Chapter

Introduction

1

1.1 Breast cancer incidence and risk factors

With more than one million woman diagnosed annually, breast cancer is the most common cancer among women worldwide. In 2011, 13.987 women were diagnosed with invasive breast cancer in the Netherlands (Integraal Kankercentrum Nederland (IKNL)). One of the greatest challenges faced by clinicians and researchers in this field is that breast cancer is not a single entity, but rather a heterogeneous group of several subtypes displaying distinct differences in biological and clinical behaviour [1,2].

Both genetic and non-genetic factors are involved in the aetiology of breast cancer. Non genetic factors include age, body mass index, alcohol intake and menstrual and reproductive history [3]. A possible genetic contribution to breast cancer risk is indicated by the increased incidence of these cancers among women with a family history of the disease and by the observation of some families in which multiple family members are affected with breast cancer in a pattern fitting that of a dominant cancer inheritance. Hereditary breast cancer accounts for approximately 5 to 10 percent of all breast cancer cases [4]. Clinical markers for hereditary breast cancer are; family history of breast cancer, early age at time of diagnosis, bilateral breast cancer, male breast cancer, multiple primary tumours and tumour phenotype.

1.2 Genetic susceptibility to breast cancer

The genes/genetic variants associated with breast cancer susceptibility can be classified according to the breast cancer risk (relative risk of carriers versus non-carriers) they confer: high-risk genes typically confer a relative risk increase of ≥5 fold, moderate-risk genes are associated with moderate increases in risk (relative risk of 2 to 4) and common low-risk polymorphisms confer smaller increases in risk (relative risk < 1.5).

1.2.1 High risk genes

High risk genes are genetic variants associated with breast cancer that are rare in the population but associated with very high breast cancer risk. Linkage studies in large families with multiple affected individuals (Hereditary Breast and Ovarian Cancer families, HBOC) conducted in the 1990s have led to the discovery that mutations in tumour suppressor genes, *BRCA1* and *BRCA2*, conferred a high risk to breast cancer [5,6].

Inherited germline mutations in the high-risk *BRCA1* or *BRCA2* genes have been identified in about 10 to 20 percent of breast cancer

families. These genes also predispose to ovarian cancer and a substantial fraction of families with breast and ovarian cancer harbour mutations in *BRCA1* or *BRCA2*. The mutations that have been associated with increased risk of cancer result in the absence of functional protein expressed from that allele, supporting the hypothesis that *BRCA1* and *BRCA2* are tumour suppressor genes. The frequency of these mutations may vary in different populations, due to certain founder mutations [7]. For example, in the Ashkenazi Jewish population two deleterious *BRCA1* and one deleterious *BRCA2* founder mutation have been described [8]. In *BRCA1*/*2* associated tumours, loss of heterozygosity of the wild-type allele is frequently observed, resulting in loss of gene function [9].

Bi-allelic germline mutations in *BRCA2* are associated with Fanconi anemia (subgroup D1) and predisposition to childhood tumours [10]. For *BRCA1,* only a single case of bi-allelic pathogenic germline mutations has been reported involving a patient diagnosed with ovarian cancer at age 28. This patient had several developmental abnormalities fitting a Fanconi anemia-like phenotype including short stature, microcephaly, and developmental delay [11].

In a large combined analysis of 22 studies regarding breast and ovarian cancer risk for *BRCA1* and *BRCA2* mutation carriers unselected for family history, the average cumulative risks in *BRCA1*-mutation carriers by age 70 years were estimated to be 65% (95% confidence interval 44%-78%) for breast cancer and 39% (95% CI 18%-54%) for ovarian cancer. The corresponding estimates for *BRCA2* were estimated to be 45% (95% CI 31%-56%) and 11% (95% CI 2.4%-19%) [12].

Other genes conferring a high risk to breast cancer have been identified as part of inherited autosomal dominant cancer syndromes. These include germline mutations in the tumour suppressor genes *TP53*, *PTEN*, *STK11*, and the cell-cell adhesion molecule encoding gene *CDH1* [13,14].

Mutations in *TP53* cause Li-Fraumeni syndrome (LFS), a cancer predisposition syndrome associated with the development of soft tissue sarcomas, osteosarcomas, very early onset breast cancer, brain tumours, adrenocortical carcinoma, and leukemias. LFS-related cancers often occur in childhood or young adulthood and survivors have an increased risk for multiple primary cancers. Breast cancer is the most frequent malignancy among female *TP53* mutation carriers, with approximately 5% of these cases being diagnosed before the age of 30 [15,16].

Germline mutations in the tumour suppressor gene *PTEN* are causative for Cowden syndrome (CS), which is one of the phenotypes included in the *PTEN* hamartoma tumour syndrome (PHTS). CS is a multiple

hamartoma syndrome with a high risk for benign and malignant tumours of the thyroid, breast, and endometrium. The lifetime risk of developing breast cancer is estimated to be as high as 85.2% [17]

Germline mutations in the serine/threonine kinase gene *STK11* are involved in Peutz-Jegher syndrome (PJS). PJS is characterized by the growth of polyps in the gastrointestinal tract, pigmented macules on the skin and mouth, and other neoplasms [13]. In female carriers, the risk for breast cancer by the age of 60 years is estimated to be 32% [18]

CDH1 (E-cadherin) germline mutations have been associated with hereditary diffuse gastric cancer. Patients with germline *CDH1* mutations carry an increased risk of breast cancer and colorectal cancer. Female carriers face a 40%–54% lifetime risk of developing breast cancer [19].

Together, germline mutations in the high-risk breast cancer susceptibility genes are estimated to account for up to 25% of familial breast cancer risk. Despite intensive efforts, genome-wide linkage studies have not discovered other high risk breast cancer genes, suggesting that no further high-risk genes of comparable importance to *BRCA1* and *BRCA2* exist [20]. If additional high-risk genes do exist these are too rare to detect by linkage analysis in unselected familial cohorts.

1.2.2 Moderate risk genes

Another group of genetic variants associated with breast cancer risk are less rare (yet uncommon) variants with a moderate effect on breast cancer risk. These variants have been identified using a candidate gene approach. This approach involves selecting genes based on their known or presumed biological function and searching for variants associated with cancer risk. There is a very large number of genetic epidemiology studies describing associations between various genetic variants and breast cancer risk. A number of these genes have been described in multiple independent studies or are supported by robust meta-analyses. However, there are also genes for which the reported associations do not replicate in follow-up studies or give contradicting results.

A well known example of a moderate risk breast cancer gene is the *CHEK2* gene. A protein-truncating germline mutation, *CHEK2**1100delC, has been shown to increase breast cancer risk by approximately 2-fold [21,22]. *CHEK2**1100delC mutation carriers have an increased risk of bilateral breast cancer [23]. A recent study describing families with homozygous *CHEK2**1100delC mutations showed that women homozygous for the mutation have a 2-fold higher risk of breast cancer when compared to heterozygotes [24]. Furthermore, excess breast cancer risk is reported in first degree relatives of *CHEK2**1100delC positive non-*BRCA1/2* familial

breast cancer patients compared to non-*CHEK2**1100delC familial breast cancer relatives [25].

Inherited mutations in the *ATM* gene are responsible for the autosomal recessive disorder ataxia-telangiectasia. Approximately 0.5% of the general population has been estimated to carry a heterozygous germline mutation in the *ATM* gene [26]. It is shown that the overall relative risk of breast cancer in *ATM* mutation carriers is approximately 2 [27].

Germline mutations in the genes *MRE11*, *RAD50* and *NBN*, together forming the MRE11-RAD50-NBS1 (MRN) complex, are also believed to increase breast cancer risk. The MRN protein complex plays an important role in maintaining genomic integrity. This protein complex integrates DNA repair with checkpoint signalling through the ATM, BRCA1, and CHEK2 proteins. For this reason, a number of studies have screened breast cancer families for pathogenic germline mutations in the genes encoding the MRN complex. The most convincing evidence to confer elevated risk to breast cancer is found for the *NBN* gene. Mutations in this gene confer a 2 to 3-fold elevated risk to breast cancer [28,29].

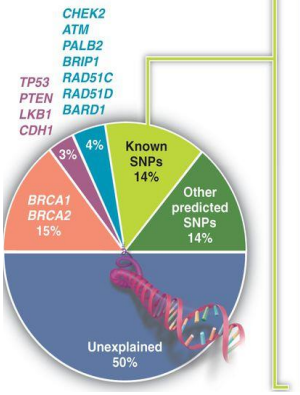
Another group of moderate-risk breast cancer genes are those causative for Fanconi anemia in case of bi-allelic germline mutations. Heterozygous germline mutations in *BRIP1* (*FANCJ*) [30] and *PALB2* (*FANCN*) [31], and possibly *RAD51C* (FANCO) [32] and *SLX4* (FANCP) [33,34] are associated with elevated breast cancer risk.

Finally, mutations in the *RAD51D* [35], *XRCC2* [36,37] and *BARD1* [38,39] genes are believed to confer susceptibility to breast cancer. However, there have been contradicting reports and large studies are needed to confirm the association of these genes with breast cancer risk.

1.2.3 Low-risk polymorphisms

Most of the unexplained genetic susceptibility to breast cancer is likely to be explained by a polygenic model involving a combination of many low-risk variants. The evaluation of common genetic variation in breast cancer has rapidly evolved by the implementation of genome-wide association studies (GWAS) in large cohort studies. GWAS of large numbers of breast cancer patients from the general population along with healthy controls have led to the discovery of more than 70 established breast cancer susceptibility loci [14]. Most of these variants or single nucleotide polymorphisms (SNPs) have very minor effects on breast cancer risk, odds-ratio (OR) <1.20, and how most of these common variants influence breast cancer risk is not understood. There is evidence for a number of SNPs identified by GWAS to be involved in regulatory mechanisms in for example regulation of gene expression. A prime example shows that breast cancer risk associated SNPs

are enriched for FOXA1 and ESR1 transcription factor binding sites, exerting their effects through these pioneer factors in ER+ related breast cancer [40]. Furthermore, a number of low-risk polymorphisms have been shown to influence the risk that high and moderate-risk genes confer [41]. Figure 1 illustrates the contribution of high, moderate and low-risk genes to familial breast cancer risk.



Locus	SNP	Odds ratio
6q14.1	rs17529111	1.97
*BRCA2	rs11571833	1.44
CCND1	rs75915166	1.38
ESR1	rs3757318	1.33
FGFR2	rs2981579	1.33
CCND1	rs554219	1.33
FGFR2	rs2981582	1.26
*MERIT40	rs8170	1.26
TERT	rs10069690	1.24
DNAJC1	rs11814448	1.22
CCDN1	rs614367	1.21
MKL1	rs6001930	1.21
MDM4	rs4245739	1.19
ESR1	rs2046210	1.16
HNF4G	rs2943559	1.14
5p12	rs10941679	1.13
12p13.1	rs12422552	1.13
MAP3K1	rs889312	1.12
TCF7L2	rs7904519	1.12
22q12.2	rs132390	1.12
2p24.1	rs12710696	1.11
2q31.1	rs2016394	1.1
SLC4A7	rs4973768	1.1
EBF1	rs1432679	1.1
FTO	rs11075995	1.1
1p11.2	rs11249433	1.09

Locus	SNP	Odds ratio
8q24	rs13281615	1.09
TOX3	rs3803662	1.09
CDYL2	rs13329835	1.09
1p13.2*	rs11552449	1.08
2q35	rs16857609	1.08
RANBP1	rs204247	1.08
ZMIZ1	rs704010	1.08
12q24	rs1292011	1.08
3p26.2	rs6762644	1.07
8p21.1	rs9693444	1.07
8q24.21	rs11780156	1.07
LSP1	rs3817198	1.07
CCDC88C	rs941764	1.07
TET2	rs9790517	1.06
CDKN2A/B	rs1011970	1.06
TGFBR2	rs12493607	1.05
9q31.2	rs10759243	1.05
DNAJC1	rs7072776	1.05
19q13.31	rs3760982	1.05
RAB3C	rs10472076	1.04
ANKRD16	rs2380205	0.98
*CASP8	rs1045485	0.97
7q35	rs720475	0.96
CHST9	rs1436904	0.96
FTO	rs17817449	0.95
18q11.2	rs527616	0.95

Locus	SNP	Odds ratio
CDCA7	rs1550623	0.94
10q26.12	rs11199914	0.94
COX11	rs6504950	0.94
*SSBP4	rs4808801	0.94
2q35	rs13387042	0.93
PDE4D	rs1353747	0.93
11q13.1	rs3903072	0.92
11q24.3	rs11820646	0.92
PAX9	rs2236007	0.92
RAD51L1	rs999737	0.92
NRIP1	rs2823093	0.92
PEX14	rs616488	0.91
FOXQ1	rs11242675	0.91
8q21.11	rs6472903	0.91
RAD51L1	rs2588809	0.91
LGR6	rs6678914	0.9
NTN4	rs17356907	0.9
2q14.2	rs4849887	0.89
9q31	rs865686	0.89
ZNF365	rs10995190	0.86
ADAM29	rs6828523	0.84
MERIT40	rs2363956	0.82
TERT	rs2736108	0.77
PTHLH	rs10771399	0.72

* Denotes coding variant

Figure 1. Genetic variants that predispose to breast cancer. The pie chart on the left shows the estimated percentage contribution of mutations in high-risk (*BRCA1/2*, *TP53*, *CDH1*, *LKB1*, and *PTEN*) and moderate-risk (e.g., *CHEK2*, *ATM*, and *PALB2*) genes and common low-risk genetic variants to familial relative risk (FRR; ratio of risk of disease for a relative of an affected individual to that for the general population). Common genetic variants are denoted as SNPs. “Known SNPs” are SNPs associated with breast cancer through GWAS, as listed on the right. The odds ratios refer to the increase (or, in some cases, the reduction) in risk conferred by the rare allele of the variants. “Other predicted SNPs” refers to the estimated contribution of all SNPs, other than known loci, that were selected for replication of breast cancer GWAS [42]. "Adapted from [14] Reprinted with permission from AAAS."

1.2.4 Targeted therapy

With the identification of mutated genes that predispose to breast cancer it is hoped that understanding the biology of the affected proteins and the pathways in which they are involved will lead to the development of new "targeted" therapies for patients. Some promising results have been achieved for targeting *BRCA1* and *BRCA2* mutant tumours. The BRCA1 and BRCA2 proteins have critical roles in the repair of double-strand DNA breaks by homologues recombination (HR). As a result, cells lacking functional BRCA1 or BRCA2 are highly sensitive in vitro to particular chemotherapeutic agents, such as cisplatin, and ionizing radiation [43]. Mutations in *BRCA1* and *BRCA2* also sensitize cells to the inhibition of poly(ADP-ribose) polymerase (PARP), an enzyme involved in base excision repair, a key

pathway in the repair of DNA single-strand breaks [44]. This seems to be because the inhibition of PARP leads to the persistence of DNA lesions normally repaired by homologous recombination. However, the exact mechanisms by which PARP inhibitors (PARPi) disrupt tumour growth is unknown. Unfortunately, not all *BRCA1/2* mutation carriers respond to therapies based on PARPi. It has been suggested that mutations in certain domains in the *BRCA1* gene may not confer hypersensitivity to PARPi [45]. In addition, tumours can become resistant to PARPi through various mechanisms [46]. It is hoped that combination therapies can overcome these complications.

1.3 Classification of breast cancer

Breast cancer is not one disease entity, it is a heterogeneous and complex disease encompassing a group of molecularly distinct neoplastic disorders with distinct biological features and clinical outcomes. Much research effort has been, and still is, invested to stratify breast tumours into clinically relevant subgroups with the purpose to select patients for optimal treatment. Breast cancer classification is mostly based on clinico-histopathological, transcriptomic and genomic features of the tumour.

1.3.1 Histopathological classification

According to the elaborate histopathological classification of the World Health Organization (WHO) [47], breast carcinomas can be divided into *in-situ* (benign) and invasive carcinomas. Of the invasive carcinomas, ductal carcinoma of no special type (IDC-NST) is most common representing ~70-80% of all breast cancers. The remaining breast carcinomas are referred to as special types, of which lobular is most common, followed by tubular, papillary and mucinous types.

Currently, treatment choice is based upon two histopathological prognostic classification methods, hormone receptor status and human epidermal growth factor receptor 2 status. The first classification method is the modified Bloom-Richardson grading system, also called the Nottingham system [48,49]. Its histological grade refers to the degree of tumour differentiation, i.e. how closely the tumour resembles its tissue of origin. The overall grade is derived from three tumour characteristics; the proportion showing gland/tubule formation, the degree of nuclear pleomorphism (variability in size, shape and staining of cells and/or their nuclei) and the mitotic count. A grade III tumour is considered to be the least differentiated and the most aggressive breast cancer type. The second method is the TNM

classification, in which T stands for tumour size, N for regional lymph node involvement and M for distant metastasis. The TNM classification is widely used in prognosis of solid tumours in general.

Nowadays, breast carcinomas are routinely scored for oestrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) protein expression using immunohistochemistry (IHC) on tissue sections. Hormone receptor-positive breast cancers (ER, PR) account for around 75–80% of all cases and standardized IHC assays for the routine testing of ER and PR are used to guide the selection of patients for hormone-based therapies. HER2 represents the only additional predictive marker currently in routine use. Approximately 10–15% of breast cancers have HER2 over-expression and/or amplification with around half of these co-expressing hormone receptors. These patients are selected for anti-HER2 based therapies, including the humanized monoclonal HER2 antibody, trastuzumab, which targets the extracellular domain of the HER2 receptor [50]. The remaining 10–15% of breast cancers are defined by hormone receptor and HER2 negativity (i.e., triple negative cancers), which represent a key clinical entity given their lack of targeted therapeutic options [51].

1.3.2 Molecular classification

In the past decade, the development of gene expression profiling using high-throughput microarray-based methods has allowed the simultaneous analysis of the expression level for thousands of genes in a single experiment. It was believed that the information provided by the quantitative assessment of multiple genes would be more precise for biological characterization of breast cancer than that offered by routine histopathology.

Seminal, class discovery, microarray-based gene expression profiling studies by the Stanford group have shown on the molecular level that breast cancer is not a single disease and that distinct molecular subtypes, often with identical histopathological features, exist [52,53]. In these studies, the five main 'intrinsic' breast cancer subtypes were characterized; the Luminal A, Luminal B, HER2-positive, Basal-like and Normal-like subtypes. Importantly, these subtypes were found to be robustly reproducible across tumour cohorts and microarray platforms and to significantly correlate with breast cancer incidence, treatment response and survival [1,54,55].

The luminal subtypes of breast cancer show gene expression signatures resembling that of normal luminal-epithelial cells of the breast. The Luminal A subtype is most common, representing 50-60% of all breast

cancers. These tumours show high expression of ER and ER-related gene networks, are mostly of low histological grade and have a relative good prognosis. Luminal B tumours also express ER but are often of higher histological grade, have higher proliferation rates and a worse prognosis compared to Luminal A tumours. It has been suggested that *BRCA2*-associated breast tumours are often of this more aggressive luminal subtype [56]. The HER2-positive breast cancers are characterized by over-expression of the *HER2* gene (by gene-amplification, mutation or otherwise) and have an aggressive clinical behaviour. A proportion of these tumours is ER+ and often cluster together with Luminal-B tumours. The Basal-like tumours show gene expression signatures similar to that of normal basal/myoepithelial cells of the breast. These basal-like breast carcinomas (BLCs) represent about 10 to 20% of all breast cancers and are generally high grade, triple negative ((ER-), (PR-), and (HER2-)) and express high molecular weight cytokeratins (CK5/6, 14 and 17). Breast tumours of *BRCA1* germline mutation carriers have been shown to be primarily of the Basal-like subtype (75-90%) [53,57]. Finally, the Normal-like subgroup shows expression of genes associated with normal breast/adipose tissue and it has been suggested that this subgroup could be an artefact of tumour sampling.

Recently, new intrinsic subtypes have been identified such as the Claudin-low subtype [58]. This subtype is characterized by low expression of genes involved in tight junctions and intercellular adhesion. This subtype is found to cluster near the Basal-like tumours and have similar histological characteristics. Despite the apparent similarities of Claudin-low tumours with Basal-like tumours, they do not show the same characteristic high expression of genes associated with proliferation. However, they do highly express genes involved in immune system responses [59].

1.3.3 Prognostic gene expression signatures

Gene expression profiling has been extensively used to develop tests that may provide better predictions of clinical outcome than the traditional clinical and pathological standards [60]. Although there are many of these prognostic signatures, there is very little overlap between the actual gene-lists used in each of them. However, it has been shown that the intrinsic subtypes, 70-gene profile, wound response, and recurrence score signatures segregate breast cancers into broadly similar prognostic subgroups and are probably tracking a common set of biologic phenotypes [61].

Two of such signatures are now commercially available to help guide treatment decisions for patients with breast cancer. These are

MammaPrint® (Agendia, The Netherlands) and Oncotype-DX® (Genomic Health, USA). MammaPrint® was developed from the first gene expression signature published on breast cancer. Using a supervised top-down approach, a gene signature was identified from a selected retrospective series of 78 patients with node-negative breast cancer who had received no systemic adjuvant therapy [62]. This assay measures the expression of 70 genes and calculates a prognostic score that categorizes patients into "good" or "poor" risk groups. The U.S. Food and Drug Administration (FDA) recently cleared the assay to aid in formulating a prognosis for patients under 61 years of age who had stage I or II disease with a tumour size of 5 cm or less.

Another molecular assay, Oncotype-DX®, was developed using the candidate-gene approach to estimating outcome [63]. It measures the expression of ER and HER2, as well as that of ER-regulated transcripts and several proliferation-related genes, with the use of the quantitative reverse-transcriptase–polymerase-chain-reaction (RT-PCR) assay. Most of these genes are associated with outcome, and several can be assessed with the use of conventional methods. The Oncotype-DX® system combines these measurements into a quantitative “recurrence score,” which can be used as a continuous variable to estimate the probability of recurrence at 10 years or to group patients into low-risk, intermediate-risk, and high-risk categories.

To evaluate the true prognostic utility of the MammaPrint® and Oncotype-DX® assays, prospective clinical trials are currently ongoing. These are MINDACT (Microarray In Node-negative and 1-3 positive lymph-node Disease may Avoid Chemotherapy) for MammaPrint® and TAILORx (Trial Assigning IndividuaLized Options for Treatment (Rx)) for Oncotype-DX® [64,65]. The outcome of these randomized trials will provide evidence for the prognostic power of both assays and guide their future appliance in daily clinical practice.

On a more critical note, it is suggested that the ability to stratify breast tumours according to clinical outcome by prognostic signatures is perhaps directly correlated to the assessment of proliferation and cell cycle-related genes. As proliferation has been shown to be prognostic in ER-positive disease and not in ER-negative disease (nearly all ER-negative tumours are highly proliferating), this seems to explain why for example both MammaPrint® and Oncotype-DX® assign the high-risk category to almost all ER-negative patients. For this reason these signatures are likely applicable in ER-positive disease only, discriminating the high-proliferating from low-proliferating luminal and HER2-positive tumours [66,67]. Evidence for this hypothesis comes from large meta-analyses of nearly 3000 breast tumours,

for which both gene expression and clinical data is available [66,68,69]. These analyses showed that nine prognostic signatures exhibited a similar prognostic performance in the entire dataset. Their prognostic abilities were found to be due mostly to the detection of proliferation activity. Although ER^--status (Basal-like) and HER2 expression status correspond to bad outcome, they seem to act through elevated expression of proliferation genes and thus contain only indirect information about prognosis. Clinical variables measuring the extent of tumour progression, such as tumour size and nodal status were found to still add independent prognostic information to proliferation genes. From these studies it has become apparent that there are strong connections between traditional (histopathological) prognostic factors, expression-based molecular subtyping, and prognostic signatures, all highlighting the important central role of differentiation and proliferation in breast cancer prognosis and that both genomic and clinical variables should perhaps be included in a common algorithm to yield the most accurate prediction model [60]. Figure 2 shows that the prognostic power of the intrinsic molecular subtypes and several prognostic gene expression signatures are driven through the assessment of genes associated with tumour differentiation and proliferation.

Gene expression profiling has contributed enormously to the understanding of the underlying biology of breast cancer. It has enabled researchers to characterize important genes and pathways that are deregulated in breast cancer, leading to the identification of novel targets for therapeutic intervention. It has also been used to refine breast cancer classification and will continue to do so. Currently, its value in the clinic is largely limited due to the close association with parameters already used within the daily histopathological classification [70].

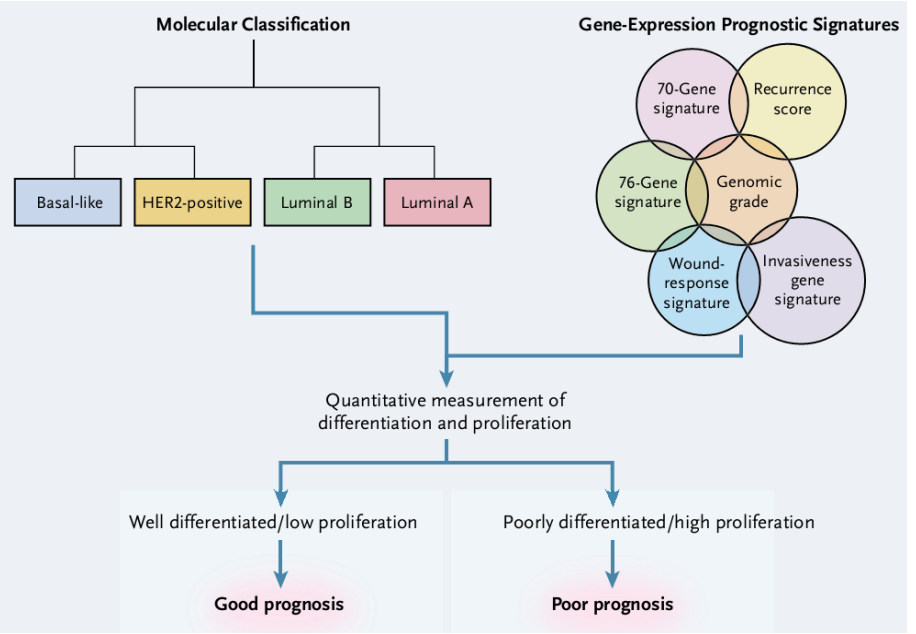


Figure 2. Adapted from [60]. Assessment of proliferation and differentiation drive the prognostic power of the intrinsic subtypes and several gene expression signatures.

1.3.4 Genomic profiling

Another means to stratify breast tumours into subgroups is by genomic copy number and loss of heterozygosity (LOH) profiling. DNA copy number aberrations (CNAs) occur frequently in breast cancer and may define key pathogenetic events, and could also potentially be useful prognostic or predictive factors. Most of the research performed on copy number profiling has been performed with the use of array comparative genomic hybridization (aCGH). This method is restricted to detecting unbalanced structural changes where there is a physical change in copy number of a region of the genome. Other structural rearrangements in the genome such as (balanced) translocations and altered ploidy cannot be identified. Its principals are as follows; test (tumour) and reference (normal genomic) DNA are differentially labelled with green and red fluorescent dyes, mixed in a 1:1 ratio. Competitive hybridization is performed on a platform composed of bacterial artificial chromosomes (BACs) containing sequence-verified, fluorescent in situ hybridization (FISH)-mapped DNA inserts spaced throughout the whole genome in predefined intervals. Hence, the resolution for the identification of genomic gains and losses is determined by the genomic distance between two contiguous BACs [71]. Nowadays, high-resolution single nucleotide

polymorphism- (SNP) arrays are widely used in copy number profiling, with the added possibility for genome-wide LOH detection.

Genome-wide copy number profiling has revealed that distinct patterns of CNAs are associated with different pathological features and gene-expression subtypes of breast cancer [72,73]. For example, the increased CNAs in Basal-like tumours indicates more genetic complexity than in the other subtypes, suggesting a greater degree of genetic instability in these tumours. Basal-like cancers are relatively enriched for low-level copy-number gains involving multiple large chromosomal regions, whereas high-level amplification at any locus is infrequent. In contrast, high-level amplifications are seen more frequently in HER2-positive and Luminal B tumours. Luminal A and especially Normal-like tumours have very few CNAs and may have normal diploid copy number profiles.

Genomic profiling has also been applied to allow discrimination between *BRCA1* and *BRCA2*-mutated and unselected sporadic/hereditary breast tumours. aCGH has been used to identify characteristic genomic aberrations in *BRCA1* and *BRCA2*-mutated breast tumours. The use of such *BRCA1/BRCA2* specific classifiers would facilitate the identification of women and their family members with unknown germline mutation status or undetected germline mutations. They might also prove valuable in assessing the pathogenicity of variants of unknown clinical significance. Furthermore, the identification of driver genes in regions characteristic for *BRCA1* and *BRCA2*-mutated tumours could lead to a better understanding of the underlying process of oncogenesis and may provide novel clues for targeted therapies. Initial studies comparing *BRCA1*-mutated and unselected sporadic breast tumours identified numerous genomic regions with differential CNAs [74-76]. However, recent studies combining gene expression and copy number profiling, have reported few or even no differential regions of genomic aberrations between Basal-like tumours with and without *BRCA1* mutations [72,77].

An integrated analysis of both genomic and transcriptomic data seems particularly powerful to studies tumours in their appropriate biological context. As mentioned earlier, *BRCA1*-mutated tumours are mostly of the Basal-like subtype which has distinct patterns of CNAs as compared to the other subtypes. If *BRCA1*-associated CNAs truly exist, these should be able to discriminate between *BRCA1*-mutated and not-mutated tumours within the Basal-like subtype, avoiding confounding subtype specific CNAs. It should be noted that a proportion of sporadic Basal-like tumours is believed to have a dysfunctional *BRCA1* gene due to alternative mechanisms such as *BRCA1*

promoter hypermethylation resulting in CNAs profiles possibly resembling that of *BRCA1*-mutated tumours [78].

Recently, such an integrative analysis without taking inherited mutations into account was performed on nearly 2000 unselected breast tumours to refine current molecular classification schemes [2]. Clustering analysis of joint copy number and gene expression data revealed 10 novel molecular subgroups. The 10 resulting integrative clusters were each associated with distinct CNAs and gene expression changes. These clusters demonstrated heterogeneity present within tumours classified according to ER, PR and HER2 expression, and they divided all of the previously identified intrinsic subtypes into separate groups. Furthermore, the 10 groups were associated with distinct clinical features and outcomes.

From these results it seems that the future of breast cancer classification will involve multiple levels of assessment incorporating clinical information about the patient, tumour specific information determined by histopathology, and molecular information revealed by genomic, and transcriptomic profiling to provide subtype-specific diagnostic, prognostic and predictive tests. At the genomic level, next generation sequencing will allow the complete genomic landscape of somatic mutations, structural rearrangements, copy number alterations and epigenetic events to be assessed. This will undoubtedly yield a wealth of information as well as add increasing complexity in our attempt to understanding breast cancer biology. Other major research efforts will be focusing on understanding inter-tumour and intra-tumour heterogeneity, which will also require consideration in the interpretation and implementation of molecular classification systems [79]. These approaches will provide the opportunity to understand the molecular events and pathways underpinning subgroups of breast cancer, and potentially allow the identification of the cell of origin or tumour-initiating cell in each subtype. These findings will undoubtedly lead to fundamental advances in our approach to the classification, biological characterization and management of breast cancer [80].

Aims and outline of this thesis

BRCA1-mutated breast carcinomas may have distinct biological features, suggesting the involvement of specific oncogenic pathways in tumour development. The identification of genomic aberrations characteristic for *BRCA1*-mutated breast carcinomas could prove valuable in clinical testing for *BRCA1*-involvement in patients and could lead to a better understanding of the underlying process of oncogenesis.

Chapter 2 describes integrated transcriptomic and genomic analyses to search for *BRCA1*-associated CNAs in a selected group of familial Basal-like breast tumours. This selection was based on an important additional finding in this study. Namely, the observation that a large proportion of breast tumour samples contained varying amounts of tumour infiltrating lymphocytes (TILs). The detrimental effect that TILs have on the ability to reliably measure genomic and expression profiles in tumour samples is largely neglected in literature. Sample selection for low amount of TILs allowed the identification of *BRCA1*-associated CNAs. We have validated these findings on H&E-stained sections of matching tumour material and in 3 publicly available data sets where appropriate. To further substantiate and validate our findings we performed DNA flow cytometry on paraffin-embedded, formalin-fixed, material of *BRCA1*-mutated breast carcinomas selected for large numbers of TILs (> 40% of nuclei). Copy number profiles obtained by shallow whole genome sequencing analysis of sorted tumour cell derived DNA clearly shows profiles that were not detected in the matching unsorted tumour samples. To our knowledge this approach is unique in breast tumour research and the results clearly demonstrate the impact of TILs on copy number profiling.

In chapter 3 we describe copy number and gene expression profiling to investigate whether *CHEK2**1100delC mutated breast cancers harbour characteristic genomic aberrations, as seen for *BRCA1* mutated breast cancers. Literature on genomic profiling of *CHEK2* associated breast tumours is scarce. Furthermore, since *CHEK2* associated tumours are reported to be of the luminal and HER2-positive intrinsic subtypes, our analysis were restricted to these subtypes. The impact of TILs is also assessed during the analyses.

The prognostic impact of the TIL signature identified in chapters II and III is discussed in chapter 4. Cancer related inflammation plays a key role in cancer progression and has been reported to be able to both promote and inhibit tumour growth. In breast cancer the prognostic value of a general tumour inflammatory cell infiltrate is controversial. This can in part be

explained by the use of small heterogeneous patient groups and varying methodologies for assessing tumour cell infiltrates. We used the mRNA immune infiltrate signature as a standardized assessment of general tumour inflammatory cell infiltrate to investigate its association with patient survival using large publically available data-sets of lymph node-negative breast tumours. Also we discuss the association between TILs, survival and breast cancer molecular subtypes.

In chapter 5, the results are shown for whole exome sequencing analysis of germline DNA from five Basal-like BRCAX (familial non-*BRCA1*/*2* mutated) cases in an effort to identify putative novel high or moderate-risk breast cancer genes. These samples were shown in the analysis described in chapter II to share characteristic CNAs and have large overlapping regions of copy neutral LOH on chromosome 17. We hypothesized these BRCAX samples to constitute a genetically more homogeneous group as compared to unselected BRCAX samples. This homogeneity could be caused by pathogenic germline mutations in the same high or moderate-risk breast cancer gene, or alternatively, multiple mutated genes acting in the same pathway. As loss of the wild-type *BRCA1* allele is suggested to be the most common mechanism of inactivation in tumours from patients who carry a deleterious *BRCA1* mutation, the overlapping regions of LOH in the BRCAX samples could hint towards the genomic region(s) harbouring novel breast cancer susceptibility gene(s), making data interpretation a less daunting task.

References

1. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de RM, Jeffrey SS et al.: **Gene expression patterns of breast carcinomas distinguish tumour subclasses with clinical implications.** *Proc Natl Acad Sci U S A* 2001, **98:**10869-10874.

2. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y et al.: **The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups.** *Nature* 2012, **486:**346-352.

3. McPherson K, Steel CM, Dixon JM: **ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics.** *BMJ* 2000, **321:**624-628.

4. **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:**1389-1399.

5. Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W et al.: **A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1.** *Science* 1994, **266:**66-71.

6. Wooster R, Bignell G, Lancaster J, Swift S, Seal S, Mangion J, Collins N, Gregory S, Gumbs C, Micklem G: **Identification of the breast cancer susceptibility gene BRCA2.** *Nature* 1995, **378:**789-792.

7. Mavaddat N, Antoniou AC, Easton DF, Garcia-Closas M: **Genetic susceptibility to breast cancer.** *Mol Oncol* 2010, **4:**174-191.

8. King MC, Marks JH, Mandell JB: **Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2.** *Science* 2003, **302:**643-646.

9. Osorio A, de la HM, Rodriguez-Lopez R, Martinez-Ramirez A, Cazorla A, Granizo JJ, Esteller M, Rivas C, Caldes T, Benitez J: **Loss of heterozygosity analysis at the BRCA loci in tumour samples from patients with familial breast cancer.** *Int J Cancer* 2002, **99:**305-309.

10. Howlett NG, Taniguchi T, Olson S, Cox B, Waisfisz Q, Die-Smulders C, Persky N, Grompe M, Joenje H, Pals G et al.: **Biallelic inactivation of BRCA2 in Fanconi anemia.** *Science* 2002, **297:**606-609.

11. Domchek SM, Tang J, Stopfer J, Lilli DR, Hamel N, Tischkowitz M, Monteiro AN, Messick TE, Powers J, Yonker A et al.: **Biallelic deleterious BRCA1 mutations in a woman with early-onset ovarian cancer.** *Cancer Discov* 2013, **3:**399-405.

12. Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL, Loman N, Olsson H, Johannsson O, Borg A et al.: **Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies.** *Am J Hum Genet* 2003, **72:**1117-1130.

13. Garber JE, Offit K: **Hereditary cancer predisposition syndromes.** *J Clin Oncol* 2005, **23:**276-292.

14. Couch FJ, Nathanson KL, Offit K: **Two decades after BRCA: setting paradigms in personalized cancer care and prevention.** *Science* 2014, **343:**1466-1470.

15. Gonzalez KD, Noltner KA, Buzin CH, Gu D, Wen-Fong CY, Nguyen VQ, Han JH, Lowstuter K, Longmate J, Sommer SS et al.: **Beyond Li Fraumeni Syndrome: clinical characteristics of families with p53 germline mutations.** *J Clin Oncol* 2009, **27:**1250-1256.

16. Chompret A, Brugieres L, Ronsin M, Gardes M, Dessarps-Freichey F, Abel A, Hua D, Ligot L, Dondon MG, Bressac-de Paillerets B et al.: **P53 germline mutations in childhood cancers and cancer risk for carrier individuals.** *Br J Cancer* 2000, **82:**1932-1937.

17. Tan MH, Mester JL, Ngeow J, Rybicki LA, Orloff MS, Eng C: **Lifetime cancer risks in individuals with germline PTEN mutations.** *Clin Cancer Res* 2012, **18:**400-407.

18. Lim W, Olschwang S, Keller JJ, Westerman AM, Menko FH, Boardman LA, Scott RJ, Trimbath J, Giardiello FM, Gruber SB et al.: **Relative frequency and morphology of cancers in STK11 mutation carriers.** *Gastroenterology* 2004, **126:**1788-1794.

19. Kluijt I, Sijmons RH, Hoogerbrugge N, Plukker JT, de Jong D, van Krieken JH, van Hillegersberg R, Ligtenberg M, Bleiker E, Cats A:

Familial gastric cancer: guidelines for diagnosis, treatment and periodic surveillance. *Fam Cancer* 2012, **11:**363-369.

20. Smith P, McGuffog L, Easton DF, Mann GJ, Pupo GM, Newman B, Chenevix-Trench G, Szabo C, Southey M, Renard H et al.: **A genome wide linkage search for breast cancer susceptibility genes.** *Genes Chromosomes Cancer* 2006, **45:**646-655.

21. Meijers-Heijboer H, van den OA, Klijn J, Wasielewski M, de Snoo A, Oldenburg R, Hollestelle A, Houben M, Crepin E, Veghel-Plandsoen M et al.: **Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations.** *Nat Genet* 2002, **31:**55-59.

22. **CHEK2*1100delC and susceptibility to breast cancer: a collaborative analysis involving 10,860 breast cancer cases and 9,065 controls from 10 studies.** *Am J Hum Genet* 2004, **74:**1175-1182.

23. Broeks A, de Witte L, Nooijen A, Huseinovic A, Klijn JG, van Leeuwen FE, Russell NS, van't Veer LJ: **Excess risk for contralateral breast cancer in CHEK2*1100delC germline mutation carriers.** *Breast Cancer Res Treat* 2004, **83:**91-93.

24. Adank MA, Jonker MA, Kluijt I, van Mil SE, Oldenburg RA, Mooi WJ, Hogervorst FB, van den Ouweland AM, Gille JJ, Schmidt MK et al.: **CHEK2*1100delC homozygosity is associated with a high breast cancer risk in women.** *J Med Genet* 2011, **48:**860-863.

25. Adank MA, Verhoef S, Oldenburg RA, Schmidt MK, Hooning MJ, Martens JW, Broeks A, Rookus M, Waisfisz Q, Witte BI et al.: **Excess breast cancer risk in first degree relatives of CHEK2 *1100delC positive familial breast cancer cases.** *Eur J Cancer* 2013, **49:**1993-1999.

26. Broeks A, Urbanus JH, Floore AN, Dahler EC, Klijn JG, Rutgers EJ, Devilee P, Russell NS, van Leeuwen FE, 't Veer LJ: **ATM-heterozygous germline mutations contribute to breast cancer-susceptibility.** *Am J Hum Genet* 2000, **66:**494-500.

27. Thompson D, Duedal S, Kirner J, McGuffog L, Last J, Reiman A, Byrd P, Taylor M, Easton DF: **Cancer risks and mortality in heterozygous ATM mutation carriers.** *J Natl Cancer Inst* 2005, **97:**813-822.

28. Zhang ZH, Yang LS, Huang F, Hao JH, Su PY, Sun YH: **Current evidence on the relationship between two polymorphisms in the NBS1 gene and breast cancer risk: a meta-analysis.** *Asian Pac J Cancer Prev* 2012, **13:**5375-5379.

29. Bogdanova N, Feshchenko S, Schurmann P, Waltes R, Wieland B, Hillemanns P, Rogov YI, Dammann O, Bremer M, Karstens JH et al.: **Nijmegen Breakage Syndrome mutations and risk of breast cancer.** *Int J Cancer* 2008, **122:**802-806.

30. Seal S, Thompson D, Renwick A, Elliott A, Kelly P, Barfoot R, Chagtai T, Jayatilake H, Ahmed M, Spanova K et al.: **Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles.** *Nat Genet* 2006, **38:**1239-1241.

31. Rahman N, Seal S, Thompson D, Kelly P, Renwick A, Elliott A, Reid S, Spanova K, Barfoot R, Chagtai T et al.: **PALB2, which encodes a BRCA2-interacting protein, is a breast cancer susceptibility gene.** *Nat Genet* 2007, **39:**165-167.

32. Meindl A, Hellebrand H, Wiek C, Erven V, Wappenschmidt B, Niederacher D, Freund M, Lichtner P, Hartmann L, Schaal H et al.: **Germline mutations in breast and ovarian cancer pedigrees establish RAD51C as a human cancer susceptibility gene.** *Nat Genet* 2010, **42:**410-414.

33. Fernandez-Rodriguez J, Quiles F, Blanco I, Teule A, Feliubadalo L, Valle JD, Salinas M, Izquierdo A, Darder E, Schindler D et al.: **Analysis of SLX4/FANCP in non-BRCA1/2-mutated breast cancer families.** *BMC Cancer* 2012, **12:**84.

34. Bakker JL, van Mil SE, Crossan G, Sabbaghian N, De Leeneer K, Poppe B, Adank M, Gille H, Verheul H, Meijers-Heijboer H et al.: **Analysis of the novel fanconi anemia gene SLX4/FANCP in familial breast cancer cases.** *Hum Mutat* 2013, **34:**70-73.

35. Osher DJ, De Leeneer K, Michils G, Hamel N, Tomiak E, Poppe B, Leunen K, Legius E, Shuen A, Smith E et al.: **Mutation analysis of RAD51D in non-BRCA1/2 ovarian and breast cancer families.** *Br J Cancer* 2012, **106:**1460-1463.

36. Park DJ, Lesueur F, Nguyen-Dumont T, Pertesi M, Odefrey F, Hammet F, Neuhausen SL, John EM, Andrulis IL, Terry MB et al.: **Rare mutations in XRCC2 increase the risk of breast cancer.** *Am J Hum Genet* 2012, **90:**734-739.

37. Hilbers FS, Wijnen JT, Hoogerbrugge N, Oosterwijk JC, Collee MJ, Peterlongo P, Radice P, Manoukian S, Feroce I, Capra F et al.: **Rare variants in XRCC2 as breast cancer susceptibility alleles.** *J Med Genet* 2012, **49:**618-620.

38. Ratajska M, Antoszewska E, Piskorz A, Brozek I, Borg A, Kusmierek H, Biernat W, Limon J: **Cancer predisposing BARD1 mutations in breast-ovarian cancer families.** *Breast Cancer Res Treat* 2012, **131:**89-97.

39. Spurdle AB, Marquart L, McGuffog L, Healey S, Sinilnikova O, Wan F, Chen X, Beesley J, Singer CF, Dressler AC et al.: **Common genetic variation at BARD1 is not associated with breast cancer risk in BRCA1 or BRCA2 mutation carriers.** *Cancer Epidemiol Biomarkers Prev* 2011, **20:**1032-1038.

40. Cowper-Sal lR, Zhang X, Wright JB, Bailey SD, Cole MD, Eeckhoute J, Moore JH, Lupien M: **Breast cancer risk-associated SNPs modulate the affinity of chromatin for FOXA1 and alter gene expression.** *Nat Genet* 2012, **44:**1191-1198.

41. Antoniou AC, Beesley J, McGuffog L, Sinilnikova OM, Healey S, Neuhausen SL, Ding YC, Rebbeck TR, Weitzel JN, Lynch HT et al.: **Common breast cancer susceptibility alleles and the risk of breast cancer for BRCA1 and BRCA2 mutation carriers: implications for risk prediction.** *Cancer Res* 2010, **70:**9742-9754.

42. Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, Schmidt MK, Chang-Claude J, Bojesen SE, Bolla MK et al.: **Large-scale genotyping identifies 41 new loci associated with breast cancer risk.** *Nat Genet* 2013, **45:**353-2.

43. Bhattacharyya A, Ear US, Koller BH, Weichselbaum RR, Bishop DK: **The breast cancer susceptibility gene BRCA1 is required for subnuclear assembly of Rad51 and survival following treatment with the DNA cross-linking agent cisplatin.** *J Biol Chem* 2000, **275:**23899-23903.

44. Farmer H, McCabe N, Lord CJ, Tutt AN, Johnson DA, Richardson TB, Santarosa M, Dillon KJ, Hickson I, Knights C et al.: **Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy.** *Nature* 2005, **434:**917-921.

45. Drost R, Bouwman P, Rottenberg S, Boon U, Schut E, Klarenbeek S, Klijn C, van dH, I, van der GH, Wientjens E et al.: **BRCA1 RING**

function is essential for tumour suppression but dispensable for therapy resistance. *Cancer Cell* 2011, **20:**797-809.

46. Rosen EM, Pishvaian MJ: **Targeting the BRCA1/2 tumour suppressors.** *Curr Drug Targets* 2014, **15:**17-31.

47. Ellis P: *WHO Classification of tumours. Pathology and Genetics of Tumours of the Breast and Female Genital Organs.* Lyon: Lyon Press; 2003.

48. Elston CW, Ellis IO: **Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up.** *Histopathology* 1991, **19:**403-410.

49. Genestie C, Zafrani B, Asselain B, Fourquet A, Rozan S, Validire P, Vincent-Salomon A, Sastre-Garau X: **Comparison of the prognostic value of Scarff-Bloom-Richardson and Nottingham histological grades in a series of 825 cases of breast cancer: major importance of the mitotic count as a component of both grading systems.** *Anticancer Res* 1998, **18:**571-576.

50. Harris L, Fritsche H, Mennel R, Norton L, Ravdin P, Taube S, Somerfield MR, Hayes DF, Bast RC, Jr.: **American Society of Clinical Oncology 2007 update of recommendations for the use of tumour markers in breast cancer.** *J Clin Oncol* 2007, **25:**5287-5312.

51. Dawson SJ, Provenzano E, Caldas C: **Triple negative breast cancers: clinical and prognostic implications.** *Eur J Cancer* 2009, **45 Suppl 1:**27-40.

52. Perou CM, Sorlie T, Eisen MB, van de RM, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA et al.: **Molecular portraits of human breast tumours.** *Nature* 2000, **406:**747-752.

53. 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 tumour subtypes in independent gene expression data sets.** *Proc Natl Acad Sci U S A* 2003, **100:**8418-8423.

54. 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 tumours are conserved across microarray platforms.** *BMC Genomics* 2006, **7:**96.

55. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, Martiat P, Fox SB, Harris AL, Liu ET: **Breast cancer classification and prognosis based on gene expression profiles from a population-based study.** *Proc Natl Acad Sci U S A* 2003, **100:**10393-10398.

56. Larsen MJ, Kruse TA, Tan Q, Laenkholm AV, Bak M, Lykkesfeldt AE, Sorensen KP, Hansen TV, Ejlertsen B, Gerdes AM et al.: **Classifications within molecular subtypes enables identification of BRCA1/BRCA2 mutation carriers by RNA tumour profiling.** *PLoS One* 2013, **8:**e64268.

57. Foulkes WD, Stefansson IM, Chappuis PO, Begin LR, Goffin JR, Wong N, Trudel M, Akslen LA: **Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer.** *J Natl Cancer Inst* 2003, **95:**1482-1485.

58. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, Rasmussen KE, Jones LP, Assefnia S, Chandrasekharan S et al.: **Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumours.** *Genome Biol* 2007, **8:**R76.

59. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, He X, Perou CM: **Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer.** *Breast Cancer Res* 2010, **12:**R68.

60. Sotiriou C, Pusztai L: **Gene-expression signatures in breast cancer.** *N Engl J Med* 2009, **360:**790-800.

61. Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DS, Nobel AB, van't Veer LJ, Perou CM: **Concordance among gene-expression-based predictors for breast cancer.** *N Engl J Med* 2006, **355:**560-569.

62. 't Veer LJ, Dai H, Van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der KK, Marton MJ, Witteveen AT et al.: **Gene expression profiling predicts clinical outcome of breast cancer.** *Nature* 2002, **415:**530-536.

63. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T et al.: **A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer.** *N Engl J Med* 2004, **351:**2817-2826.

64. Zujewski JA, Kamin L: **Trial assessing individualized options for treatment for breast cancer: the TAILORx trial.** *Future Oncol* 2008, **4:**603-610.

65. Cardoso F, Piccart-Gebhart M, Van't Veer L, Rutgers E: **The MINDACT trial: the first prospective clinical validation of a genomic tool.** *Mol Oncol* 2007, **1:**246-251.

66. Haibe-Kains B, Desmedt C, Sotiriou C, Bontempi G: **A comparative study of survival models for breast cancer prognostication based on microarray data: does a single gene beat them all?** *Bioinformatics* 2008, **24:**2200-2208.

67. Haibe-Kains B, Schroeder M, Bontempi G, Sotiriou C, Quakenbush J. **genefu: Relevant Functions for Gene Expression Analysis, Especially in Breast Cancer.** *R package version 1.12.0.* 2013.

68. Desmedt C, Haibe-Kains B, Wirapati P, Buyse M, Larsimont D, Bontempi G, Delorenzi M, Piccart M, Sotiriou C: **Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes.** *Clin Cancer Res* 2008, **14:**5158-5165.

69. Wirapati P, Sotiriou C, Kunkel S, Farmer P, Pradervand S, Haibe-Kains B, Desmedt C, Ignatiadis M, Sengstag T, Schutz F et al.: **Meta-analysis of gene expression profiles in breast cancer: toward a unified understanding of breast cancer subtyping and prognosis signatures.** *Breast Cancer Res* 2008, **10:**R65.

70. Weigelt B, Reis-Filho JS: **Molecular profiling currently offers no more than tumour morphology and basic immunohistochemistry.** *Breast Cancer Res* 2010, **12 Suppl 4:**S5.

71. Reis-Filho JS, Simpson PT, Gale T, Lakhani SR: **The molecular genetics of breast cancer: the contribution of comparative genomic hybridization.** *Pathol Res Pract* 2005, **201:**713-725.

72. Jonsson 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:**R42.

73. 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:**1033-1040.

74. Jonsson G, Naylor TL, Vallon-Christersson J, Staaf J, Huang J, Ward MR, Greshock JD, Luts L, Olsson H, Rahman N et al.: **Distinct genomic profiles in hereditary breast tumours identified by array-based comparative genomic hybridization.** *Cancer Res* 2005, **65:**7612-7621.

75. Tirkkonen M, Johannsson O, Agnarsson BA, Olsson H, Ingvarsson S, Karhu R, Tanner M, Isola J, Barkardottir RB, Borg A et al.: **Distinct somatic genetic changes associated with tumour progression in carriers of BRCA1 and BRCA2 germ-line mutations.** *Cancer Res* 1997, **57:**1222-1227.

76. Wessels LF, van Welsem T, Hart AA, van't Veer LJ, Reinders MJ, Nederlof PM: **Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumours.** *Cancer Res* 2002, **62:**7110-7117.

77. 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:**661-677.

78. Birgisdottir V, Stefansson OA, Bodvarsdottir SK, Hilmarsdottir H, Jonasson JG, Eyfjord JE: **Epigenetic silencing and deletion of the BRCA1 gene in sporadic breast cancer.** *Breast Cancer Res* 2006, **8:**R38.

79. Caldas C: **Cancer sequencing unravels clonal evolution.** *Nat Biotechnol* 2012, **30:**408-410.

80. Dawson SJ, Rueda OM, Aparicio S, Caldas C: **A new genome-driven integrated classification of breast cancer and its implications.** *EMBO J* 2013, **32:**617-628.