimethylation of H3K27. Studies of homogenous cell populations of normal and cancerous breast cells, respectively, would be intriguing to be able to describe the PRC2-occupational patterns in these cells. In addition, large whole genome-covering methylation analyses will reveal the extent of the differential methylation patterns of the subtypes in the near future.

- The gene expression patterns of *EZH2* were validated and mRNA levels and protein levels correlate perfectly across the subtypes. High expression of EZH2 does not immediately lead to high occurrence of H3K27me3. A question raised from this study is: which gene or event alters the patterns of methylation and/or histone modification leading to tumour development? Future sequencing studies might find the answer to that question.

General Conclusions

In summary, this thesis adds further pieces to the puzzle on the heterogeneity among breast tumours, demonstrating that, at the molecular level, each breast tumour carries its own specific set of genetic and epigenetic aberrations resulting in collective phenotypes as demonstrated by the molecular subtypes.

In the current thesis results from genetic analyses on different levels (genetic, genomic, epigenomic) are presented. In Table 1, I summarise some of the current knowledge as well as results from this thesis about the molecular subtypes of breast tumours.

The heterogeneity among breast tumours is reflected in the identification of at least five molecular subtypes. However, heterogeneity also exists within each subtype. Potentially, with the aid of large next generation sequencing projects and subsequent data mining the diversity can be narrowed down within the subtypes to alterations of collective pathways or parts of pathways. Alternatively, the alterations could be different but the end result phenotypically similar. Co-existence of several genetic alterations in different pathways adds another dimension to the phenotypic behaviour of breast tumours and complicates the wish to categorise the patients for treatment purposes. Fruitful collaborations between biologists and bioinformaticians during analyses of the data could lead to the identification of subgroups within the subtypes with common phenotypes, leading to the opportunity to develop a library of drugs for personalised treatment and improved survival.

Table 1. Simplified summary of the phenotypic behaviour of the molecular subtypes of breast tumours based on results from this thesis as well as prior publications [16-20, 26, 28-30].

	Basal-like	**HER2-enriched**	**Luminal A**	**Luminal B**	**Normal-like**
ER	Neg	Neg/Pos	Pos	Pos	Neg/Pos
Proliferation	High	High	Low	High	Low
Prognosis	Poor	Poor	Good	Intermediate	Intermediate
Specific genetic alterations	*BRCA1*^{mut} *BRCA1*^{met}	HER2^{amp}		*BRCA2*^{mut}	
Genomic subtype	Basal-complex	17q12	Luminal-simple	Luminal-complex	
PI3K pathway alterations	PTEN^{loss}	HER2^{amp}	*PIK3CA*$^{mutNKD?}$	*PIK3CA*$^{mutKD?}$	
CCND1	Low	Low	High	High	Intermediate
Promoter DNA methylation	Low	Low	Intermediate	High	Low
EZH2	High	High	Low	Intermediate	Low
H3K27me3	Low	Intermediate	High	Low	High

Neg=negative; Pos=positive; mut=mutant; met=methylation; amp=amplified; mutKD=kinase domain mutation; mutNKD=non-kinase domain mutation

Summary in Swedish

Populärvetenskaplig sammanfattning

Bröstcancer är den vanligaste cancerformen bland kvinnor i Sverige. Ungefär var tionde kvinna beräknas drabbas av bröstcancer, och i Sverige insjuknar ungefär 7.000 kvinnor årligen. Glädjande nog har överlevnaden förbättrats och nästan 90% lever 5 år efter sin diagnos och cirka två tredjedelar kan anses botade. Trots detta dör cirka 1.500 kvinnor i sjukdomen i Sverige varje år. Bröstcancer är vanligare i västvärlden än i utvecklingsländer, sannolikt beroende på annorlunda diet och livsstil men möjligen även beroende på screeningprogram som t.ex. mammografi, där tumörer som eventuellt skulle förblivit asymptomatiska ibland upptäcks. Riskfaktorer för att drabbas av bröstcancer inkluderar en familjehistoria för sjukdomen, stigande ålder, rökning, alkohol, hormonersättningsbehandling i och efter menopaus samt reproduktiva faktorer såsom barnlöshet eller hög ålder vid första graviditet, tidig pubertet och sen menopaus. Medfödda förändringar i två gener, *BRCA1* och *BRCA2*, har visat sig orsaka en stor andel av de cancerfall som drabbar personer med en stark familjehistoria av framförallt bröst- och äggstockscancer. Sedan mitten av 90-talet när sekvens och funktion för *BRCA1* och *BRCA2* karakteriserats, utförs mutationsscreening för att möjliggöra prediktiv testning av friska familjemedlemmar, prevention eller uppföljning för tidig upptäckt av sjukdomen. Kvinnor med en mutation i antingen *BRCA1* eller *BRCA2* har en risk av upp till 85% att utveckla bröstcancer.

Bröstcancer är en heterogen sjukdom både ur kliniskt och tumörbiologiskt perspektiv med ett stort antal olika genetiska förändringar som ansamlas under tumörutvecklingen. Gemensamt för all cancer är förlust av de kontrollmekanismer för celldelning vilka normalt finns i alla celler. När denna kontroll är satt ur spel kan cellen dela sig ett obegränsat antal gånger och eftersom även dottercellerna ärver dessa fel leder detta till att en tumörmassa bildas. Om cellerna även erhåller förmågan att växa invasivt i bröstet, sprida sig och börja växa på annan plats i kroppen (metastasera) benämns

tumören malign, medan en tumör som inte kan sprida sig kallas benign. Även benigna tumörer kan dock orsaka skada på den plats de växer.

Genom senare års forskning har man lyckats identifiera minst fem olika tumörsubtyper av bröstcancer. Varje subtyp har sina egna karakteristiska genetiska och epigenetiska förändringar vilka leder till speciella genuttrycksmönster. Medan genetik definieras som läran om vårt genetiska material (dvs. DNA), handlar epigenetik om sådana faktorer som påverkar geners uttryck utan att påverka DNA-sekvensen. Mycket forskning riktas nu mot att hitta och mer exakt karakterisera tumörspecifika förändringar med målet att utveckla bättre och mer riktade behandlingar. I denna avhandling presenteras resultat från fyra studier i vilka vi på olika nivåer (genetiska, epigenetiska och på genuttrycksnivå) har analyserat förändringar i brösttumörer och relaterat dessa till tumörsubtyper och patientöverlevnad.

I studie I genomfördes en mutationsscreening av generna *PIK3CA* och *PTEN*. Även förekomsten av det protein som *PTEN* kodar för mättes. Proteiner samverkar i nätverk för att styra celldelning och *PIK3CA* och *PTEN* finns i samma signalväg, där *PIK3CA* ökar cellens tillväxt medan *PTEN* bromsar den. Vi fann mutationer i *PIK3CA* som ökar dess normala funktion i ca 26% av alla tumörer och fr.a. i de som var hormonkänsliga, dvs. uttryckte receptorer för östrogen. Förlust av PTEN fann vi i 31% av tumörerna och oftast i tumörer som inte uttryckte östrogenreceptorer. Förlust av PTEN och mutationer i *PIK3CA* leder båda till ökad celltillväxt. Vi drog slutsatsen att det nätverk som PIK3CA och PTEN är en del av är förändrat i en stor andel av bröstcancer och kan vara en förklaring till tumörutveckling och aggressivt växtsätt, och vidare att tumörens hormonreceptorstatus påverkar hur denna signalväg påverkas.

I studie II har vi studerat mönster av ökning (amplifiering) eller förlust (deletion) av genetiskt material i tumörer som har en amplifiering av genen *CCND1*. *CCND1* kodar för ett protein som heter cyclin D1 vilket är involverat i cellcykelkontroll och celldelning. Om genen är amplifierad kan det stimulera celldelning och därmed ökad celltillväxt. Vi fann att amplifiering av gener både precis före och efter *CCND1* på kromosom 11 följer ett visst mönster och att dessa andra gener också kan vara viktiga för tumörtillväxten. Vi fann även att amplifiering och överuttryck av *CCND1* oftast sker i östrogenreceptorpositiva tumörer.

I studie III studerade vi om det fanns epigenetiska mönster i tumörer i likhet med de mönster man tidigare funnit baserat på genuttryck för olika subtyper av brösttumörer.

Vi studerade en viss typ av epigenetisk förändring som benämns CpG metylering och som innebär tillförsel av en metylgrupp (CH_3) till basen cytosin (C) när den efterföljs av en guanosin (G) i DNA-kedjan. När metylering sker av CpG i anslutning till geners kontrollregioner leder detta till att genen inte längre uttrycks. Vi fann specifika mönster av gener som inaktiverats via metylering och att dessa var associerade med åtminstone tre av de tidigare definierade tumörsubtyperna samt att metyleringsfrekvensen varierade mellan subtyperna. Utöver detta fann vi ett högt genuttryck av genen *EZH2* i en viss tumörsubtyp som samtidigt hade låg metyleringsfrekvens.

Studie IV är en fortsättning på studie III där vi validerade genuttrycket för *EZH2* genom att mäta förekomsten av det protein som *EZH2* kodar för. Vi fann en god överensstämmelse mellan genuttryck (RNA) och proteinuttryck för EZH2. EZH2 är involverat i utförandet av en annan epigenetisk mekanism som kallas för histonmodifiering. Histoner är en grupp proteiner som gör att DNA-kedjan kan vindas upp till kromatin och packas ihop till kromosomer. Olika typer av modifieringar av histonerna leder till att DNA packas olika tätt vilket i sin tur påverkar hur generna kan uttryckas. EZH2 trimetylerar lysin nummer 27 på histon nummer tre vilket orsakar tät packning av DNA och därmed tystade gener. Vi mätte även förekomsten av denna histonmodifiering och fann en avsevärd variation över tumörsubtyperna.

Förhoppningen är att i framtiden förbättra prognosen för bröstcancerpatienter genom att utveckla riktade behandlingar mot de genetiska eller epigenetiska defekter som varje individ har. Sammantaget bidrar denna avhandling med mer kunskap kring de olika bröstcancersubtyperna och vilka genetiska och epigenetiska förändringar som karakteriserar dem.

Acknowledgements

Perhaps my PhD years were both more numerous and more winding than I hoped when I began this journey. Many have been involved in the making of this thesis, everybody with their own role as supervisor, colleague, family or friend. I am so grateful for all help, support and cheerful comments I have received during these years. Nobody is forgotten!

Foremost, I would like to thank my past and present main supervisors *Åke Borg* and *Markus Ringnér*. *Åke*, thank you for taking me in as a PhD student in your group. It has been a pleasure to work with you and see how you both successfully and skilfully have quintupled the group size since I began, as well as have created a tremendously nice working environment. *Markus*, you are an impressive person with a never-ending enthusiasm and energy storage. I have learned a lot from you and am happy to be part of your Computational unit…even though computations never have been my cup of tea. I am very much looking forward to our coming research projects.

My co-supervisor *Cecilia Hegardt*, thank you for all support and for always being there no matter what the topic was. Not many are as thorough and kind-hearted as you!

My co-supervisor *Mårten Fernö*, thank you for all discussions and support especially in the beginning of my PhD-years. Who knows how my thesis would have looked if that first project had worked out.

My former roommates and PhD-colleagues, *Göran Jönsson*, *Johan Staaf* and *Johan Vallon-Christersson*, thank you for all your help, for spilling bits and pieces of your vast knowledge about microarrays on me, and for all fun times we have had during the years! My present roommate *Martin Lauss,* thank you for nice discussions and help with R.

Lao Saal, thank you for support and fruitful collaborations.

Gabriella Honeth and *Anna Isinger Ekstrand*, thank you for all good times in the

"corner office", and for cheering and supporting whenever needed. I am so glad that we have somewhat the same working hours Gabriella, when panic attacks at 2 a.m., I know whom to turn to.

A special thank you to all personnel at the Department of Oncology, especially *Ulla*, *Tessan*, *Steina*, *Carina, Kristina L*, *Susanne A*, *PhD-colleagues*, and everybody in the *BRCA-team* and *SCIBLU-facility*, for always being so helpful.

All my dear friends for cheering me on. A special thank you to my neighbour *Kristina Persson* for being such a good friend and for drawing the illustration on the cover of this thesis.

I would also like to thank *Pat* and *Roy Townsend* for always helping Martin, me and our boys out whenever needed.

My sister *Jenny*, thank you for persistently talking about the Biotechnology program at SLU, Uppsala, and for (I think a bit unconsciously) making me interested in science.

My parents, *Birgit* and *Bobby,* thank you for endless support and help no matter what. It is a relief to know that we always can turn to you.

Martin, thank you for putting up with me in spite of me spending much more time with my computer than with you.

Mina snuttisar, min *Melvin* och min *Kalle*. Mamma älskar er så mycket att det gör ont. Det är ni som är verkligheten och det viktigaste som finns! ♥

My thoughts and gratitude also go to all patients who generously let us use their tissue in our research.

The work presented herein is supported by funding from the Swedish Cancer Society, the Knut & Alice Wallenberg Foundation, the Foundation for Strategic Research through the Lund Centre for Translational Research (CREATE Health), the Mrs Berta Kamprad Foundation, the Gunnar Nilsson Cancer Foundation, the Swedish Research Council, the Lund University Hospital Research Funds, and the IngaBritt and Arne Lundberg Foundation.

References

1. Russo J, Russo IH: **Development of the human breast**. *Maturitas* 2004, 49(1):2-15.
2. Hanahan D, Weinberg RA: **The hallmarks of cancer**. *Cell* 2000, 100(1):57-70.
3. Hanahan D, Weinberg RA: **Hallmarks of cancer: the next generation**. *Cell* 2011, 144(5):646-674.
4. Socialstyrelsen, Cancerfonden: **Cancer i siffror 2009**. In.; 2009.
5. Ziegler RG, Hoover RN, Pike MC, Hildesheim A, Nomura AM, West DW, Wu-Williams AH, Kolonel LN, Horn-Ross PL, Rosenthal JF *et al*: **Migration patterns and breast cancer risk in Asian-American women**. *J Natl Cancer Inst* 1993, 85(22):1819-1827.
6. Ryden L, Haglund M, Bendahl PO, Hatschek T, Kolaric A, Kovacs A, Olsson A, Olsson H, Strand C, Ferno M: **Reproducibility of human epidermal growth factor receptor 2 analysis in primary breast cancer: a national survey performed at pathology departments in Sweden**. *Acta Oncol* 2009, 48(6):860-866.
7. van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT *et al*: **Gene expression profiling predicts clinical outcome of breast cancer**. *Nature* 2002, 415(6871):530-536.
8. van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ *et al*: **A gene-expression signature as a predictor of survival in breast cancer**. *N Engl J Med* 2002, 347(25):1999-2009.
9. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T *et al*: **A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer**. *N Engl J Med* 2004, 351(27):2817-2826.
10. Sotiriou C, Wirapati P, Loi S, Harris A, Fox S, Smeds J, Nordgren H, Farmer P, Praz V, Haibe-Kains B *et al*: **Gene Expression Profiling in Breast Cancer: Understanding the Molecular Basis of Histologic Grade To Improve Prognosis**. *J Natl Cancer Inst* 2006, 98(4):262-272.
11. Hedenfalk I, Duggan D, Chen Y, Radmacher M, Bittner M, Simon R, Meltzer P, Gusterson B, Esteller M, Kallioniemi OP *et al*: **Gene-expression profiles in hereditary breast cancer**. *N Engl J Med* 2001, 344(8):539-548.
12. Bogaerts J, Cardoso F, Buyse M, Braga S, Loi S, Harrison JA, Bines J, Mook S, Decker

N, Ravdin P *et al*: **Gene signature evaluation as a prognostic tool: challenges in the design of the MINDACT trial**. *Nat Clin Pract Oncol* 2006, 3(10):540-551.

13. Paik S: **Development and clinical utility of a 21-gene recurrence score prognostic assay in patients with early breast cancer treated with tamoxifen**. *Oncologist* 2007, 12(6):631-635.
14. Eden P, Ritz C, Rose C, Ferno M, Peterson C: **"Good Old" clinical markers have similar power in breast cancer prognosis as microarray gene expression profilers**. *Eur J Cancer* 2004, 40(12):1837-1841.
15. Gruvberger SK, Ringner M, Eden P, Borg A, Ferno M, Peterson C, Meltzer PS: **Expression profiling to predict outcome in breast cancer: the influence of sample selection**. *Breast Cancer Res* 2003, 5(1):23-26.
16. Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA *et al*: **Molecular portraits of human breast tumours**. *Nature* 2000, 406(6797):747-752.
17. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS *et al*: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications**. *Proc Natl Acad Sci U S A* 2001, 98(19):10869-10874.
18. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S *et al*: **Repeated observation of breast tumor subtypes in independent gene expression data sets**. *Proc Natl Acad Sci U S A* 2003, 100(14):8418-8423.
19. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, Livasy C, Carey LA, Reynolds E, Dressler L *et al*: **The molecular portraits of breast tumors are conserved across microarray platforms**. *BMC Genomics* 2006, 7:96.
20. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z *et al*: **Supervised risk predictor of breast cancer based on intrinsic subtypes**. *J Clin Oncol* 2009, 27(8):1160-1167.
21. Bergamaschi A, Kim YH, Wang P, Sorlie T, Hernandez-Boussard T, Lonning PE, Tibshirani R, Borresen-Dale AL, Pollack JR: **Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer**. *Genes Chromosomes Cancer* 2006, 45(11):1033-1040.
22. Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, Kuo WL, Lapuk A, Neve RM, Qian Z, Ryder T *et al*: **Genomic and transcriptional aberrations linked to breast cancer pathophysiologies**. *Cancer Cell* 2006, 10(6):529-541.
23. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K, Hess KR, Stec J, Ayers M, Wagner P *et al*: **Breast cancer molecular subtypes respond**

differently to preoperative chemotherapy. *Clin Cancer Res* 2005, 11(16):5678-5685.

24. Smid M, Wang Y, Zhang Y, Sieuwerts AM, Yu J, Klijn JG, Foekens JA, Martens JW: **Subtypes of breast cancer show preferential site of relapse**. *Cancer Res* 2008, 68(9):3108-3114.
25. Honeth G, Bendahl PO, Ringner M, Saal LH, Gruvberger-Saal SK, Lovgren K, Grabau D, Ferno M, Borg A, Hegardt C: **The CD44+/CD24- phenotype is enriched in basal-like breast tumors**. *Breast Cancer Res* 2008, 10(3):R53.
26. Lim E, Vaillant F, Wu D, Forrest NC, Pal B, Hart AH, Asselin-Labat ML, Gyorki DE, Ward T, Partanen A *et al*: **Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers**. *Nat Med* 2009, 15(8):907-913.
27. Laakso M, Loman N, Borg A, Isola J: **Cytokeratin 5/14-positive breast cancer: true basal phenotype confined to BRCA1 tumors**. *Mod Pathol* 2005, 18(10):1321-1328.
28. Jönsson G, Staaf J, Vallon-Christersson J, Ringner M, Holm K, Hegardt C, Gunnarsson H, Fagerholm R, Strand C, Agnarsson BA *et al*: **Genomic subtypes of breast cancer identified by array comparative genomic hybridization display distinct molecular and clinical characteristics**. *Breast Cancer Res* 2010, 12(3):R42.
29. Melchor L, Honrado E, Garcia MJ, Alvarez S, Palacios J, Osorio A, Nathanson KL, Benitez J: **Distinct genomic aberration patterns are found in familial breast cancer associated with different immunohistochemical subtypes**. *Oncogene* 2008, 27(22):3165-3175.
30. Waddell N, Arnold J, Cocciardi S, da Silva L, Marsh A, Riley J, Johnstone CN, Orloff M, Assie G, Eng C *et al*: **Subtypes of familial breast tumours revealed by expression and copy number profiling**. *Breast Cancer Res Treat* 2010, 123(3):661-677.
31. Loo LW, Grove DI, Williams EM, Neal CL, Cousens LA, Schubert EL, Holcomb IN, Massa HF, Glogovac J, Li CI *et al*: **Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes**. *Cancer Res* 2004, 64(23):8541-8549.
32. Fridlyand J, Snijders AM, Ylstra B, Li H, Olshen A, Segraves R, Dairkee S, Tokuyasu T, Ljung BM, Jain AN *et al*: **Breast tumor copy number aberration phenotypes and genomic instability**. *BMC Cancer* 2006, 6:96.
33. Natrajan R, Lambros MB, Rodriguez-Pinilla SM, Moreno-Bueno G, Tan DS, Marchio C, Vatcheva R, Rayter S, Mahler-Araujo B, Fulford LG *et al*: **Tiling path genomic profiling of grade 3 invasive ductal breast cancers**. *Clin Cancer Res* 2009, 15(8):2711-2722.

34. Chin SF, Wang Y, Thorne NP, Teschendorff AE, Pinder SE, Vias M, Naderi A, Roberts I, Barbosa-Morais NL, Garcia MJ *et al*: **Using array-comparative genomic hybridization to define molecular portraits of primary breast cancers**. *Oncogene* 2007, 26(13):1959-1970.

35. Hicks J, Krasnitz A, Lakshmi B, Navin NE, Riggs M, Leibu E, Esposito D, Alexander J, Troge J, Grubor V *et al*: **Novel patterns of genome rearrangement and their association with survival in breast cancer**. *Genome Res* 2006, 16(12):1465-1479.

36. Russnes HG, Vollan HK, Lingjaerde OC, Krasnitz A, Lundin P, Naume B, Sorlie T, Borgen E, Rye IH, Langerod A *et al*: **Genomic architecture characterizes tumor progression paths and fate in breast cancer patients**. *Sci Transl Med* 2010, 2(38):38ra47.

37. Kinzler KW, Vogelstein B: **Cancer-susceptibility genes. Gatekeepers and caretakers**. *Nature* 1997, 386(6627):761, 763.

38. Lane DP: **Cancer. p53, guardian of the genome**. *Nature* 1992, 358(6381):15-16.

39. Borresen-Dale AL: **TP53 and breast cancer**. *Hum Mutat* 2003, 21(3):292-300.

40. Hemminki K, Vaittinen P: **Familial cancers in a nationwide family cancer database: age distribution and prevalence**. *Eur J Cancer* 1999, 35(7):1109-1117.

41. **Familial breast cancer: collaborative reanalysis of individual data from 52 epidemiological studies including 58,209 women with breast cancer and 101,986 women without the disease**. *Lancet* 2001, 358(9291):1389-1399.

42. Hall JM, Lee MK, Newman B, Morrow JE, Anderson LA, Huey B, King MC: **Linkage of early-onset familial breast cancer to chromosome 17q21**. *Science* 1990, 250(4988):1684-1689.

43. Wooster R, Neuhausen SL, Mangion J, Quirk Y, Ford D, Collins N, Nguyen K, Seal S, Tran T, Averill D *et al*: **Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13**. *Science* 1994, 265(5181):2088-2090.

44. Narod SA, Foulkes WD: **BRCA1 and BRCA2: 1994 and beyond**. *Nat Rev Cancer* 2004, 4(9):665-676.

45. Easton DF, Bishop DT, Ford D, Crockford GP: **Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium**. *Am J Hum Genet* 1993, 52(4):678-701.

46. Ford D, Easton DF, Stratton M, Narod S, Goldgar D, Devilee P, Bishop DT, Weber B, Lenoir G, Chang-Claude J *et al*: **Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium**. *Am J Hum Genet* 1998, 62(3):676-689.

47. Ford D, Easton DF, Bishop DT, Narod SA, Goldgar DE: **Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium**. *Lancet* 1994, 343(8899):692-695.

48. Knudson AG, Jr.: **Mutation and cancer: statistical study of retinoblastoma**. *Proc Natl Acad Sci U S A* 1971, 68(4):820-823.

49. Pardee AB: **G1 events and regulation of cell proliferation**. *Science* 1989, 246(4930):603-608.

50. Canepa ET, Scassa ME, Ceruti JM, Marazita MC, Carcagno AL, Sirkin PF, Ogara MF: **INK4 proteins, a family of mammalian CDK inhibitors with novel biological functions**. *IUBMB Life* 2007, 59(7):419-426.

51. Matsushime H, Quelle DE, Shurtleff SA, Shibuya M, Sherr CJ, Kato JY: **D-type cyclin-dependent kinase activity in mammalian cells**. *Mol Cell Biol* 1994, 14(3):2066-2076.

52. Dyson N: **The regulation of E2F by pRB-family proteins**. *Genes Dev* 1998, 12(15):2245-2262.

53. Roy PG, Thompson AM: **Cyclin D1 and breast cancer**. *Breast* 2006, 15(6):718-727.

54. Hui R, Campbell DH, Lee CS, McCaul K, Horsfall DJ, Musgrove EA, Daly RJ, Seshadri R, Sutherland RL: **EMS1 amplification can occur independently of CCND1 or INT-2 amplification at 11q13 and may identify different phenotypes in primary breast cancer**. *Oncogene* 1997, 15(13):1617-1623.

55. Hui R, Ball JR, Macmillan RD, Kenny FS, Prall OW, Campbell DH, Cornish AL, McClelland RA, Daly RJ, Forbes JF *et al*: **EMS1 gene expression in primary breast cancer: relationship to cyclin D1 and oestrogen receptor expression and patient survival**. *Oncogene* 1998, 17(8):1053-1059.

56. Ormandy CJ, Musgrove EA, Hui R, Daly RJ, Sutherland RL: **Cyclin D1, EMS1 and 11q13 Amplification in Breast Cancer**. *Breast Cancer Res Treat* 2003, 78(3):323-335.

57. Musgrove EA, Lee CS, Buckley MF, Sutherland RL: **Cyclin D1 induction in breast cancer cells shortens G1 and is sufficient for cells arrested in G1 to complete the cell cycle**. *Proc Natl Acad Sci U S A* 1994, 91(17):8022-8026.

58. Zwijsen RM, Wientjens E, Klompmaker R, van der Sman J, Bernards R, Michalides RJ: **CDK-independent activation of estrogen receptor by cyclin D1**. *Cell* 1997, 88(3):405-415.

59. Samuels Y, Diaz LA, Jr., Schmidt-Kittler O, Cummins JM, Delong L, Cheong I, Rago C, Huso DL, Lengauer C, Kinzler KW *et al*: **Mutant PIK3CA promotes cell growth and invasion of human cancer cells**. *Cancer Cell* 2005, 7(6):561-573.

60. Bachman KE, Argani P, Samuels Y, Silliman N, Ptak J, Szabo S, Konishi H, Karakas B, Blair BG, Lin C *et al*: **The PIK3CA gene is mutated with high frequency in human breast cancers**. *Cancer Biology & Therapy* 2004, 3(8):772-775.

61. Lee JW, Soung YH, Kim SY, Lee HW, Park WS, Nam SW, Kim SH, Lee JY, Yoo NJ, Lee SH: **PIK3CA gene is frequently mutated in breast carcinomas and hepatocellular carcinomas**. *Oncogene* 2005, 24(8):1477-1480.

62. Kadota M, Sato M, Duncan B, Ooshima A, Yang HH, Diaz-Meyer N, Gere S, Kageyama S, Fukuoka J, Nagata T *et al*: **Identification of novel gene amplifications in breast cancer and coexistence of gene amplification with an activating mutation of PIK3CA**. *Cancer Res* 2009, 69(18):7357-7365.

63. Shayesteh L, Lu Y, Kuo WL, Baldocchi R, Godfrey T, Collins C, Pinkel D, Powell B, Mills GB, Gray JW: **PIK3CA is implicated as an oncogene in ovarian cancer**. *Nat Genet* 1999, 21(1):99-102.

64. Samuels Y, Wang Z, Bardelli A, Silliman N, Ptak J, Szabo S, Yan H, Gazdar A, Powell SM, Riggins GJ *et al*: **High frequency of mutations of the PIK3CA gene in human cancers**. *Science* 2004, 304(5670):554.

65. Isakoff SJ, Engelman JA, Irie HY, Luo J, Brachmann SM, Pearline RV, Cantley LC, Brugge JS: **Breast cancer-associated PIK3CA mutations are oncogenic in mammary epithelial cells**. *Cancer Res* 2005, 65(23):10992-11000.

66. Zhao JJ, Liu Z, Wang L, Shin E, Loda MF, Roberts TM: **The oncogenic properties of mutant p110alpha and p110beta phosphatidylinositol 3-kinases in human mammary epithelial cells**. *Proc Natl Acad Sci U S A* 2005, 102(51):18443-18448.

67. Bader AG, Kang S, Vogt PK: **Cancer-specific mutations in PIK3CA are oncogenic in vivo**. *Proc Natl Acad Sci U S A* 2006, 103(5):1475-1479.

68. Katso R, Okkenhaug K, Ahmadi K, White S, Timms J, Waterfield MD: **Cellular function of phosphoinositide 3-kinases: implications for development, homeostasis, and cancer**. *Annu Rev Cell Dev Biol* 2001, 17:615-675.

69. Li J, Yen C, Liaw D, Podsypanina K, Bose S, Wang SI, Puc J, Miliaresis C, Rodgers L, McCombie R *et al*: **PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer**. *Science* 1997, 275(5308):1943-1947.

70. Leslie NR, Downes CP: **PTEN function: how normal cells control it and tumour cells lose it**. *Biochem J* 2004, 382(Pt 1):1-11.

71. Depowski PL, Rosenthal SI, Ross JS: **Loss of expression of the PTEN gene protein product is associated with poor outcome in breast cancer**. *Mod Pathol* 2001, 14(7):672-676.

72. Saal LH, Gruvberger-Saal SK, Persson C, Lovgren K, Jumppanen M, Staaf J, Jonsson G, Pires MM, Maurer M, Holm K *et al*: **Recurrent gross mutations of the PTEN**

tumor suppressor gene in breast cancers with deficient DSB repair. *Nat Genet* 2008, 40(1):102-107.

73. Garcia JM, Silva J, Pena C, Garcia V, Rodriguez R, Cruz MA, Cantos B, Provencio M, Espana P, Bonilla F: **Promoter methylation of the PTEN gene is a common molecular change in breast cancer**. *Genes Chromosomes Cancer* 2004, 41(2):117-124.

74. Saal LH, Johansson P, Holm K, Gruvberger-Saal SK, She QB, Maurer M, Koujak S, Ferrando AA, Malmstrom P, Memeo L *et al*: **Poor prognosis in carcinoma is associated with a gene expression signature of aberrant PTEN tumor suppressor pathway activity**. *Proc Natl Acad Sci U S A* 2007, 104(18):7564-7569.

75. Khan S, Kumagai T, Vora J, Bose N, Sehgal I, Koeffler PH, Bose S: **PTEN promoter is methylated in a proportion of invasive breast cancers**. *Int J Cancer* 2004, 112(3):407-410.

76. Shaw RJ, Cantley LC: **Ras, PI(3)K and mTOR signalling controls tumour cell growth**. *Nature* 2006, 441(7092):424-430.

77. Vivanco I, Sawyers CL: **The phosphatidylinositol 3-Kinase AKT pathway in human cancer**. *Nat Rev Cancer* 2002, 2(7):489-501.

78. Luo J, Manning BD, Cantley LC: **Targeting the PI3K-Akt pathway in human cancer: rationale and promise**. *Cancer Cell* 2003, 4(4):257-262.

79. Collado M, Medema RH, Garcia-Cao I, Dubuisson ML, Barradas M, Glassford J, Rivas C, Burgering BM, Serrano M, Lam EW: **Inhibition of the phosphoinositide 3-kinase pathway induces a senescence-like arrest mediated by p27Kip1**. *J Biol Chem* 2000, 275(29):21960-21968.

80. Soucek T, Yeung RS, Hengstschlager M: **Inactivation of the cyclin-dependent kinase inhibitor p27 upon loss of the tuberous sclerosis complex gene-2**. *Proc Natl Acad Sci U S A* 1998, 95(26):15653-15658.

81. Alt JR, Cleveland JL, Hannink M, Diehl JA: **Phosphorylation-dependent regulation of cyclin D1 nuclear export and cyclin D1-dependent cellular transformation**. *Genes Dev* 2000, 14(24):3102-3114.

82. Mayo LD, Donner DB: **A phosphatidylinositol 3-kinase/Akt pathway promotes translocation of Mdm2 from the cytoplasm to the nucleus**. *Proc Natl Acad Sci U S A* 2001, 98(20):11598-11603.

83. Muise-Helmericks RC, Grimes HL, Bellacosa A, Malstrom SE, Tsichlis PN, Rosen N: **Cyclin D expression is controlled post-transcriptionally via a phosphatidylinositol 3-kinase/Akt-dependent pathway**. *J Biol Chem* 1998, 273(45):29864-29872.

84. Puc J, Keniry M, Li HS, Pandita TK, Choudhury AD, Memeo L, Mansukhani M, Murty VV, Gaciong Z, Meek SE *et al*: **Lack of PTEN sequesters CHK1 and initiates**

genetic instability. *Cancer Cell* 2005, 7(2):193-204.

85. Cohen P, Goedert M: **GSK3 inhibitors: development and therapeutic potential**. *Nat Rev Drug Discov* 2004, 3(6):479-487.
86. Ayala JE, Streeper RS, Desgrosellier JS, Durham SK, Suwanichkul A, Svitek CA, Goldman JK, Barr FG, Powell DR, O'Brien RM: **Conservation of an insulin response unit between mouse and human glucose-6-phosphatase catalytic subunit gene promoters: transcription factor FKHR binds the insulin response sequence**. *Diabetes* 1999, 48(9):1885-1889.
87. Durham SK, Suwanichkul A, Scheimann AO, Yee D, Jackson JG, Barr FG, Powell DR: **FKHR binds the insulin response element in the insulin-like growth factor binding protein-1 promoter**. *Endocrinology* 1999, 140(7):3140-3146.
88. Bird A: **Perceptions of epigenetics**. *Nature* 2007, 447(7143):396-398.
89. Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A: **An operational definition of epigenetics**. *Genes Dev* 2009, 23(7):781-783.
90. Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, Li BF: **Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1**. *Science* 1997, 277(5334):1996-2000.
91. Okano M, Bell DW, Haber DA, Li E: **DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development**. *Cell* 1999, 99(3):247-257.
92. Chedin F, Lieber MR, Hsieh CL: **The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a**. *Proc Natl Acad Sci U S A* 2002, 99(26):16916-16921.
93. Jia D, Jurkowska RZ, Zhang X, Jeltsch A, Cheng X: **Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation**. *Nature* 2007, 449(7159):248-251.
94. Tost Jr: **Epigenetics**. Wymondham: Caister Academic; 2008.
95. Bird A: **DNA methylation patterns and epigenetic memory**. *Genes Dev* 2002, 16(1):6-21.
96. Takai D, Jones PA: **Comprehensive analysis of CpG islands in human chromosomes 21 and 22**. *Proc Natl Acad Sci U S A* 2002, 99(6):3740-3745.
97. Wang Y, Leung FC: **An evaluation of new criteria for CpG islands in the human genome as gene markers**. *Bioinformatics* 2004, 20(7):1170-1177.
98. Esteller M, Herman JG: **Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours**. *J Pathol* 2002, 196(1):1-7.
99. Rodriguez-Paredes M, Esteller M: **Cancer epigenetics reaches mainstream oncology**. *Nat Med* 2011, 17(3):330-339.

100. Margueron R, Reinberg D: **The Polycomb complex PRC2 and its mark in life**. *Nature* 2011, 469(7330):343-349.

101. Squazzo SL, O'Geen H, Komashko VM, Krig SR, Jin VX, Jang SW, Margueron R, Reinberg D, Green R, Farnham PJ: **Suz12 binds to silenced regions of the genome in a cell-type-specific manner**. *Genome Res* 2006, 16(7):890-900.

102. Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K *et al*: **Control of developmental regulators by Polycomb in human embryonic stem cells**. *Cell* 2006, 125(2):301-313.

103. Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K: **Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions**. *Genes Dev* 2006, 20(9):1123-1136.

104. Cao R, Zhang Y: **SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex**. *Mol Cell* 2004, 15(1):57-67.

105. Kim H, Kang K, Kim J: **AEBP2 as a potential targeting protein for Polycomb Repression Complex PRC2**. *Nucleic Acids Res* 2009, 37(9):2940-2950.

106. Landeira D, Sauer S, Poot R, Dvorkina M, Mazzarella L, Jorgensen HF, Pereira CF, Leleu M, Piccolo FM, Spivakov M *et al*: **Jarid2 is a PRC2 component in embryonic stem cells required for multi-lineage differentiation and recruitment of PRC1 and RNA Polymerase II to developmental regulators**. *Nat Cell Biol* 2010, 12(6):618-624.

107. Nekrasov M, Klymenko T, Fraterman S, Papp B, Oktaba K, Kocher T, Cohen A, Stunnenberg HG, Wilm M, Muller J: **Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes**. *EMBO J* 2007, 26(18):4078-4088.

108. Walker E, Chang WY, Hunkapiller J, Cagney G, Garcha K, Torchia J, Krogan NJ, Reiter JF, Stanford WL: **Polycomb-like 2 associates with PRC2 and regulates transcriptional networks during mouse embryonic stem cell self-renewal and differentiation**. *Cell Stem Cell* 2010, 6(2):153-166.

109. Jones A, Wang H: **Polycomb repressive complex 2 in embryonic stem cells: an overview**. *Protein Cell* 2010, 1(12):1056-1062.

110. Kondo Y, Shen L, Cheng AS, Ahmed S, Boumber Y, Charo C, Yamochi T, Urano T, Furukawa K, Kwabi-Addo B *et al*: **Gene silencing in cancer by histone H3 lysine 27 trimethylation independent of promoter DNA methylation**. *Nat Genet* 2008, 40(6):741-750.

111. Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG, Ghosh D, Pienta KJ, Sewalt RG, Otte AP *et al*: **The polycomb group protein EZH2**

is involved in progression of prostate cancer. *Nature* 2002, 419(6907):624-629.

112. Kleer CG, Cao Q, Varambally S, Shen R, Ota I, Tomlins SA, Ghosh D, Sewalt RG, Otte AP, Hayes DF *et al*: **EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells**. *Proc Natl Acad Sci U S A* 2003, 100(20):11606-11611.

113. Raman JD, Mongan NP, Tickoo SK, Boorjian SA, Scherr DS, Gudas LJ: **Increased expression of the polycomb group gene, EZH2, in transitional cell carcinoma of the bladder**. *Clin Cancer Res* 2005, 11(24 Pt 1):8570-8576.

114. Bachmann IM, Halvorsen OJ, Collett K, Stefansson IM, Straume O, Haukaas SA, Salvesen HB, Otte AP, Akslen LA: **EZH2 expression is associated with high proliferation rate and aggressive tumor subgroups in cutaneous melanoma and cancers of the endometrium, prostate, and breast**. *J Clin Oncol* 2006, 24(2):268-273.

115. Pietersen AM, Horlings HM, Hauptmann M, Langerod A, Ajouaou A, Cornelissen-Steijger P, Wessels LF, Jonkers J, van de Vijver MJ, van Lohuizen M: **EZH2 and BMI1 inversely correlate with prognosis and TP53 mutation in breast cancer**. *Breast Cancer Res* 2008, 10(6):R109.

116. Bracken AP, Pasini D, Capra M, Prosperini E, Colli E, Helin K: **EZH2 is downstream of the pRB-E2F pathway, essential for proliferation and amplified in cancer**. *EMBO J* 2003, 22(20):5323-5335.

117. Muller H, Bracken AP, Vernell R, Moroni MC, Christians F, Grassilli E, Prosperini E, Vigo E, Oliner JD, Helin K: **E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis**. *Genes Dev* 2001, 15(3):267-285.

118. Gonzalez ME, Li X, Toy K, DuPrie M, Ventura AC, Banerjee M, Ljungman M, Merajver SD, Kleer CG: **Downregulation of EZH2 decreases growth of estrogen receptor-negative invasive breast carcinoma and requires BRCA1**. *Oncogene* 2009, 28(6):843-853.

119. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K *et al*: **A bivalent chromatin structure marks key developmental genes in embryonic stem cells**. *Cell* 2006, 125(2):315-326.

120. Hawkins RD, Hon GC, Lee LK, Ngo Q, Lister R, Pelizzola M, Edsall LE, Kuan S, Luu Y, Klugman S *et al*: **Distinct epigenomic landscapes of pluripotent and lineage-committed human cells**. *Cell Stem Cell* 2010, 6(5):479-491.

121. Mohn F, Weber M, Rebhan M, Roloff TC, Richter J, Stadler MB, Bibel M, Schubeler D: **Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors**. *Mol Cell* 2008, 30(6):755-766.

122. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM *et al*: **The Polycomb group protein EZH2 directly controls DNA methylation**. *Nature* 2006, 439(7078):871-874.

123. Keshet I, Schlesinger Y, Farkash S, Rand E, Hecht M, Segal E, Pikarski E, Young RA, Niveleau A, Cedar H *et al*: **Evidence for an instructive mechanism of de novo methylation in cancer cells**. *Nat Genet* 2006, 38(2):149-153.

124. Schlesinger Y, Straussman R, Keshet I, Farkash S, Hecht M, Zimmerman J, Eden E, Yakhini Z, Ben-Shushan E, Reubinoff BE *et al*: **Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer**. *Nat Genet* 2007, 39(2):232-236.

125. Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, Paul JE, Boyle M, Woolcock BW, Kuchenbauer F *et al*: **Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin**. *Nat Genet* 2010, 42(2):181-185.

126. Bodor C, O'Riain C, Wrench D, Matthews J, Iyengar S, Tayyib H, Calaminici M, Clear A, Iqbal S, Quentmeier H *et al*: **EZH2 Y641 mutations in follicular lymphoma**. *Leukemia* 2011, 25(4):726-729.

127. Ceol CJ, Houvras Y, Jane-Valbuena J, Bilodeau S, Orlando DA, Battisti V, Fritsch L, Lin WM, Hollmann TJ, Ferre F *et al*: **The histone methyltransferase SETDB1 is recurrently amplified in melanoma and accelerates its onset**. *Nature* 2011, 471(7339):513-517.

128. Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC, Harview CL, Brunet JP, Ahmann GJ, Adli M *et al*: **Initial genome sequencing and analysis of multiple myeloma**. *Nature* 2011, 471(7339):467-472.

129. Yan XJ, Xu J, Gu ZH, Pan CM, Lu G, Shen Y, Shi JY, Zhu YM, Tang L, Zhang XW *et al*: **Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia**. *Nat Genet* 2011, 43(4):309-315.

130. Jones S, Wang TL, Shih Ie M, Mao TL, Nakayama K, Roden R, Glas R, Slamon D, Diaz LA, Jr., Vogelstein B *et al*: **Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma**. *Science* 2010, 330(6001):228-231.

131. Malmstrom P, Bendahl PO, Boiesen P, Brunner N, Idvall I, Ferno M: **S-phase fraction and urokinase plasminogen activator are better markers for distant recurrences than Nottingham Prognostic Index and histologic grade in a prospective study of premenopausal lymph node-negative breast cancer**. *J Clin Oncol* 2001, 19(7):2010-2019.

132. Ryden L, Jonsson PE, Chebil G, Dufmats M, Ferno M, Jirstrom K, Kallstrom AC, Landberg G, Stal O, Thorstenson S *et al*: **Two years of adjuvant tamoxifen in**

premenopausal patients with breast cancer: a randomised, controlled trial with long-term follow-up. *Eur J Cancer* 2005, 41(2):256-264.

133. **Gene Expression Omnibus (GEO)** [http://www.ncbi.nlm.nih.gov/geo/]
134. **SCIBLU Genomics** [http://www.lth.se/sciblu]
135. Jönsson G, Staaf J, Olsson E, Heidenblad M, Vallon-Christersson J, Osoegawa K, de Jong P, Oredsson S, Ringner M, Hoglund M *et al*: **High-resolution genomic profiles of breast cancer cell lines assessed by tiling BAC array comparative genomic hybridization**. *Genes Chromosomes Cancer* 2007, 46(6):543-558.
136. Autio R, Hautaniemi S, Kauraniemi P, Yli-Harja O, Astola J, Wolf M, Kallioniemi A: **CGH-Plotter: MATLAB toolbox for CGH-data analysis**. *Bioinformatics* 2003, 19(13):1714-1715.
137. Saal LH, Troein C, Vallon-Christersson J, Gruvberger S, Borg A, Peterson C: **BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data**. *Genome Biol* 2002, 3(8):SOFTWARE0003.
138. Staaf J, Jonsson G, Ringner M, Vallon-Christersson J: **Normalization of array-CGH data: influence of copy number imbalances**. *BMC Genomics* 2007, 8:382.
139. Tusher VG, Tibshirani R, Chu G: **Significance analysis of microarrays applied to the ionizing radiation response**. *Proc Natl Acad Sci U S A* 2001, 98(9):5116-5121.
140. Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M *et al*: **TM4: a free, open-source system for microarray data management and analysis**. *Biotechniques* 2003, 34(2):374-378.
141. **The R Project for Statistical Computing** [http://cran.r-project.org]
142. Wu G, Xing M, Mambo E, Huang X, Liu J, Guo Z, Chatterjee A, Goldenberg D, Gollin SM, Sukumar S *et al*: **Somatic mutation and gain of copy number of PIK3CA in human breast cancer**. *Breast Cancer Res* 2005, 7(5):R609-616.
143. Stemke-Hale K, Gonzalez-Angulo AM, Lluch A, Neve RM, Kuo WL, Davies M, Carey M, Hu Z, Guan Y, Sahin A *et al*: **An integrative genomic and proteomic analysis of PIK3CA, PTEN, and AKT mutations in breast cancer**. *Cancer Res* 2008, 68(15):6084-6091.
144. Kalinsky K, Jacks LM, Heguy A, Patil S, Drobnjak M, Bhanot UK, Hedvat CV, Traina TA, Solit D, Gerald W *et al*: **PIK3CA mutation associates with improved outcome in breast cancer**. *Clin Cancer Res* 2009, 15(16):5049-5059.
145. Li SY, Rong M, Grieu F, Iacopetta B: **PIK3CA mutations in breast cancer are associated with poor outcome**. *Breast Cancer Res Treat* 2006, 96(1):91-95.
146. Zhang H, Liu G, Dziubinski M, Yang Z, Ethier SP, Wu G: **Comprehensive analysis of oncogenic effects of PIK3CA mutations in human mammary epithelial cells**. *Breast Cancer Res Treat* 2008, 112(2):217-227.

147. Zhao L, Vogt PK: **Helical domain and kinase domain mutations in p110alpha of phosphatidylinositol 3-kinase induce gain of function by different mechanisms**. *Proc Natl Acad Sci U S A* 2008, 105(7):2652-2657.

148. Rhei E, Kang L, Bogomolniy F, Federici MG, Borgen PI, Boyd J: **Mutation analysis of the putative tumor suppressor gene PTEN/MMAC1 in primary breast carcinomas**. *Cancer Res* 1997, 57(17):3657-3659.

149. Ueda K, Nishijima M, Inui H, Watatani M, Yayoi E, Okamura J, Yasutomi M, Nakamura Y, Miyoshi Y: **Infrequent mutations in the PTEN/MMAC1 gene among primary breast cancers**. *Jpn J Cancer Res* 1998, 89(1):17-21.

150. Perez-Tenorio G, Alkhori L, Olsson B, Waltersson MA, Nordenskjold B, Rutqvist LE, Skoog L, Stal O: **PIK3CA mutations and PTEN loss correlate with similar prognostic factors and are not mutually exclusive in breast cancer**. *Clin Cancer Res* 2007, 13(12):3577-3584.

151. Campbell IG, Russell SE, Choong DY, Montgomery KG, Ciavarella ML, Hooi CS, Cristiano BE, Pearson RB, Phillips WA: **Mutation of the PIK3CA gene in ovarian and breast cancer**. *Cancer Res* 2004, 64(21):7678-7681.

152. Barbareschi M, Buttitta F, Felicioni L, Cotrupi S, Barassi F, Del Grammastro M, Ferro A, Dalla Palma P, Galligioni E, Marchetti A: **Different prognostic roles of mutations in the helical and kinase domains of the PIK3CA gene in breast carcinomas**. *Clin Cancer Res* 2007, 13(20):6064-6069.

153. Barbareschi M, Pelosio P, Caffo O, Buttitta F, Pellegrini S, Barbazza R, Dalla Palma P, Bevilacqua G, Marchetti A: **Cyclin-D1-gene amplification and expression in breast carcinoma: relation with clinicopathologic characteristics and with retinoblastoma gene product, p53 and p21WAF1 immunohistochemical expression**. *Int J Cancer* 1997, 74(2):171-174.

154. Vaziri SA, Tubbs RR, Darlington G, Casey G: **Absence of CCND1 gene amplification in breast tumours of BRCA1 mutation carriers**. *Mol Pathol* 2001, 54(4):259-263.

155. Brown LA, Johnson K, Leung S, Bismar TA, Benitez J, Foulkes WD, Huntsman DG: **Co-amplification of CCND1 and EMSY is associated with an adverse outcome in ER-positive tamoxifen-treated breast cancers**. *Breast Cancer Res Treat* 2010, 121(2):347-354.

156. Ringnér M, Fredlund E, Häkkinen J, Borg Å, Staaf J: **GOBO: Gene expression-based Outcome for Breast cancer Online**. *PLoS One* 2011, 6(3):e17911.

157. **GOBO** [http://co.bmc.lu.se/gobo]

158. Hughes-Davies L, Huntsman D, Ruas M, Fuks F, Bye J, Chin SF, Milner J, Brown LA, Hsu F, Gilks B *et al*: **EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer**. *Cell* 2003, 115(5):523-535.

159. Rodriguez C, Hughes-Davies L, Valles H, Orsetti B, Cuny M, Ursule L, Kouzarides T, Theillet C: **Amplification of the BRCA2 pathway gene EMSY in sporadic breast cancer is related to negative outcome**. *Clin Cancer Res* 2004, 10(17):5785-5791.

160. Kirkegaard T, Nielsen KV, Jensen LB, Campbell FM, Muller S, Tovey SM, Brown S, Cooke TG, Bartlett JM: **Genetic alterations of CCND1 and EMSY in breast cancers**. *Histopathology* 2008, 52(6):698-705.

161. Brown LA, Irving J, Parker R, Kim H, Press JZ, Longacre TA, Chia S, Magliocco A, Makretsov N, Gilks B *et al*: **Amplification of EMSY, a novel oncogene on 11q13, in high grade ovarian surface epithelial carcinomas**. *Gynecol Oncol* 2006, 100(2):264-270.

162. Hu X, Stern HM, Ge L, O'Brien C, Haydu L, Honchell CD, Haverty PM, Peters BA, Wu TD, Amler LC *et al*: **Genetic alterations and oncogenic pathways associated with breast cancer subtypes**. *Mol Cancer Res* 2009, 7(4):511-522.

163. Ronneberg JA, Fleischer T, Solvang HK, Nordgard SH, Edvardsen H, Potapenko I, Nebdal D, Daviaud C, Gut I, Bukholm I *et al*: **Methylation profiling with a panel of cancer related genes: Association with estrogen receptor, TP53 mutation status and expression subtypes in sporadic breast cancer**. *Mol Oncol* 2011, 5(1):61-76.

164. Kamalakaran S, Varadan V, Giercksky Russnes HE, Levy D, Kendall J, Janevski A, Riggs M, Banerjee N, Synnestvedt M, Schlichting E *et al*: **DNA methylation patterns in luminal breast cancers differ from non-luminal subtypes and can identify relapse risk independent of other clinical variables**. *Mol Oncol* 2011, 5(1):77-92.

165. Bediaga NG, Acha-Sagredo A, Guerra I, Viguri A, Albaina C, Ruiz Diaz I, Rezola R, Alberdi MJ, Dopazo J, Montaner D *et al*: **DNA methylation epigenotypes in breast cancer molecular subtypes**. *Breast Cancer Res* 2010, 12(5):R77.

166. Collett K, Eide GE, Arnes J, Stefansson IM, Eide J, Braaten A, Aas T, Otte AP, Akslen LA: **Expression of enhancer of zeste homologue 2 is significantly associated with increased tumor cell proliferation and is a marker of aggressive breast cancer**. *Clin Cancer Res* 2006, 12(4):1168-1174.

167. Puppe J, Drost R, Liu X, Joosse SA, Evers B, Cornelissen-Steijger P, Nederlof P, Yu Q, Jonkers J, van Lohuizen M *et al*: **BRCA1-deficient mammary tumor cells are dependent on EZH2 expression and sensitive to Polycomb Repressive Complex 2-inhibitor 3-deazaneplanocin A**. *Breast Cancer Res* 2009, 11(4):R63.

168. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL *et al*: **Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis**. *Nature* 2010, 464(7291):1071-1076.

169. Tan J, Yang X, Zhuang L, Jiang X, Chen W, Lee PL, Karuturi RK, Tan PB, Liu ET, Yu Q: **Pharmacologic disruption of Polycomb-repressive complex 2-mediated**

gene repression selectively induces apoptosis in cancer cells. *Genes Dev* 2007, 21(9):1050-1063.

170. Bracken AP, Helin K: **Polycomb group proteins: navigators of lineage pathways led astray in cancer**. *Nat Rev Cancer* 2009, 9(11):773-784.

171. Wei Y, Xia W, Zhang Z, Liu J, Wang H, Adsay NV, Albarracin C, Yu D, Abbruzzese JL, Mills GB *et al*: **Loss of trimethylation at lysine 27 of histone H3 is a predictor of poor outcome in breast, ovarian, and pancreatic cancers**. *Mol Carcinog* 2008, 47(9):701-706.

172. Gonzalez ME, Duprie ML, Krueger H, Merajver SD, Ventura AC, Toy KA, Kleer CG: **Histone Methyltransferase EZH2 Induces Akt-Dependent Genomic Instability and BRCA1 Inhibition in Breast Cancer**. *Cancer Res* 2011, 71(6):2360-2370.

173. Holm K, Hegardt C, Staaf J, Vallon-Christersson J, Jonsson G, Olsson H, Borg A, Ringner M: **Molecular subtypes of breast cancer are associated with characteristic DNA methylation patterns**. *Breast Cancer Res* 2010, 12(3):R36.

174. Varambally S, Cao Q, Mani RS, Shankar S, Wang X, Ateeq B, Laxman B, Cao X, Jing X, Ramnarayanan K *et al*: **Genomic loss of microRNA-101 leads to overexpression of histone methyltransferase EZH2 in cancer**. *Science* 2008, 322(5908):1695-1699.

175. Bosco EE, Knudsen ES: **RB in breast cancer: at the crossroads of tumorigenesis and treatment**. *Cell Cycle* 2007, 6(6):667-671.

176. Martin-Subero JI, Kreuz M, Bibikova M, Bentink S, Ammerpohl O, Wickham-Garcia E, Rosolowski M, Richter J, Lopez-Serra L, Ballestar E *et al*: **New insights into the biology and origin of mature aggressive B-cell lymphomas by combined epigenomic, genomic, and transcriptional profiling**. *Blood* 2009, 113(11):2488-2497.

177. O'Riain C, O'Shea DM, Yang Y, Le Dieu R, Gribben JG, Summers K, Yeboah-Afari J, Bhaw-Rosun L, Fleischmann C, Mein CA *et al*: **Array-based DNA methylation profiling in follicular lymphoma**. *Leukemia* 2009, 23(10):1858-1866.

178. Ohm JE, McGarvey KM, Yu X, Cheng L, Schuebel KE, Cope L, Mohammad HP, Chen W, Daniel VC, Yu W *et al*: **A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing.** *Nat Genet* 2007, 39(2):237-242.

179. Widschwendter M, Fiegl H, Egle D, Mueller-Holzner E, Spizzo G, Marth C, Weisenberger DJ, Campan M, Young J, Jacobs I *et al*: **Epigenetic stem cell signature in cancer.** *Nat Genet* 2007, 39(2):157-158.